Simulation and optimisation in direct marketing. Part 1: Using simulation models to develop forecasts

Received: 31st May, 2001

P. Bruce McDoniel

Bruce is the Vice-President of Database and Analysis at Summit Marketing Group in St. Louis, Missouri — an integrated marketing service company. Bruce has been in the database marketing industry for over ten years holding positions at GMAC and Mellon Bank. He is currently completing an Executive MBA at Washington University.

J. Patrick Monteleone

Pat is currently the Senior Modeling Analyst at Summit Marketing Group. Pat has 12 years of experience in industry and academia and has held several positions in financial service marketing. Pat also holds a PhD in Management and Decision Sciences from Saint Louis University.

Abstract This study is the first in a two-part series that proposes the use of simulation and optimisation modelling in the direct marketing environment to improve marketing results. The first part of this series outlines the benefits of probabilistic simulations of the business process to improve forecasting information. The second part focuses on how an optimisation model, after simulation of the business process, can produce significant gains in direct marketing efficiency.

The problem addressed in this paper focuses on creating a forecast for a marketing effort, comparing different marketing allocation scenarios and determining which one produces the most revenue per marketing dollar. The study proposes a solution to this direct marketing problem by making contributions in two areas: 1) A forecasting method using Monte Carlo simulation techniques; and 2) an adaptation of return on promotion as the key indicator of marketing results and its use to compare alternative mailing scenarios.

A simulation methodology is proposed and a simulation model is presented. The simulation model is then compared to deterministic methods that use point estimates and evaluated against actual results.

INTRODUCTION

The Monte Carlo methods used today have roots in the Los Alamos National Laboratory's work with thermonuclear weapons and the random neutron diffusion in fissile materials.¹ The scientists who first worked on this simulation project, John Von Neumann

and Stanislaw Ulam, attempted to model a complex situation that still adhered to the laws of physics in order to forecast neutron behaviour.^{2,3} This is similar to what companies are doing today in managing uncertain situations (ie risk analysis and sales forecasts) of different management decisions given a set of

Dr J. Patrick Monteleone

Senior Modeling Analyst,
Summit Marketing Group,
132 East Monroe Avenue,
St Louis, MO 63122-6125,
USA.

Tel: +1 (314) 821 1135;
Fax: +1 (314) 821 7511;
e-mail: Patrick.Monteleone@
summitmg.com

Figure 1 Observed distributions

historically measured results (ie revenue per sale, cost per sale).

As shown later in this study, deterministic methods, until recently, were used to generate point estimates as the basis for the forecast (ie average revenue per sale, average number of sales). Point estimates, however, do not consider variation or the distribution around the point estimate (See Figure 1). While the sample mean represents an accurate measure of central tendency, it does not account for dispersion, skew, or kurtosis of the variable. Forecasts using just means would not consider these other elements of the distribution. Instead of using deterministic measures for forecasting, this study recommends using simulation methods that account for central tendency and dispersion.

Marketing allocation decisions are made in an environment of uncertainty. Monte Carlo simulation provides a means of managing the risk probability associated with decision making that deterministic methods do not. The organisation makes judgments based on uncertainty and limited knowledge of the probabilities of certain outcomes. Using Monte Carlo simulation, a portion of the business model can be modelled, as a system, complete with consideration of uncertainties.

Kautt and Hopewell reinforce the use

of probabilistic measures over deterministic measures in their use of simulation in the financial services industry saying that deterministic forecasts are 'becoming passe for planning and modeling purposes'. The authors further add that deterministic models are not wrong but inaccurate.⁴

In a related article, Fogarty and Goldwater highlight the nature of accounting research using deterministic simulations as the primary research tool. The authors note that deterministic simulation, although popular, can be improved by incorporating elements that have a probabilistic nature that represent the variation and uncertainty in any system. Monte Carlo simulation 'represents the uncertain relationships to predict system output'.⁵

Prior to 1992, the Environmental Protection Agency (EPA) used single-point estimates of risk that provided minimal information about variation of the risk assessment. After 1992, the EPA recommended developing 'multiple descriptors' of risk and to allow factor variation. The EPA described Monte Carlo simulation as 'a highly effective way to produce these multiple risk descriptors'.⁶

Perhaps the best summary of the use of simulation comes from a 1999

simulation and gaming article that states, 'business simulations provide the researcher with a laboratory within which a wide range of issues can be explored under different scenarios'.⁷

SIMULATION IN DIRECT MARKETING

Simulation has been used extensively in the marketing arena but infrequently within direct marketing. Research in Monte Carlo simulation can be found in the areas of consumer behaviour models,^{8,9} consideration set-brand choice models and the impact of marketing mix,¹⁰ risk reduction in marketing strategy building,¹¹ and the impact of network effects on market share.¹² While Alan Rosenspan in a 1999 article emphasises the need for risk reduction as a major objective of direct mail testing,¹³ no publications were found that discuss the use of simulation as a risk reduction tool. This study fills that gap by exploring the use of simulation to decide among competing allocation decisions.

SIMULATION ILLUSTRATED

Table 1 shows a simple example in which several lists were tested promoting multiple products. (The creative package, offer, and seasonality are assumed to remain constant for this illustration.)

Determining the optimal blend across lists to maximise the next 100,000 piece mailing becomes difficult. List D performs well across all product lines in terms of sales rate. List C performs poorly with Product 1 but has high average revenue. List C also has 80,000 names available for the next mailing whereas List D only has 25,000. The best allocation decision can be overlooked using judgment alone.

As mentioned earlier, non-quantitative

allocation methods frequently enter into the decision process. Quantitative simulation models can be used to test allocation decisions and their outcomes in a low-cost, controlled, inexpensive laboratory of the computer, removing non-quantitative methods such as judgment and politics.

The example above only illustrated two inputs that were allowed to vary — list and product. In a situation with many more variables, the forecast results become even more complicated to estimate. The interaction between a large group of variables prohibits large-scale experimentation or testing of allocation decisions.

With historical data and a simulation model, a range of decision outcomes can be generated without the marketing cost of testing the inputs that contribute to the outcomes. Law and Kelton recommend using a systems approach for the development of a simulation model.¹⁴ The system approach that was selected for use in this study was open systems theory described by Katz and Kahn.¹⁵ With a foundation in social science, open systems theory is a general systems theory that is easily adaptable for use in a systems approach with simulation models.

The key feature of open systems theory is the relationship between a system and its supporting environment. From a consistency point of view, the environment the system operates in must be continually considered when evaluating the operation of the system. The different aspects of the environment impact the system in varying degrees. By allowing environmental factors to vary in the model, the impact these variations have on the business model can be measured. In turn, understanding the relationships and their strengths allows decisions to be made under circumstances that reflect the market environment.

Table 1: Marketing results

List	Quantity mailed	Quantity available	Sales rate				Average revenue product			
			Product 1 %	Product 2 %	Product 3 %	Product 4 %	Product 1 \$	Product 2 \$	Product 3 \$	Product 4 \$
A	25,000	50,000	0.25	0.30	0.30	0.20	35	30	30	25
B	30,000	60,000	0.30	0.20	0.40	0.10	30	30	25	30
C	15,000	80,000	0.10	0.40	0.20	0.20	40	30	35	35
D	20,000	25,000	0.50	0.34	0.10	0.30	30	40	30	35
E	10,000	75,000	0.40	0.20	0.30	0.20	30	35	30	35
Totals	100,000	290,000	0.31	0.28	0.28	0.19	31	33	28	32

In its simplest form, an equation can be created based on open systems theory usable in a simulation model:^{16,17}

Input → Transformation → Output

The input variables are variables the organisation controls. Typical direct marketing examples are marketing budget, mailing cost, mail quantity, products to be sold, premiums offered and product price. Prior to conducting a marketing campaign, a firm will allocate a budget and, given a cost per thousand (CPM), determine how much to mail. The firm will also decide which product(s) to market, the offer (ie '10% off', 'Buy one get one free', entry into sweepstakes) and the product price. These are the decisions a firm makes prior to conducting a marketing campaign and are controllable by the firm itself.

Transformation variables are the forces within a simulation model that cannot be determined by the organisation but are a result of its business process and must be measured. Typical transformation variables are response, sales and profitability metrics such as response rate, sales rate and revenue per sale. The development of the transformation variables is largely dependent on historical data and is used to forecast future rates for the next marketing effort. These estimates are the heart of the simulation model. The

simulation model incorporates the transformation variables and allows them to vary as they do in reality, based on probability distributions. In the case of the Monte Carlo simulation model, risk is contained in the variability of these environmental factors.

Table 2 shows an example of a sample data set. Using normal, deterministic methods, a forecasting model would use a mean of 2.5 per cent as an estimate of the response rate. Looking at the plot of these values in Figure 2, however, it can be seen that these rates range from a low of 2.10 per cent to 2.99 per cent having a standard deviation of 0.3 per cent and a 95 per cent confidence interval between 1.9 per cent and 3.1 per cent — a large variation. Using the simple mean would not adequately represent this distribution.

To use a distribution to represent variation of a variable in a simulation model, it is critical that distributions or variables selected have been tested to support the assumption that they vary in a manner that is consistent with how the variables behave in reality. To test a variable's distribution against known distributions, chi-squared is a good generalised goodness-of-fit test to test whether significant differences exist between the two distributions. For the data in this illustration, a normal distribution was selected since it had the

Table 2: Sample data set

Observation	Response rate %	Observation	Response rate %
1	2.62	12	2.49
2	2.20	13	2.71
3	2.65	14	2.46
4	2.99	15	2.56
5	2.35	16	2.34
6	2.77	17	2.10
7	2.39	18	2.92
8	2.76	19	2.48
9	2.27	Mean	2.5
10	2.85	Standard deviation	0.3
11	2.56

Figure 2 Distribution by response rate

highest p -value in detecting no significant difference between the distributions.

The output variables are the result of the interaction between the input and transformation variables. The output variables are typically financial or response performance indicators. Some examples are number of sales, dollar volume of sales, profit and risk distribution.

In addition to receiving point estimates of output variables, simulation also produces variances (something deterministic models do not do). With variances a company can then understand the range of possible outcomes, calculate the probability of

achieving target goals, calculate expected values, and also generate a relative risk indicator. For example: if the company has a sales goal of 5,000 units, simulation forecasts assign a probability to achieving that estimate. So the $P(\text{sales} \geq 5,000) = 40\%$.

VALIDATING SIMULATION RESULTS

Recalling, Law and Kelton suggest validating a simulation model using three techniques.¹⁸ These techniques are:

- checking for high face validity
- empirically testing model assumptions
- testing distributions of the simulation output.

Checking for high face validity: the first step can be considered as a 'common sense test'. This step reviews, with the business process experts, model assumptions in determining whether or not the model is reasonable and credible. This process entails effectively quantifying observations from the actual system while gaining model acceptance from key people.

Empirically testing model assumptions: the second step entails quantitative testing of the assumptions made in the face validity phase of validation. When distributions are fitted to variables in the model, the distributions must be empirically tested to support the goodness-of-fit assumption. Kolmogorov-Smirnoff (K-S) and chi-squared tests can be used to test for differences between observed data and selected distributions. Sensitivity analysis is also useful in measuring the relationship between output variables and the variation of input parameters.

Testing distributions of the simulation output: the last step of simulation model validation is the most definitive of the three steps. Output data distributions are tested to establish that the output is reflective of actual output from the system being simulated. If the simulated output and the actual output are not significantly different, then the system can be considered reasonably valid. A variety of techniques exist to test the final output of the simulation, including stationarity tests, tests for auto-correlation, among other standard tests for differences between distributions.

For the simulation model discussed later in the paper, Law and Kelton's techniques were applied in the following manner. For checking for high face validity a log and 'check-off' list was prepared so the development process could move forward only after all model

assumptions were agreed upon. Upon completion of the model, a complete overview of the model was conducted by analysing historical results to establish model credibility. For empirically testing model assumptions, distributions were determined for marketing metrics by using historical results. Distributions were selected that accurately reflect variation and were supported by K-S and chi-squared tests. For testing distributions of the simulation output most output distributions were tested against their historical distributions. Two input variables, response cost and mailing cost, did not have distributions as these were fixed estimates. As a result, output variables including cost estimates did not have directly testable distributions. Credibility and reasonability of these variables were supported through discussions with the business experts.

RETURN ON PROMOTION

This study proposes the use of a modified indicator called return on promotion index (ROPI) that is an extension of Shepard's return on promotion illustrated in his book.¹⁹ (See formula 1).

Formula 1:

$$\begin{aligned}\text{ROPI} &= (\text{Gross Revenue} - \text{Variable} \\ &\quad \text{Expenses} - \text{Premium}) / \\ &\quad \text{Marketing Cost} \\ &= \text{Net Contribution} / \text{Marketing} \\ &\quad \text{Cost}\end{aligned}$$

Where gross revenue is defined as the revenue generated by product sales, variable expenses are defined as the direct and indirect expenses generated as a result of response and sales activity (for example commissions, processing costs, fulfilment costs, cost of goods), premium is defined as the discount used to induce prospects to respond and marketing cost

Table 3: ROPI illustrated

Item	Marketing Project 1 \$	Marketing Project 2 \$
Gross revenue	100,000	150,000
Variable expenses	70,000	90,000
Premium	10,000	15,000
Contribution	20,000	45,000
Mailing cost	25,000	30,000
ROPI	0.80	1.50

is the cost of the direct marketing programme (such as the cost of the mailing). Marketing cost is viewed as fixed, therefore any project with a profitability index greater than one is acceptable and the most efficient project has the highest ROPI.²⁰ Table 3 illustrates an example of the calculation of ROPI.

Marketing Project 1 with a ROPI = 0.80 would be rejected as its contribution is less than the cost of the mailing. Marketing Project 2 with a ROPI of 1.50 would be acceptable as it is producing a positive contribution (\$45,000) to reduce fixed costs (such as interest expense, rental expense) and add to corporate profits.²¹

APPLIED SIMULATION ... A CASE STUDY

In this section, the paper explores a business process simulation and the need

to develop a model to measure and quantify an acquisition business process. The simulation model was undertaken to develop better forecasting methods and evaluate multiple mailing allocations across multiple variables: variable pricing structure, response model's decile performance, fixed marketing budget, one product and no response premium. The goal was to achieve a ROPI of at least \$1.15.

The input variables were mailing quantity and mailing CPM. There were a total of 127 transformation variables that had to be forecast. This included response rates by state by response decile, sales rates by state by response decile, response costs, sales costs and many other measurable steps in the business process. See Table 4 for the output results.

The deterministic and simulation forecasts gave nearly identical results, a not unsurprising result, since the same means used in the deterministic forecast were used in the simulated forecast. The simulation, however, also produced measures of dispersion that are not present in a deterministic method. (Actual September results are also displayed for illustrative purposes.)

The probability of achieving the \$1.15 ROPI, using the mean and standard deviations from the simulation forecast, was calculated as $p = 0.76$ — giving the

Table 4: Actual performance vs. deterministic and simulation forecast

Output measure	September actual \$	Deterministic forecast \$	Simulation forecast	
			Mean \$	Standard deviation \$
Gross revenue	235,000	250,000	262,500	15,264
Variable expenses	115,000	125,000	140,000	7,896
Premium	0	0	0	n/a
Contribution	120,000	125,000	122,500	7,368
Mailing cost	100,000	100,000	100,000	n/a
ROPI	1.20	1.25	1.23	0.11

Numbers have been weighted to maintain client confidentiality. Illustrated results are consistent with actuals.

marketing programme a moderate probability of success.

CONCLUSIONS

Using simulation forecasting techniques to project direct marketing results generates more useful information for evaluating alternative allocation options. Primarily, it allows for probabilistic estimates of reaching targeted goals and allows for sensitivity analysis in testing 'what if' scenarios. It also reduces the high test of costing by comparing different results before 'rolling-out'.

Simulation forecasting produces many output variables that allow companies to choose the best course of action depending on corporate goals — such as maximising sales, return on investment (ROI) or revenue. The authors recommend taking a financial approach to using simulation forecasts and choosing the marketing option that produces the best financial evaluation metric — ROPI, ROI, EVA or other appropriate measure of return on investment.

Prior to creating any forecasting model, however, an internal study needs to be undertaken to develop metrics to monitor the business process. This includes metrics such as response rates, sales rates, 'no-pays', returns, retention rates and average sales. Cost metrics, including direct and indirect costs, should be developed to quantify the cost of each processing step. This will allow the company to produce quality simulated forecasts of the impact the marketing effort will have on its key performance indicators.

As the simulation model was developed, the understanding of the business process also improved. The simulation model allowed future marketing efforts to be more aligned with corporate profitability metrics than

marketing response and sales metrics. Even if a company undertakes the process to develop a simulation model and then decides not to deploy it, the company has accumulated information across many departments and will have gained valuable information regarding costs and reporting structures.

As direct marketers, there may be a tendency to focus on the more obvious parts of the direct marketing mix such as designing and creating new packages, choosing lists, managing production and measuring campaign results. It is vital to remember, however, that marketing is the driver for all new customers and the driver for all revenues and costs. So understanding how a new marketing programme impacts financial results is key to understanding its additive (or reductive) effect on profitability. The development of a simulation model not only helps forecast results but also produces probability estimates of achieving marketing goals.

Simulation models integrate well with predictive models. In many cases this can produce eye-opening results. For example, companies typically target prospects that are the most responsive by mailing only those who are most likely to respond. The most responsive prospects are not necessarily, however, the most profitable target market due to such factors as no-pays, returns and retention rates. In one simulation model, response deciles three, four and five (out of ten deciles) proved to be the most profitable.

With information about customer profitability (such as lifetime value calculations) and attrition probability in addition to acquisition metrics, simulation models can help marketing managers compare different marketing allocation mixes between acquisition and retention strategies. This allows managers to determine various revenue generating

capabilities for a fixed marketing budget.

While simulation excels in reviewing different business options, it can be improved further. Since simulation allows companies to project the results of different business options (such as marketing allocation mix and the budget mix between acquisition and retention strategies), should it not also be able to determine the optimal business option — the one that generates the highest ROI, ROPI or EVA? Every direct marketing programme juggles among lists, creative package, offers, products, timing and many other factors to achieve the 'best' mix. Part 2 of this study, 'Optimised marketing', focuses on just this — allocating scarce marketing dollars considering all the different marketing options to achieve the largest contribution to fixed expenses and profits.

Limitations and areas for improvement

A simulation model is only as good as the forecasting methodology, the quality of the business process measurements and the cost of relevant activities in the business process. Anything affecting the collection of information weakens the ability of the simulation model to weigh alternative options accurately.

While it is an added benefit to make the simulation an accurate forecast of future results, what is ultimately important is to develop a relative ranking of the different marketing scenarios — list, response decile, creative, geographic area, product and anything else that factors into the marketing mix decision. This is an important distinction to make considering the difficulty of projecting seasonality effects inherent in most direct marketed products and the amount of creative and list testing that occurs within the industry.

References

- 1 Pillana, S. (2000) 'History of Monte Carlo method', available from the World Wide Web at <http://stud4.tuwien.ac.at/~e9527412/history.html>.
- 2 *Ibid.*
- 3 Veach, E. (1997) 'Robust Monte Carlo methods for light transportation simulation', PhD dissertation, Stanford University, December.
- 4 Kautt, G. and Hopewell, L. (2000) 'Modeling the future', *Journal of Financial Planning*, Vol. 13, No. 10, pp. 90–100.
- 5 Fogarty, T. and Goldwater, P. (1995) 'Overcoming the false certainty of accounting education: The use of Monte Carlo simulation by accounting students', *Journal of Accounting and Computers*, Vol. 10.
- 6 Environmental Protection Agency (1994) 'Use of Monte Carlo simulation in risk assessments', 'Technical Guidance Manual: Risk Assessment'.
- 7 Neal, D. (1999) 'How consistent are winning strategies? The role of competitor analysis and budgets on performance in a simulation', *Simulation & Gaming*, Vol. 30, Issue 2, p. 118, June.
- 8 McConnell, A., Niedermeier, K., Leibold, J., El-Alayli, A. *et al.* (2000) 'What if I find it cheaper somewhere else? Role of prefactual thinking and anticipated regret in consumer behavior', *Psychology & Marketing*, Vol. 17, No. 4, pp. 281–298, April.
- 9 Orme, B. and Huber, J. (2000) 'Improving the value of conjoint simulations', *Marketing Research*, Vol. 12, No. 4, pp. 12–20, Winter.
- 10 Chiang, J., Chib, S. and Narasimhan, C. (1999) 'Markov chain Monte Carlo and models of consideration set and parameter heterogeneity', *Journal of Econometrics*, Vol. 89, No. 1, pp. 223–248, March/April.
- 11 Walker, J. (1999) 'Taking the guesswork out of strategy building', *Marketing Week*, Vol. 22, No. 7, p. 14.
- 12 Clark, B. and Chatterjee, S. (1999) 'The evolution of dominant market shares', *Journal of Marketing Theory and Practice*, Vol. 7, No. 2, pp. 83–96.
- 13 Rosenspan, A. (1999) 'The science of marketing', *Direct Marketing*, Vol. 62, No. 8, pp. 64–65, December.
- 14 Law, A. and Kelton, W. (1991) 'Simulation Modeling and Analysis', McGraw-Hill, New York.
- 15 Katz, D. and Kahn, R. L. (1978) 'The social psychology of organizations', 2nd ed., John Wiley and Sons, New York.
- 16 *Ibid.*
- 17 Hughes, A. M. (2000) 'Strategic database marketing: The masterplan for starting and managing a profitable, customer-based marketing program', 2nd ed., McGraw-Hill, New York.
- 18 Law and Kelton (1991) *op. cit.*

- 19 Shepard, D. (1999) 'The new direct marketing: How to implement a profit-driven database marketing strategy', 3rd ed., McGraw-Hill, New York.
- 20 Shapiro, A. C. and Balbirer, S. D. (1999) 'Modern corporate finance', 1st ed., Prentice Hall, New Jersey.
- 21 When comparing projects of equal cost the ROP and the ROPI will yield identical marketing project decisions.