

## Paper AA-08-2015

# Get the highest bangs for your marketing bucks using Incremental Response Models in SAS® Enterprise Miner™

Delali Agbenyegah, Alliance Data Systems, Columbus, Ohio

## 0.0 ABSTRACT

Traditional marketing predictive models target customers who are likely shop, make more trips or spend more. While this approach generally yields higher marketing campaign performance results over random selection, it can sometimes lead to money wasting on customers who will shop regardless of marketing offers and 'do not disturb' customers who will rather stop shopping if you 'disturb' them with marketing offers. Net lift models are used to identify 'persuadable' customers who have higher likelihood to respond to marketing campaigns and help marketers maximize their return on marketing investments. This paper simplifies the basic concept of Net Lift modeling using real life examples and shows how this can easily be accomplished using SAS® Enterprise Miner™. The paper concludes with real life challenges of Net Lift modeling and suggests ways to handle some of those challenges as well as recommended situations where Net Lift models work best.

## 2.0 INTRODUCTION

Targeting customers for direct marketing has evolved over the past decade. About thirty years ago, marketing professionals target customers using grids computed by customer Recency (Time since last purchase), Frequency (number of trips a customer make in a defined period) and Monetary (amount customer spent) which is popularly known as RFM strategy. This strategy proved to work better than a random selection of customers for marketing campaign.

With the introduction of Predictive Analytics in Marketing during late 1980s, marketing professionals now build predictive models for customer selection in direct marketing campaigns. These models include response models that rank order customers' likelihood to make purchases or spend more regardless of marketing offers. Other marketing professionals build model on customers that have marketing incentives (treatment) to separate the likely respondents from the non-responders. Various modeling methods such as Regression models, Decision trees, Random Forest or neural networks are used to accomplish these goals.

The problem of these conventional response model is that they ignore customers who could have shopped naturally without receiving marketing incentives or customers who perhaps decided not to shop due to the 'disturbance' of marketing offers. Relying on these conventional methods therefore will not ensure the optimal return on marketing investment. The ideal method is to identify the customers who otherwise will not have shopped without receiving marketing offers and to provide incentive to those customers with the offers to encourage them to make purchases. Incremental response (or Net Lift) models are designed to identify that profitable group of customers that need incentive to shop.

## 3.0 INCREMENTAL RESPONSE (NET LIFT) MODELS

Incremental response model attempts to distinguish or identify customer groups that need marketing incentives to shop from customer groups that will shop naturally or will not shop if given marketing offer in order to allocate resources appropriately to optimize returns on marketing investments. The model is built on a treatment (receive marketing promotion) and control (do not receive marketing promotion) data set up combined with other demographic and transactional attributes available to determine the key drivers that differentiate customers who respond under treatment from natural responders and apply these key drivers in scoring the model for future campaign selections.

**Table 1: Customer groups in a Direct Mail offer set up.**

PURCHASE IF TARGETED	PURCHASE IF NOT TARGETED	
	YES	NO
NO	Do not disturb	Never mind
YES	Natural shopper	Highest bangs for your bucks

The figure above clearly illustrates the power and the importance of the Incremental response models. Accurately identifying the ‘highest bangs for your bucks’ group will ensure an optimal allocation of marketing investments.

#### **4.0 INCREMENTAL RESPONSE MODELING DATA**

Availability of ‘good’ campaign history data with reasonable treatment and control groups is crucial for building an accurate Incremental Response model. In addition to all available transactional data attributes, demographic attributes as well as any other available data that can determine a customer’s future purchases, the modeling data has to contain a treatment versus control indicator in the selected promotion window as well as the customer responses during the selected promotion window. The data should be large enough to allow for a model building set and a model validation data split.

#### **5.0 INCREMENTAL RESPONSE MODELING METHODS**

Due to the inherent nonlinearity in net lifts generally and the unbalanced distribution of treatment and control groups, it is important to select modeling techniques that will be robust enough to handle your particular case. Another important complication arises as a result of the practical impossibility of including a customer in both the treatment and control group simultaneously. Although this paper is focused on modeling using SAS Enterprise Miner, a brief overview of some Incremental Response Modeling techniques are provided for your review. Further reading is encouraged for a deeper understanding of the techniques.

##### **5.1 LINEAR DIFFERENCE METHOD**

The simplest method of approaching incremental response modeling is to develop two different independent Logistic regression models, a separate model for the Treatment group (MT) and another for the control group (MC). Scoring is then accomplished by scoring customers with the separate models, taking the difference between the predictions of the two models ( $D_i = P_{mt} - P_{mc}$ ) and finally ranking the difference ( $D_i$ ) to obtain the predictive model ranks. While this model makes in principle and easy to implement, it has some shortfalls due to the second order phenomenon modeling and relative small signal strength as well as the nonlinearity of incremental lifts in practice. Refer to Radcliffe and Surry (2011) for further explanation regarding the shortfalls of the linear difference method.

##### **5.2 BIFURCATED LOGISTIC REGRESSION METHOD**

Lo (2002) proposed a single Logistic Regression model by adding interaction terms created by multiplying treatment indicator ( $T=1$  for treatment group and  $T=0$  for control group) with all the independent variables ( $x_i$ ) and include all these variables as potential predictors to model the binary outcome ( $Y=1$  if customer made a purchase during promotion window and  $Y=0$  if no purchase is made). Lo argued his motivation of introducing a single model rather than two independent regressions on the basis that estimated incremental lifts can be sensitive to statistically insignificant differences in the parameters of the treatment and control models. He applied this approach to his work on direct marketing with Fidelity and reported good results (Lo, 2005). Refer to Lo(2005) for further explanation using the bifurcated regression modeling approach.

### **5.3 DECISION TREE BASED UPLIFT MODELING**

Hansotia and Rukstales (2001, 2002) applied a decision tree based approach to model what they called Incremental Value Modeling. In their approach, Hansotia and Rukstales used raw uplifts in the two subpopulations as a split criterion, which is a natural approach but it has a disadvantage of not taking the population size into account. Refer to Hansotia and Rukstales (2001, 2002) for details using this method.

A different version of a decision tree based approach was adopted by Radcliffe and Surry in what they called Significance-Based Uplift Trees. They used the difference in uplift between the two sub populations as a split criterion and used the statistical significance of that measure as a quality measure of the split in a decision tree set up. Radcliffe and Surry (2011) for further explanation about the Significance-Based Uplift trees which has seen many successes in real life applications. Their method is what is implemented in Portrait Uplift product, a commercial uplift modeling product from Pitney Bowes.

Rzepakowski and Jaroszewicz (2010) also introduced a decision based approach which used a splitting criteria based on information theory. They applied their approach in a clinical trial set up and reported good success.

### **5.4 UPLIFT MODELING USING K NEAREST NEIGHBORS**

Another mathematically appealing method of modeling incremental response is by using K nearest Neighbor approach. This method first conducts neighborhood analysis to find optimal distance measurements, and then find k nearest neighbors for a candidate customer and finally develop a net lift score for the neighborhood. While this method seems very appealing, it is computationally intensive.

### **5.5 NAÏVE BAYES APPROACH**

Larson(2010) introduce a new approach of uplift modeling using a modification of the popularly known Naïve Bayes Classifier by relaxing the Naïve assumption of the conditional independence of the predictor variables( $X_i$ ) given the response ( $Y$ ) and directly modeling the net lift between the treatment and control groups with what he called the Generalized Naïve Bayes Classifier. In his work, Larson emphasized variable selection as a crucial component in uplift modeling and he proposed the Net weight of evidence (NWOE) and consequently the Net Information Value (NIV) which is a modification of the popularly known Information value used in variable reduction for binary responses. SAS Enterprise miner uses the NIV and applies its variation in variable selection in the incremental response modeling. The method is explained in detail below.

## **6.0 INCREMENTAL RESPONSE MODELING IN SAS ENTERPRISE MINER**

In the following sub section, we discussed in detail the methodology implemented in SAS Enterprise Miner to model incremental response.

### **6.1 VARIABLE SELECTION**

At the heart of any predictive modeling is variable selection. It is even more crucial in incremental response modeling since we are modeling a second order effect that in most applications has relatively small magnitude. Identifying the variables that influence incremental response such set up described above is a big challenge. SAS Enterprise Miner implements the NIV method proposed by Larson (2010) described below.

#### **6.1A THE POPULAR WEIGHT OF EVIDENCE (WOE) AND INFORMATION VALUE (IV)**

Let  $Y$  denote a binary response variable  $\{y=1$  if response is a success,  $y=0$  otherwise $\}$ . Suppose  $X$  is a predictor variable of  $Y$  grouped into  $I$  mutually exclusive bins. The weight of evidence is defined and computed for each bin as

$$WOE_i = \log \frac{P(x = x_i | Y = 1)}{P(x = x_i | Y = 0)} \text{ for } i = 1, 2, \dots, I$$

And the information value is calculated as

$$IV = \sum_i [P(x = x_i | Y = 1) - P(x = x_i | Y = 0)] WOE_i$$

Many practitioners, including the author has used the IV as a measure of the strength of the correlation between the response and the explanatory variable and have recorded success and a rule of thumb frequently used is IV greater 0.3 indicates the variable has a strong predictive power. Larson (2010) extended similar idea in a treatment versus control population set up which he called the Net Weight of Evidence described below.

### 6.1B THE NET WEIGHT OF EVIDENCE (WOE) AND NET INFORMATION VALUE (IV)

Consider a marketing or clinical trial set up where one population receive a Treatment (T) and the other held as a Control(C), and again suppose X is a predictor variable of Y grouped into I mutually exclusive bins, Larson(2010) defined the NWOE for each bin as

$$NWOE = \log \frac{P(x = x_i | Y = 1)_T / P(x = x_i | Y = 0)_T}{P(x = x_i | Y = 1)_C / P(x = x_i | Y = 0)_C}$$

And consequently,

$$NIV = \sum_i [P(x = x_i | Y = 1)_T P(x = x_i | Y = 0)_C - P(x = x_i | Y = 0)_T P(x = x_i | Y = 1)_C] NWOE_i$$

This method helps in identifying key predictors that differentiates customers who responds due to treatment from the remaining customer groups. In SAS Enterprise Miner, the NIV is implemented with some variation to select the best candidate variables, usually above a pre-determined threshold.

### 6.1C THE PENALIZED NET INFORMATION VALUE

Given the challenges associated with second order effect with relatively small magnitude modeling, Larson further introduced a penalty factor to evaluate the performance of a predictor within the training data compared to a validation data. A variable that lacks consistency in its performance between the training and validation data is penalized using the following method. For each variable, compute the NWOE for both the training data and validation data and then take the difference ( $w_i$ ) as

$$W_i = |NWOE_{\text{train}} - NWOE_{\text{valid}}|$$

A penalty factor (PF) is calculated as

$$PF = \sum_i |P(x = x_i | Y = 1)_T P(x = x_i | Y = 0)_C - P(x = x_i | Y = 0)_T P(x = x_i | Y = 1)_C| W_i$$

And the penalized NIV is computed as

$$PNIV = NIV - PF$$

This provides a more robust variable selection approach and SAS Enterprise Miner automatically uses the PNIV as long as validation data is provided.

## 6.2 INCREMENTAL RESPONSE MODEL FITTING AND VALIDATION

If you are familiar with SAS Enterprise Miner, then you already know that fitting any model and validating it is pretty straightforward as long as you have created your library, created your data source and have your variable roles assigned correctly. For Incremental Response modeling, create your data source using the Advance Adviser and use the Data Partition node to split your data into training and validation. Be sure to assign partition roles to both your target variable and treatment variable which are all binary in incremental response modeling. This will ensure there is no bias in the data partitions with respect to the target and the treatment variables. The Incremental Response node is then connected to the Data Partition node as shown below. The default settings under the Incremental sales node are pretty standard in most applications. It is however advised that you check and make all settings suitable to your needs.



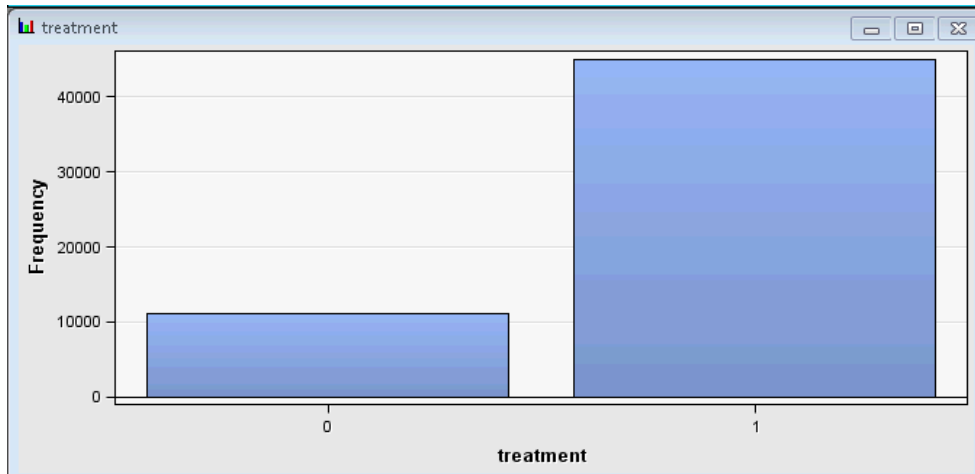
**Figure1: Process Flow Diagram for Incremental Response Modeling**

Once the flow diagram is executed, SAS Enterprise Miner will fit the model and create Incremental Response Model Diagnostic Plots and tables for your review. Various other useful outputs will be generated in the results window for your review. The best way to ensure that your model is performing as expected is to check the incremental response model diagnostic plot is to ensure that the observed and predicted increments are relatively close and the increments decrease monotonically from the top to the bottom percentiles.

## 7.0 A REAL LIFE EXAMPLE

We describe a simple Incremental response model using a real life data with masked input variables for a loyalty marketing retailer. The data set has 58 variables including a treatment variable (Treatment=1 if targeted and Treatment=0 if not targeted), an ID variable and a target variable (Target=1 if customer responded to marketing treatment and Target=0 if customer did not respond). There are total of 56,000 records with 11,000 in the control group.

Once we create the SAS library and the Data Source in SAS Enterprise miner using the Advance adviser, we can quickly explore all the variables in the data set on the fly by selecting the variable of interest and clicking the explore menu in the variables window. The graph below provides a summary of the treatment variable in our dataset.



**Figure2: Summary of the Treatment variable.**

To complete Incremental sales modeling, we first add the data set to the process flow diagram. After that, we add the data partition node to split our data set into training and validation sets. Under the data partition node, be sure to assign partition roles to both the treatment and the target variables to allow unbiased and accurate sampling with respect to both the target and treatment variables. We set 60% of the dataset for training and use the remaining 40% for validation. The snapshot of the settings under the data partition node is shown below.

Name ▾	Partition Role	Role	Level
treatment	<b>Stratification</b>	Treatment	Binary
Target	<b>Stratification</b>	Target	Binary
input9	<b>Default</b>	Input	Interval
input8	<b>Default</b>	Input	Interval
input7	<b>Default</b>	Input	Interval
input6	<b>Default</b>	Input	Interval

**Figure3: Partition Role set up under Data Partition Node**

The next step is to add the Incremental Response node which is under the applications tab of the process flow diagram window. Most of the default settings under the incremental response Node are pretty standard, but be sure to check and change settings to your preference, especially the Rank Percentage Cutoff which determines the percent of the variables that should be selected into the final model using the Penalized Information Value criteria. The process flow diagram is shown below:



**Figure4: Process Flow Diagram for Incremental Response Modeling**

Finally, we run the above process flow diagram to complete our Incremental Response Modeling. The results window shows several model diagnostics output usually by deciles and this is helpful in determining which rank groups will generate incremental response for the marketing activity under consideration and can be used to determine cut off percentiles. Below is a snapshot of our Incremental Response Model Diagnostics plot. If the model is accurate, we expect the observed increment to be very close to the predicted increment as well as decrease monotonically from percentile 10 to 100. Overall, our data has 2.3% incremental response but our Incremental Response model identified customers in the top rank that has about 3 times higher Incremental response rate than the average. Figure 3 depicts the result by each percentile. Using this model as well as cost and revenue information, we recommended targeting the top 50% of the customers for the next campaign.

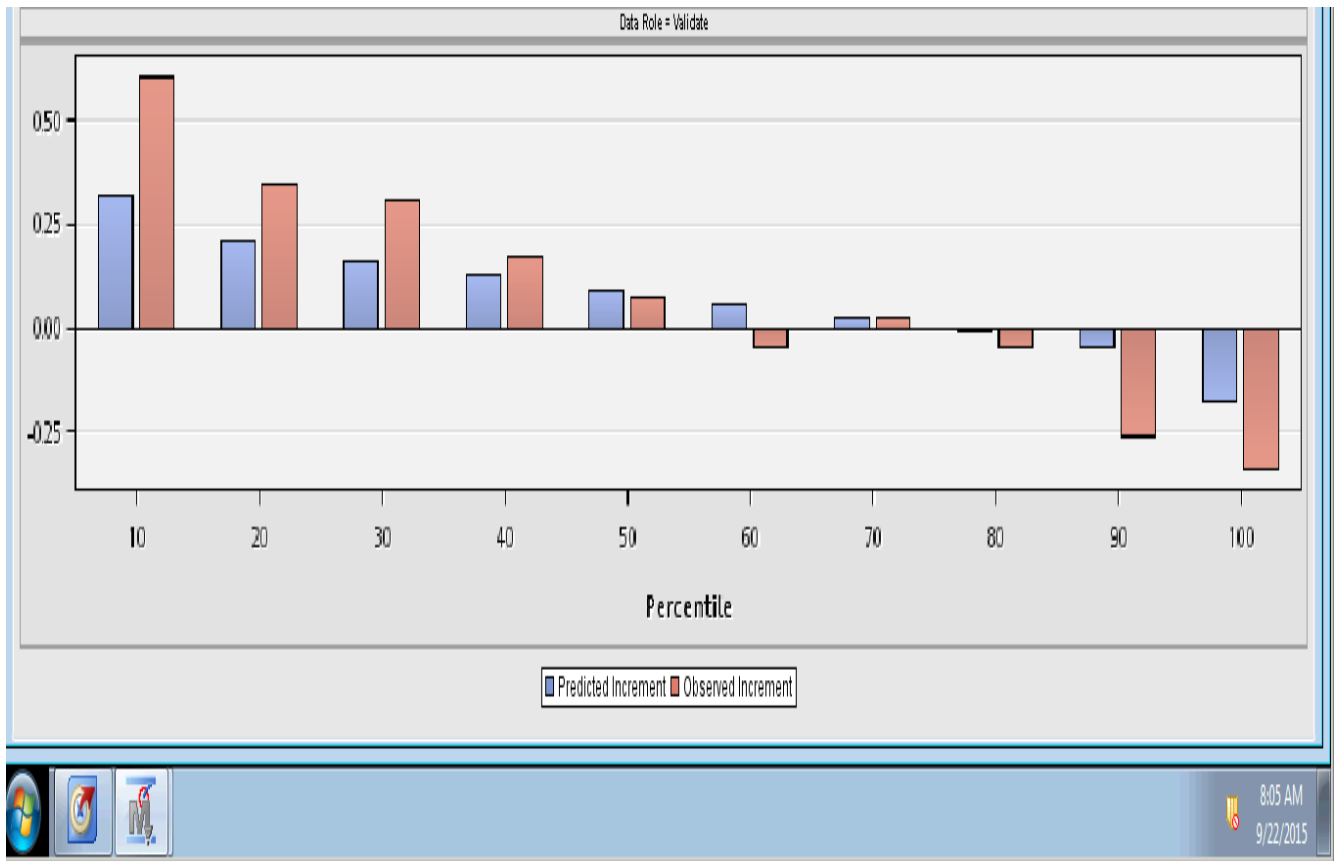


Figure5: Incremental Response Model Diagnostic Plot

## 8.0 CONCLUSION

Incremental Response Model is useful in maximizing response to direct marketing offers, but it comes with a lot of practical challenges. Firstly, there has to be good history of a valid control group of a reasonable size. Without a treatment and control split, it is impossible to build Incremental Response Model. It is also important to check if the goal of your marketing campaign is solely to increase response rates or maximize sales. In the author's experience this is not always the case. There are situations where certain customer groups have very high incremental responses relative to others but do not have higher absolute dollar values. It is therefore advisable to ensure that Incremental Response Model is suitable for your specific situation before implementing. In situations where the business is very offer driven and frequently sending offers to customers through different channels such as email and mobile, Incremental Response Models may not show impressive results.

As discussed in section 5 under Incremental Sales Modeling Methods, the inherent non linearity in net lift, the impossibility of a customer to be in the test and control groups simultaneously and the fact that we are modeling a second order phenomenon present some technical challenges with modeling methods and model stability and hence it is advised to cautiously select the best method that is suitable to your situation and conservatively conduct a rigorous validation.

## REFERENCES

- Lee, Taiyeong, Zhang, Ruiwen, Meng Xiangxiang, Ryan, Laura (2013) *Incremental Response Modeling Using SAS® Enterprise Miner TM*, SAS Global Forum 2013 Paper 096-2013, San Francisco, CA
- Larsen, K. (2010). *Net Lift Models: Optimizing the Impact of Your Marketing Efforts*. SAS Course Notes. Cary, NC: SAS Institute Inc.
- Lo, V. (2002). *The True Lift Model: A Novel Data Mining Approach to Response Modeling in Database Marketing*. ACM SIGKDD Explorations Newsletter 4:78–86.
- Lo, V. (2005). *Marketing data mining – new opportunities*. J. Wang, editor, Encyclopedia of Data Warehousing and Mining. Idea Reference Group.
- Radcliffe, Nicholas J., Surry, Patrick D. (2011). *Real-World Uplift Modeling with Significance-Based Uplift Trees*, Boston, MA: Stochastic Solutions Limited.
- Rzepakowski, P., Jaroszewicz, S (2010.) *Decision trees for uplift modeling*. IEEE Conference on Data Mining, pages 441–450.
- Hansotia, B. Rukstales, B. (2001). *Incremental value modeling*. DMA Research Council Journal, pages 1–11.
- Hansotia, B. Rukstales, B. (2002). *Direct marketing for multichannel retailers: Issues, challenges and solutions*. Journal of Database Marketing, 9(3):259–266.

## ACKNOWLEDGMENTS

The author will like to thank Yin Chen and the entire Predictive Analytics team at Alliance Data Card Services for their support.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Delali Agbenyegah  
Alliance Data Systems  
3100 Easton Square Place  
Columbus, OH, 43219  
[delali.agbenyegah@alliancedata.com](mailto:delali.agbenyegah@alliancedata.com)  
[www.delaliagbenyegah.com](http://www.delaliagbenyegah.com)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.