



Stochastic Solutions

Generating Incremental Sales

Maximizing the incremental impact of cross-selling,
up-selling and deep-selling through uplift modelling

Abstract

There is a subtle but important difference between

targeting people who are likely to buy if they are included in a campaign

and

targeting people who are *only* likely to buy if they are included in a campaign.

It transpires that this single-word distinction is often the difference between a strongly profitable and a severely loss-making campaign. We have seen many cases in which moving to targeting on the second basis (for *incremental* sales) has more than doubled the extra sales generated by a campaign. Conventional “response” models—despite their name—target on the former basis, and have a marked tendency to concentrate on people who would have bought anyway, thus misallocating marketing resources by increasing costs and failing to maximize sales. This paper discusses the use of a radical new type of predictive modelling—*uplift* modelling—that allows campaigns to be targeted on the second basis, i.e. so as to maximize incremental sales from cross-sell, up-sell and other sales-generation campaigns.

Management Implications

- Measuring incremental sales, while essential, is not enough: the goal is to *maximize* incremental sales.
- Traditional “response” models have a strong tendency to direct resources towards customers who would have bought anyway; this often results in comparatively few incremental sales.
- The customers who spend most after being subject to a marketing intervention are *not* necessarily the ones whose spending increases most as a result of that intervention.
- Because campaigns are often limited by volume (e.g. the top 20% of the customer base by score is mailed), the price of mistargeting is not merely an increase in the cost of each incremental sale but also a reduction in the total volume and value of incremental sales (and therefore of total sales).
- Particular care must be taken when assessing and optimizing the financial impact of incentive-based campaigns. While there may be collateral benefits, offering an incentive to a customer who would have bought anyway has a double cost—the contact cost and the (unnecessary) incentive cost.
- Response codes, while useful indicators, do not actually prove incremental impact and can be misleading; this is particularly true when they are associated with an incentive.
- Businesses not currently using rigorous control groups should adopt them as a matter of priority and always report the incremental impact (uplift) of initiatives.
- Businesses using control groups should consider adopting uplift modelling to drive greater incremental sales and significantly enhanced campaign profitability and effectiveness.

“ *Measuring incremental sales is not enough. The goal is to maximise incremental sales* ”

1. Introduction

It is often said that success has many parents, but failure is always an orphan. Nowhere is this more true than in sales and marketing, where any sale will typically be claimed by at least a product team, an advertising team and a direct marketing team, while missed targets will invariably be blamed by each function on the others. In truth, of course, outcomes have multiple causes, but this does not diminish the importance of correctly assessing and maximizing the contribution of each.

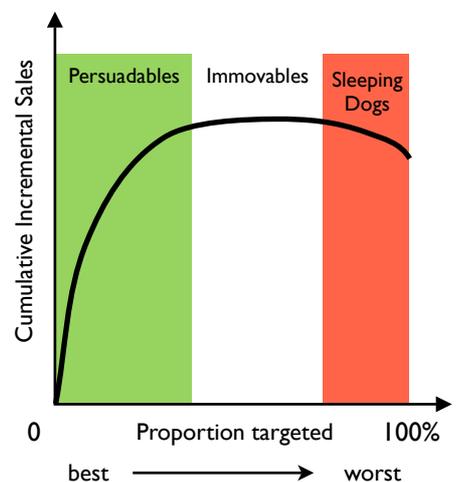
One of the strengths of direct marketing is that it is possible to prove with great confidence that a particular campaign has generated a certain level of incremental business. The key to such proof is the systematic use of control groups. However, while control groups now form an essential part of best practice in direct marketing, their use is almost always limited to *assessing* incrementality after the event. The propensity models normally used for targeting are not built with the goal of maximizing this incrementality. This paper discusses the benefits that can be gained by switching to modelling approaches that are designed with the specific goal of maximizing the incremental sales generated by a cross-sell, up-sell or deep-sell¹ campaign.

2. Incremental Impact or Uplift

Different individuals naturally react differently to any given marketing campaign. For simplicity, consider the case of a campaign that aims to sell a single unit of a particular product or service to people. In general, there will be some people for whom the campaign will be effective, meaning that their probability of purchase will increase if they are included in the campaign; we call such people “Persuadables”. There will others for whom the campaign has little or no impact; we call these “Immovables”. Finally, we have to accept at least the logical possibility that there will some people for whom the campaign will have a negative effect, i.e. they will be less likely to purchase the product in question as a result of the intervention; we call such people “Sleeping Dogs”. As we shall see later, compelling evidence shows that such negative effects do exist, and are more frequent than we might expect, across a range of marketing campaigns.

If we measure the incremental sales generated by the campaign at different targeting volumes, starting with those people most positively influenced and working our way down the customer base to those most negatively affected, the result is the graph to the right—the “Italian Flag Diagram”. If our goal is to maximize incremental sales at minimum cost, it is clear that we should ideally target only the Persuadables segment, as money spent on the Immovables has no meaningful impact on sales and money spent on Sleeping Dogs, when they exist, is counterproductive.

¹ Whereas cross-selling aims to sell an extra product to a customer, and up-selling attempts to persuade a customer to upgrade a product or service, *deep-selling* simply aims to increase the frequency or quantity of a customer's purchases.



3. The Fundamental Campaign Segmentation

We now examine propensity models and consider how they relate to the Italian Flag diagram.

It is obviously the case that each of our customers either will or will not buy the product in question if we do not treat them (i.e. do not include them in our campaign). Equally, it must be the case that each customer will either buy or not buy that same product (during our chosen outcome period) if we do treat them. We cannot directly measure both of these outcomes for any customer, because we cannot simultaneously treat and not treat an individual; however, in principle, everyone has a position in the segmentation shown to the right, which we call the *Fundamental Campaign Segmentation*. In effect, this splits our previous “Immovables” segment into two sub-segments—the “Sure Things”, who purchase whether we include them in the campaign or not, and the “Lost Causes”, who do not purchase in either case. Again, in the ideal world, if the goal of our cross-sell campaign is to generate as many extra sales of our product as possible, at minimum cost, it is clear that we should target only the Persuadables segment.

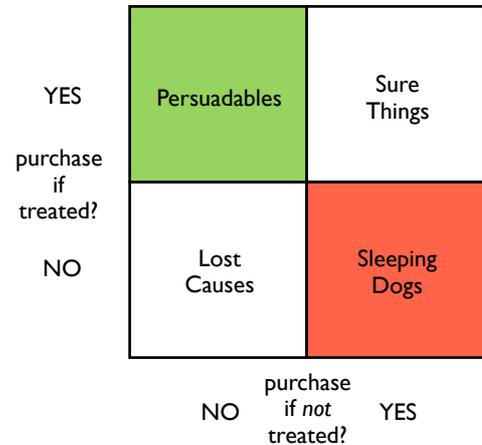
In reality, the most we might reasonably expect to be able to do is to assess the *propensity* of customers to buy in the two cases. So it is interesting to look at a “soft” version of the Fundamental Segmentation, which corresponds more closely to the Italian Flag Diagram we introduced earlier. Now, instead of binary outcomes, we consider the probability of purchase for each customer in the two scenarios—treated and non-treated. Again, the Persuadables segment is at the top left because Persuadables have a significantly higher probability of purchase when treated. Conversely, the Sleeping Dogs are at the bottom right because their probability of purchase is materially reduced by inclusion in the campaign. Thus incremental impact, or *uplift*, increases from bottom right to top left: the higher this is for an individual, the higher is the probability of an incremental sale as a result of treatment.

Despite their name, so-called “response” models do not, in general, predict incremental sales (uplift), as we shall see in section 5. Before that, however, we need briefly to review control groups.

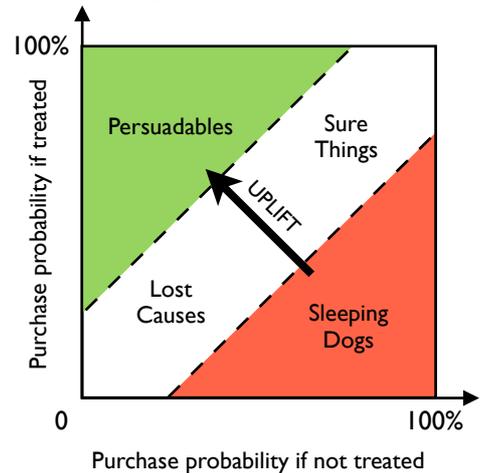
4. Control Groups

Control groups are now widely recognized as being crucial for measuring the true impact of targeted marketing and form a standard part of best practice. The most widely used control group is the Mailing Control Group²—a randomly chosen set of people who meet the campaign selection criteria but are deliberately omitted from the campaign to allow measurement of the incremental impact of that campaign. Provided that the only systematic difference between the controls and the treated group is the treatment, and that both are of adequate size, a mailing control group allows accurate assessment of the extra sales generated by a campaign during whatever out-

The Fundamental Campaign Segmentation (“Hard”)



The Fundamental Campaign Segmentation (“Soft”)



- Persuadables** are more likely to purchase if treated
- Sure Things** are likely to buy whether treated or not
- Lost Causes** are unlikely to buy whether treated or not
- Sleeping Dogs** are discouraged from buying by treatment

² More generally, we call these Campaign Controls, or Treatment Controls, because they are used to allow assessment of the change in behaviour that results from a campaign or treatment. Other kinds of controls are Targeting Controls and Fallows. Targeting controls are people who do not meet the targeting criteria, but who are treated to allow assessment of the quality of targeting. Fallows are customers excluded from all (or at least a range of) marketing activity to allow assessment of the impact on behaviour of a complete marketing programme.

come period is chosen. It is worth noting that mailing controls are quite capable of isolating the effect of a single marketing action, whatever other activity is happening at the same time provided, again, that they are genuinely chosen at random from the initial target population.

In the case of a binary outcome such as a purchase, we define the uplift simply as the difference between the purchase rate in the treated group and that in the control group. For example, if the purchase rate in the treated group is 1.7% and the purchase rate in the control group is 1.3%, the uplift is 0.4 percentage points. If the value of sales is being considered, we would normally look at the difference in spend per head of population, so that if the average purchase size in the treated group were €11.70 and in the control group were €10.20, this would represent an uplift of €1.50 per head.

5. “Response” Models and Penetration Models

We repeatedly enclose the “response” of “response models” in quotation marks not to belittle them, but because the term “response” carries strong connotations of causality. Thus “response” models sound as if they estimate outcomes *caused* by the intervention in question, which they do not. In fact, that is precisely what uplift models do. Even when a mailing control group has been used in a campaign, the almost universal approach with “response” modelling is to remove the controls from the modelling population and to fit a model to predict the *probability of purchase given treatment*, which we denote p^T . More formally, we define p^T by

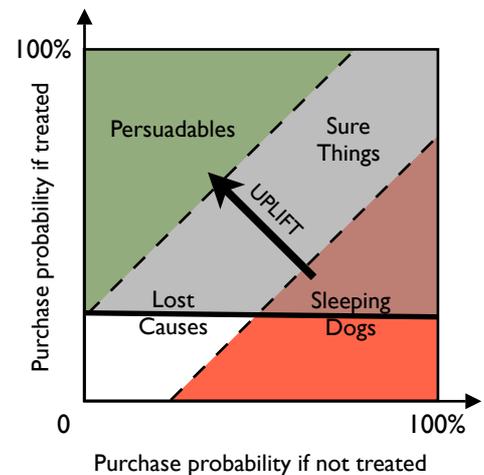
$$p^T = \text{Prob}(\text{purchase} \mid \text{treatment}).$$

This is exactly the quantity plotted on the vertical axis of the soft form of the Fundamental Campaign Segmentation. As a result, targeting the “best” people as identified by a “response” model corresponds to selecting those above some horizontal line. As is illustrated to the right, to catch all the Persuadables in this way, it is normally necessary to include all of the Sure Things (who will buy anyway), and a fair number of the Lost Causes and Sleeping Dogs as well.³

Though less often used, there is another kind of propensity model, namely the *penetration* model. Penetration models simply take the customer base and model how likely a customer is to have the product in question, i.e. the *probability of having purchased*. Depending on the nature of marketing that has been carried out previously, this can be quite similar to the quantity plotted on the horizontal axis, the *probability of purchase given no treatment*, which we denote p^U , i.e.

$$p^U = \text{Prob}(\text{purchase} \mid \text{no treatment}).$$

Of course, the penetration model could be restated as estimating the *probability of having purchased given the marketing mix in operation since that customer joined*. In practice, penetration models are mostly used when



³ Though be aware that population density is not shown on the diagram and that the exact positions of all segment boundaries are subject to interpretation and choice.

there is no previous campaign to analyse, so this is quite similar to p^U . Allowing this approximation, a penetration model predicts something close to the quantity plotted on the horizontal axis of the (soft) Fundamental Segmentation, so that selecting a population on this basis corresponds largely to selecting a population to the right of some line on the diagram. As can be seen, this is even worse than targeting with a so-called response model, since it gives priority to Sleeping Dogs and Sure Things, while missing significant numbers of Persuadables.

6. Case Study: Cross-Selling

A major North American bank repeatedly ran (and runs) a cross-sell campaign for a high-value product to selected segments of its customer base.

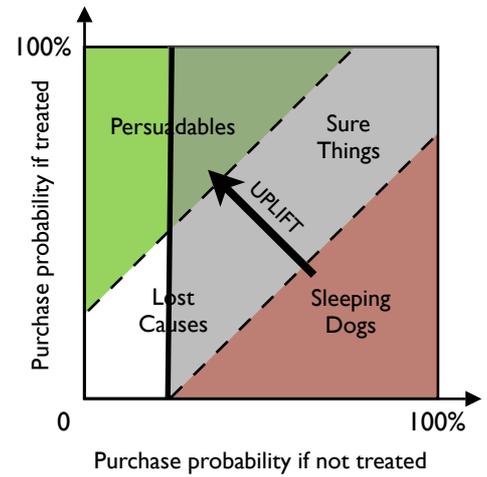
The first run of the campaign was an untargeted trial, sent to approximately 10% of the base. This provided good baseline data for conventional “response” modelling. The background purchase rate in the control group was around 0.9%, while in the treated group this rose to about 1.1%, giving a small uplift of around 0.2 percentage points. Since the mailing cost was around \$0.50, and the value of a sale was over \$1,000, this was a highly profitable campaign.

The second run of the campaign targeted the best 30% of those identified by a standard “response” model. A mailing control group (around a tenth of the total) was kept, and around 10% of the lower seven deciles were also mailed as a “targeting control”. (These allow the quality of the targeting to be assessed.)

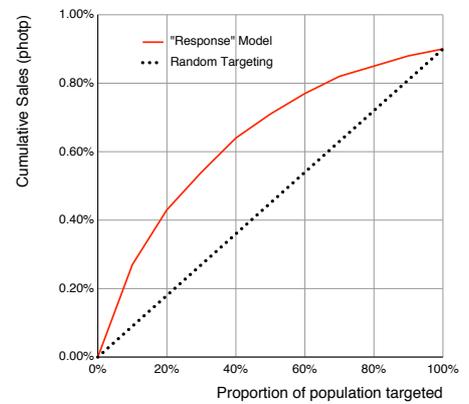
The result of targeting with a conventional “response” model, when assessed on the basis of uplift, is shown on the two graphs to the right. The upper graph (the gains chart) shows a conventional view of the model, and it is quite encouraging. But when a conventional response-based gains chart is replaced with one that shows the *incremental* sales generated at different targeting volumes, the results look almost unbelievably bad, showing that targeting the “best” 30% (as chosen by the “response” model) yields no incremental sales at all. In fact, the rate of sales is very slightly lower than in the control group, though not significantly so.

But there is no mistake here. The explanation is simply that the “response” models are not supposed to predict incremental purchases, but merely purchases: in this case, clearly, the people it targeted would have bought anyway.

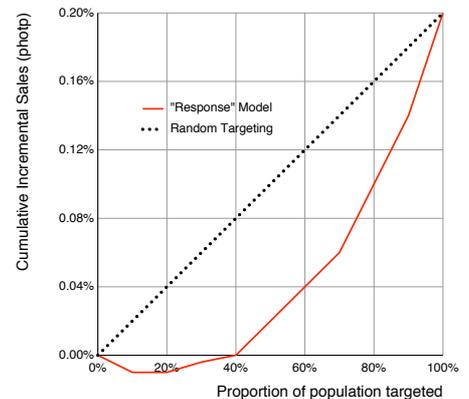
Showing some fortitude, the bank in question then tried two approaches to modelling incremental impact. The first was to build two models, one on the treated population, and the other on the untreated population.⁴ Subtracting the latter from the former gives an unbiased estimate of the uplift. We refer to this as “poor-man’s” uplift modelling. The problem with this technique is



Gains Chart for "Response" Model



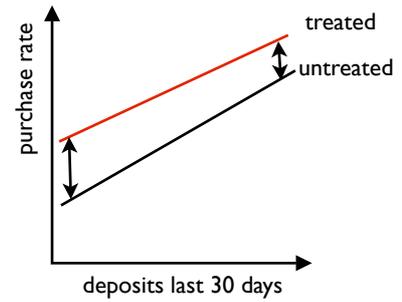
Targeting with "Response" Model



The vertical axis label “photp” means “per head of total population”. So if the total population is 1,000,000, cumulative incremental sales of 0.2% photp means 2,000 extra sales.

⁴ This required careful weighting to take account of the different treatment proportions in the top three and lower seven deciles

that neither of the two models has the prediction of uplift as its goal, i.e. neither is *supposed* to predict it. While it is the case that if both models were perfect, their difference would perfectly model uplift, as the models depart from perfection we would expect that the errors in their prediction of uplift would increase significantly faster, if only because errors add when models are subtracted. The sketch graph to the right illustrates the problem.

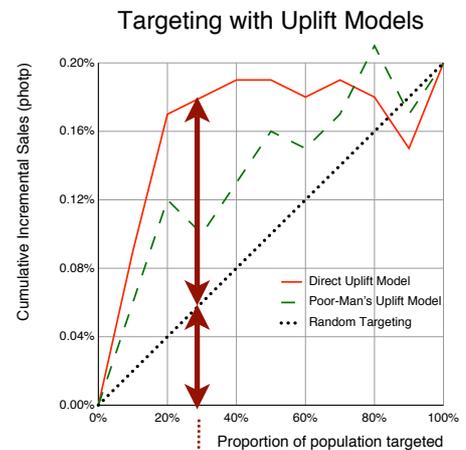


In this example, purchase rate increases with the level of deposits in the last 30 days, but the effectiveness of the intervention decreases (as can be seen from the convergence of the lines). It may be unrealistic to expect a difference of two models to predict this variation in incremental impact reliably.

The bank also used a dedicated uplift modelling package to model the incremental impact directly. This was a tree-based method⁵ that directly works with both the treated and untreated populations and has as its specific modelling goal the fitting of variation in uplift as a function of the predictor variables. We refer to this as “direct” uplift modelling. Another campaign was run, again targeting around 30% of the base. In this case, however, some of the targets were chosen by the poor-man’s uplift model, and some by direct uplift modelling.

⁵ Uplift Optimizer, from Portrait Software

The results are shown in the graph to the right.⁶ This shows a number of interesting things. First, both uplift approaches not only led to significant incremental sales at the chosen cut-off (30%), but significantly (unlike the “response” model) out-performed an untargeted approach. In the case of the direct uplift model, incremental sales were about three times as high as with random targeting, and with the poor-man’s uplift model they were around double. Secondly, at all reasonable targeting volumes the direct uplift model significantly outperformed the poor-man’s approach, as well as exhibiting considerably more stability. Thirdly, the last couple of deciles exhibit very unstable behaviour, and indeed there is quite a lot of noise in general beyond the third decile. This is largely because treated volumes were much smaller beyond this point, increasing noise. We suspect that most of the variation beyond 40% (for the direct uplift model) and beyond 60% for the poor-man’s model is actually noise. There were also some discrepancies between the modelling population and the campaign population, and we would expect to see greater stability if the models were rebuilt.



*“Three times the extra sales,
Same mailing volume:
Three times the profit.”*

In terms of financial impact, the campaign (which was profitable even with the initial random sampling strategy) was over three times as profitable using direct uplift modelling and cutting off at 30%. The increase results from two factors—a cost reduction from targeting fewer people and a higher absolute volume of sales as a result of driving more incremental business (and therefore more total business). It could have been even more profitable had the cut-off been reduced to 20%, with only a marginal impact on incremental sales. In the case of the poor-man’s uplift model, the optimal targeting volume was rather larger (more like 50%) and incremental sales somewhat weaker at that volume than targeting 20% or 30% with the direct uplift model. However, even this poor-man’s approach to uplift modelling is significantly more profitable than random targeting, which in turn, for this campaign, was obviously vastly better than using a conventional “response” model, which generated no incremental sales at all (previous page).

⁶ Numbers have been altered slightly for reasons of commercial sensitivity, but reflect all essential aspects and understate, rather than overstate, the impact, and preserve the relationship between the performance of the approaches

As well as the general evidence that this example provides that uplift modelling can be very efficient at generating incremental sales profitably, it is impor-

tant to note that when the contact volume is fixed (as is frequently the case), better targeting not only reduces the cost per incremental sale, but also increases the absolute volume and value of incremental sales. The summary impact, for this campaign, is shown in the table to the right.

7. Negative Effects

The cross-sell example in the previous section has illustrated the difference between targeting on the basis of purchase probability and targeting on the basis of uplift. However, it certainly has not provided compelling evidence for the existence of negative effects in direct marketing. We now consider these.

Some kinds of marketing interventions frequently have negative effects for significant proportions of the customer base. Perhaps the clearest example of this is the case of retention activity. Here, the mainstream approach is to target people who are perceived to have a high risk of attrition (often weighted by customer value). Unfortunately, such people are, almost by definition, often dissatisfied, and because of this virtually any approach by the business carries a significant risk of back-firing and bringing forward the very attrition it is designed to prevent.⁷ Indeed, occasionally the *overall* impact of retention campaigns is to increase attrition.

There are a number of factors that seem to increase the likelihood of negative effects. These include intrusive contact mechanisms (particularly outbound calling) and excessive contact frequency. It is also obviously the case that some creative messages may offend or otherwise actively repel certain customers, or may emphasize features of a product or service that are unappealing to a particular customer. There is also a suggestion that contact may in some cases turn what would have been an impulse purchase into a considered purchase that may be more likely to be subject to comparison shopping. For example, certain claims (“one of the highest interest rates available”) may inadvertently encourage a customer to go to a comparison site when she might otherwise simply have stuck with her current institution.

Across the various cross-sell, up-sell and deep-sell campaigns we have looked at and modelled, we have frequently found that the last one or two deciles are negative on the historical datasets used to build models. However, just as any model usually performs slightly more strongly on historical data than new data (if only because causal relationships change over time), these effects usually weaken on deployment, though are still sometimes present. Certainty in this area is made more difficult because, of course, once uplift models have been built and validated, campaigns are usually heavily biased in favour of those showing strong positive uplift, with at most low volumes of targeting controls being employed in the bottom deciles.

For sales-generation campaigns, it is of course only critical to understand whether there are negative effects if the intention is otherwise to target eve-

Impact of targeting best 30% by score		
	Uplift (pc pt)	Profit per million customers mailed
Response Model	-0.004%	-\$540,000
Poor-Man's Uplift Model	0.05%	\$0
Uplift Model	0.18%	\$1,300,000

⁷ For more details on this topic, see “Identifying Who Can be Saved and Who will be Driven Away by Retention Activity”, also available from Stochastic Solutions in two versions, one focusing on churn in mobile telecoms and another focusing on attrition in financial services.

“*Triggered attrition, intrusive contact, induced comparisons — just three aspects of negative effects.*”

ryone. Whenever the contact cost is non-negligible, campaigns tend to become unprofitable well before any negative effects set in.

It is also worth noting that there are other kinds of applications, such as changing credit limits, interest rates or service levels, where the potential for various kinds of negative effects is much greater. In those cases, clearly this becomes a very salient issue.

8. Incentives and Response Codes

There are special factors to consider when a marketing campaign involves some kind of incentive, such as a discount. There are, of course, many different reasons for giving discounts including generation of good-will and customer loyalty, protection of market share, response to competitor activity, a belief that such an “investment” will increase long-term customer value and reductions in costs of production or other business efficiencies. If these or similar factors are the main motivation behind an incentivized campaign, then the arguments in this paper may not be relevant.

If, however, the principal purpose of an incentivized direct marketing campaign is the generation of incremental sales, and more particularly, incremental profit, accurate assessment and prediction of uplift becomes critical. The problem here is two-fold. The first is that it is fairly common for campaign codes to be used as a mechanism for assigning “credit” for a sale to a particular marketing initiative. While we would argue that using a control group is a much better way to assess incremental impact, it is clearly the case that quoting a campaign code or presenting a coupon provides reasonable evidence that a customer has been at least somewhat influenced by the corresponding campaign. However, this evidence is much more dubious when the campaign code gives access to a discount, as it is manifestly the case that many—perhaps even most—customers who plan to buy anyway will take advantage of any incentives of which they are aware. Indeed, helpful staff often inform customers of incentives that are available, frequently to the long-term benefit of the business. Clearly, however, as well as providing various benefits, incentives carry a direct cost, and if a sale that would have occurred anyway now attracts a discount, that has an immediate negative effect on profitability. Manifestly, if the goal of the campaign is to stimulate profits, the ideal campaign targets are people who will buy with the incentive, but not without. Once again, uplift modelling is designed precisely to identify such people.

Even in this context, however, it will sometimes be the case that a business will feel that it is in some way “wrong” or dangerous to offer discounts to some customers and not to others. Indeed, there is the wholly hard-headed point that loyal customers can be particularly infuriated to discover that less-loyal people are given preferential treatment, and this can lead to strong negative effects very similar to those discussed in the previous section. Once again, it is not the purpose of this paper to discourage businesses

from giving discounts or treating customers “fairly”, nor indeed to take a one-dimensional view of the potential benefits of a marketing campaign. We simply argue that it is important to be clear as to what the goals of a campaign are, and that if those goals are mainly to maximize the profit contribution from an incentivized cross-sell, up-sell or deep-sell campaign, it is important to take account of the fact that customers who would buy anyway are extremely likely to use incentives given to them.

9. When is Uplift Modelling Needed?

We have discussed in detail one example in which uplift modelling is clearly vastly superior to “response” modelling, and have made reference to other cases, particularly retention initiatives, in which it is even more effective (because of the existence of significant negative effects for anything from 10% to 70% of the population). We have applied the method to various other cross-sell and related campaigns, and in most cases it has significantly outperformed “response” modelling, but there have been a few cases in which the difference has not been significant. It would clearly be useful to characterize when it is superior, and when it is not.

We begin with an obvious point: uplift is the difference between purchase probability with and without treatment. So if the purchase probability without treatment is zero or negligible, uplift models reduce to response models.⁸ In practice, this means that uplift modelling is more important in situations in which there are many factors potentially driving a customer to purchase. For example, well known companies with large advertising spend and major high-street presence have to be relatively careful when attributing an apparently incremental sale to a particular direct marketing campaign.

Perhaps a more subtle point is that uplift modelling is less important if sales are strongly correlated with incremental sales, i.e. if the campaign is most effective for people in segments with a high background purchase rate. This is particularly true if only the rank ordering from the model is to be used, as is the case when a fixed targeting volume is chosen. One situation in which we have seen this several times is in retail, for example with catalogue retailers. It is not hard to imagine that customers with an affinity to a particular catalogue might be more likely than others to open and read a newly received copy, and might therefore have both a high background purchase rate, and a high incremental purchase rate. On the other hand, we have also seen retail situations, including other catalogue mailings, in which uplift modelling has been able to identify good targets significantly better than a conventional “response” model; so the evidence is not clear-cut.

Other considerations are model complexity and noise. Although this paper is not primarily concerned with techniques, uplift modelling is more complex than “response” or penetration modelling, since it seeks to fit the difference in behaviour between two populations.^{9, 10, 11} So uplift models are *second order* and so are harder to fit. This also means that there is more intrinsic noise to overcome with uplift models, and this is usually exacerbated in prac-

⁸ Equivalently, when the purchase rate in the control group is zero, “response” models actually do predict true (i.e. incremental) response.

⁹ Radcliffe N. J. & Surry, P. D. (1999). “Differential response analysis: Modeling true response by isolating the effect of a single action.” Proceedings of Credit Scoring and Credit Control VI. Credit Research Centre, University of Edinburgh Management School.

¹⁰ Radcliffe, N. J. (2007). “Using Control Groups to Target on Predicted Lift: Building and Assessing Uplift Models”, Direct Marketing Analytics Journal, Direct Marketing Association, 2007.

¹¹ Lo, V. S. Y.. (2002). “The true lift model”. ACM SIGKDD Explorations Newsletter. Vol. 4 No. 2, 78–86. 1

tice because control groups are rarely of truly adequate size. (In practice, control groups, being sized primarily for *measuring* uplift, are typically only around 10% of the size of treated groups.) For these reasons, uplift models are sometimes no better than conventional “response” models, and could in principle be worse in some cases.

Notwithstanding these caveats, in general, we believe uplift modelling to be the better approach, on the basis (following John Tukey¹²) that an approximate solution to the right problem is normally more useful than a more precise solution to the wrong problem. However, the effectiveness of uplift models should always be checked, and it is never a bad idea to benchmark against a conventional model.

10. Conclusion

Best-practice post-campaign analysis already uses control groups to allow calculation of the incremental impact—or uplift—of a marketing campaign. Unfortunately, the corresponding mainstream best practice for targeting campaigns uses so-called “response” models, which are not designed to maximize the incremental impact that we use as the yardstick for judging the success of campaigns. In this sense, there is a fundamental mismatch between accepted direct marketing goals and mainstream approaches to targeting.

We have argued that from a theoretical perspective targeting should normally be performed using a model specifically designed to maximize incremental impact. We have also shown one of several cross-sell campaigns we have seen in which this difference was extremely important in practice. As we discussed, there are both simplistic (“poor-man’s”) approaches to building models that predict uplift and more sophisticated, direct methods for doing so.

Of course, in some cases all of the apparent responses are truly incremental, and when this is so an uplift model will offer no advantage; indeed, it may be less good, because it is a more complex approach. Similarly, if response and uplift are strongly positively correlated, the approach may offer not any advantage if scores are only to be used for rank ordering. However, overall, the weight of evidence strongly supports uplift modelling as better way of targeting, both in theory and in practice.

In our experience, organizations that adopt uplift modelling rarely if ever go back to “response” modelling. We take this as evidence that—as we like to say—*one day, all targeting will be done this way.*

¹² “Far better an approximate answer to the right question, than the exact answer to the wrong question, which can always be made precise.” — John Tukey, (1962), “*The future of data analysis.*” *Annals of Mathematical Statistics* **33** (1), pp. 1-67

“ *One day,
all targeting
will be done
this way.* ”