Challenges of Incremental Sales Modeling in Direct Marketing

Andrew Pole

Sr. Manager, Guest Data & Analytical Services

Target

September 14, 2009

Twin Cities Area SAS User Group

Agenda

- Project objectives
- Incremental sales measurement methodology
- Models used in Target campaigns
- Incremental sales modeling methods investigated
- Data available for modeling
- Model development and selection
- Model results
- Hypotheses: Why the models didn't work

Project objectives

Incremental sales modeling project

Business objective

Maximize revenue and profit, by identifying and contacting guests (customers) that are likely to spend *incrementally* upon receiving a Target direct mail piece

Analytical objective

Build a model to predict who is most likely to spend *incrementally* upon receiving a Target direct mail piece

How Target measures direct mail campaigns

How we analyze campaigns

- We measure the sales lift of mail vs. no-mail guests over the length of the coupon redemption period (about six weeks)

Segment Name	Mail / No Mail	Guests	PRE $/guest % diff	PROMO $/guest % diff	Adjusted PROMO $/guest % diff	INCR $/guest	Total INCR $
A	Mail	500,000	0.4%	4.3%	3.9%	$4.54	$2,269,076
A	No Mail	100,000					
B	Mail	200,000	-0.7%	1.3%	1.9%	$1.52	$304,636
B	No Mail	75,000					

We calculate *incremental sales* by comparing promo $/guest for TEST and NO MAIL, after adjusting for pre-period differences

[Data are fictional and only for illustrative purposes]

The challenge

- This report gives us performance at a guest segment level, but not at a guest level

Segment Name	Mail / No Mail	Guests	INCR $/guest
A	Mail	500,000	$4.54
A	No Mail	100,000	
B	Mail	200,000	$1.52
B	No Mail	75,000	

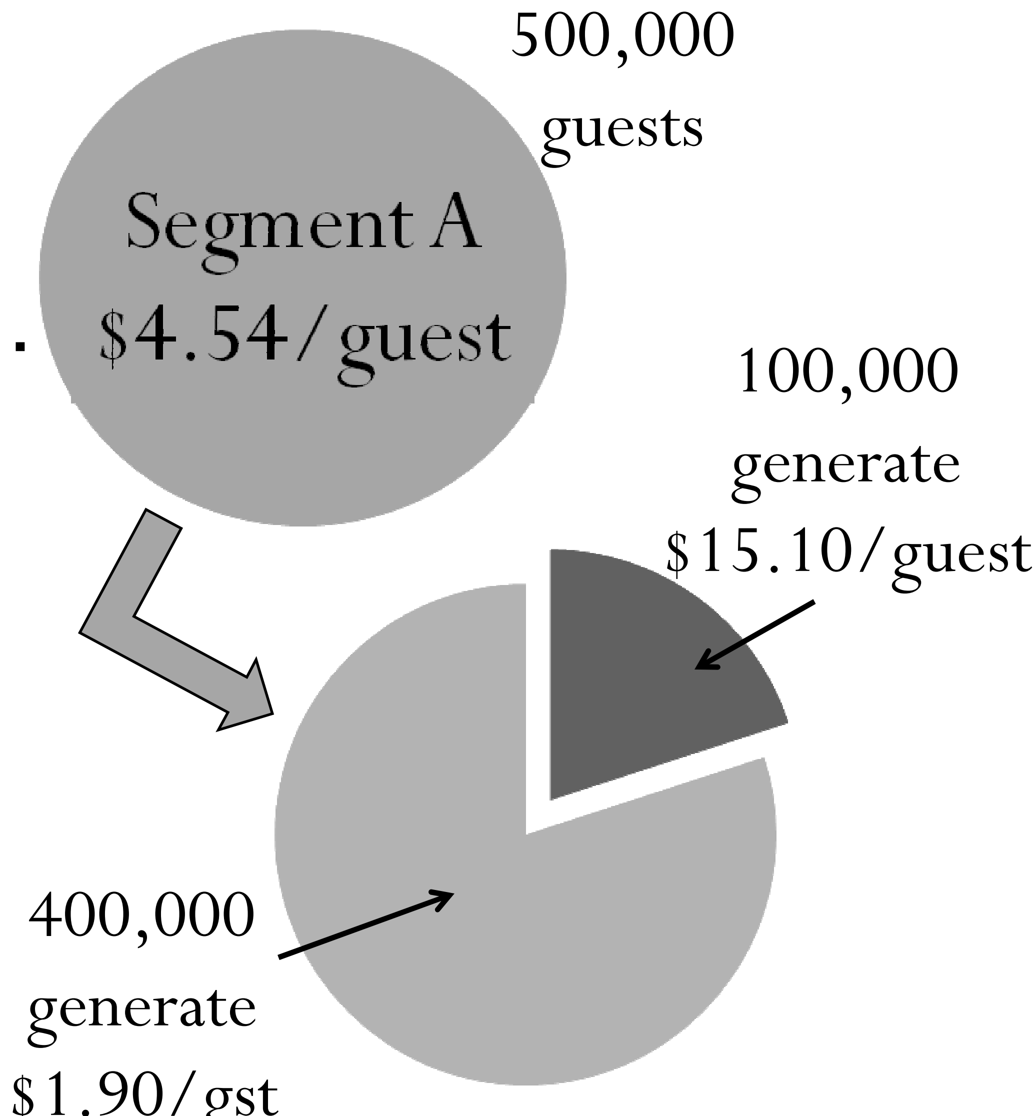
- So, how do we know which guests spent incrementally and which ones didn't?

[Data are fictional and only for illustrative purposes]

Who should we target?

So, on average, guests in segment A generated $4.54/guest…

…but what if we knew that 20% of them generated 70% of the incremental sales?



Models currently used in direct mail campaigns

Using predictive models in campaigns

- At Target, we have hundreds of models that are scored and used each month in campaigns
  - Response/Conversion models predict whether a guest is likely to purchase a product that the guest has/hasn't purchased before *(e.g. Should we give Andrew a cookie offer?)*
  - Payout models predict how much a guest will spend on a category *(e.g. Will Andrew spend over $50 on Home next month?)*
  - Demographic-inference models predict whether or not a guest has a particular demographic attribute or lifestage segment *(e.g. Does Andrew have 0-12 month year old in his household?)*
  - Redemption models predict whether a guest is likely to redeem a coupon *(e.g. Will Andrew redeem the coupon for $1 off laundry detergent?)*

  - Incremental sales models predict the incremental sales that a guest will generate if he or she receives a marketing contact

    (e.g. If I send Andrew the grocery coupon book, will he spend incrementally, and if so, how much?)

The challenge of building ISMs

- Incremental sales models (ISMs) are difficult to build because we don't know incremental sales
  at a guest level
  - We know incremental sales at a total campaign or segment level

- To build a model you have to have look in the past for known responses. In our database, we can find:
  - Guests that have shopped the category
  - Mothers with a 0-12 month old baby
  - Guests that have redeemed a TargetMail coupon

- So how do we build an ISM? We don't know actual incremental sales for each guest!

Incremental sales modeling methodologies investigated

ISM methodologies

Five primary methods were investigated:

1. Likelihood to shop ISM
2. Expected spend ISM
3. Two-stage spend ISM
4. Coupon redemption model
5. "Specialized" decision tree
   - Model difference between Test \$/gst and Control

Method 1: Likelihood to shop ISM

- In this methodology, we look to find guests that are significantly more likely to shop Target if contacted

Methodology

1. Build one model on TEST guests to predict likelihood that guest will shop Target, then build another model for CONTROL guests
2. Score every guest using both models
3. Subtract the TEST likelihood score from the CONTROL likelihood score to get the incremental likelihood to shop (given a contact)

Guest	P(shop | TEST)	P(shop | CONTROL)	Incr P(shop | contact)
A	80%	78%	2%
B	52%	52%	0%
C	68%	35%	33%

Method 2: Expected spend ISM

- In this methodology, we build two models, one to estimate incremental guest spend given TEST and one given CONTROL, and the final score is difference between the two

Methodology

1. Build one model on TEST guests to predict promo $, then build another model for CONTROL guests to predict promo $
2. Score every guest using both models, and subtract the TEST model score from the CONTROL model score to get predicted incremental gu[illegible]

Guest	Est $ | TEST	Est $ | CONTROL	Predicted incr $ | TEST
A	$120	$115	$5
B	$100	$100	$0
C	$55	$52	$3

Method 3: Two-stage spend ISM

- In this methodology, we use the results from the expected spend ISM (method 2), and then we build a model on these results to come up with one final model

Methodology

1. Build one model on TEST guests to predict promo $, then build another model for CONTROL guests to predict promo $
2. Score every guest using both models, and subtract the TEST model score from the CONTROL model score to get estimated incremental guest $
3. Build a model to predict incremental spend $ (using results from step 2)

Guest	Est $ | TEST	Est $ | CONTROL	Est incr $ | TEST	Model (incr $ | TEST)
A	$120	$115	$5	$4.52
B	$100	$100	$0	-$0.32
C	$55	$52	$3	$2.57

Method 4: Likelihood to redeem coupon

- In this methodology, we build a model to determine the likelihood a guest will redeem a coupon included in mail piece, and see if this correlates with incremental sales.

Methodology

1. Build a model on TEST guests to estimate the likelihood that the guest will redeem the coupon incentive included in the direct mail.
2. Look to see if guest segments with high coupon redemption scores also have high incremental sales.

Guest	Redeemed coupon?	P(redeem coupon)
A	Yes	80%
B	No	12%
C	Yes	68%

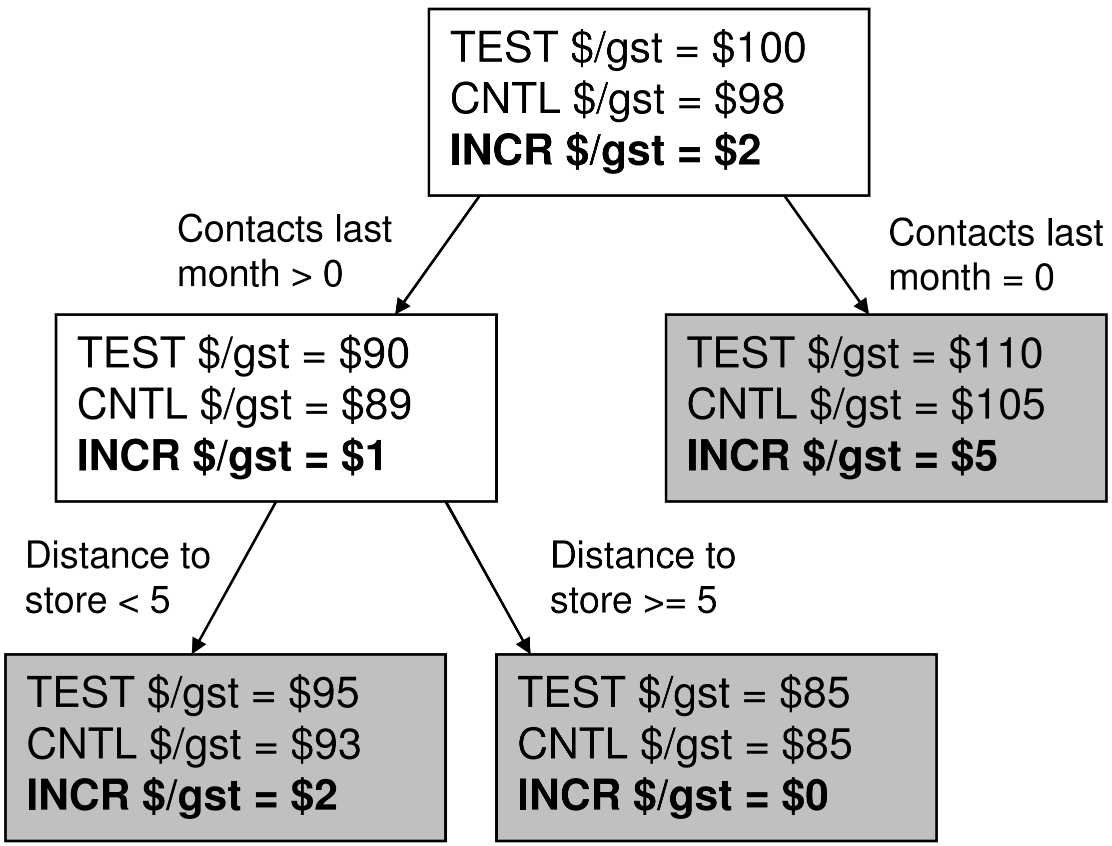
Method 4: Likelihood to redeem coupon

- While we were able to build fairly accurate models to predict likely coupon redeemers, high redemption likelihood <u>did not correlate</u> with high incremental sales

Because of these results, we <u>did not pursue</u> this methodology further

Method 5: Decision tree approach

- This method starts with overall campaign reporting results, and then uses variables to segment the guests into different groups in order to maximize the Test and Control difference between segments



Pros

- Only need to develop one model instead of two
- Logic and segments are easy to explain to clients

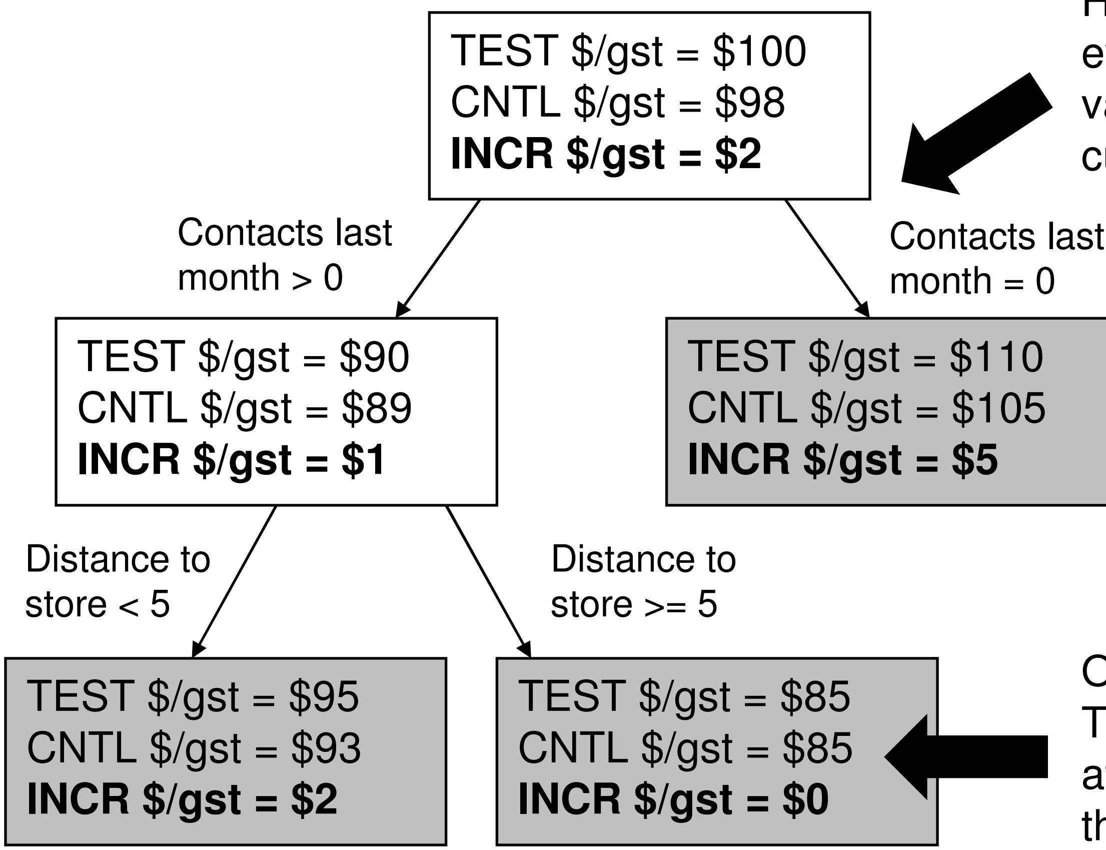
Cons

- Very manual development method
- Final model only found 3-4 variables that were useful in prediction
- Run out of sample very quickly

Method 5: Decision tree approach

- Because of the manual nature of model development, and because guest counts were too small as we grew the tree, we <u>did not pursue this methodology</u> further

Had to <u>manually</u> determine, evaluate, and iterate what variables to split on and at what cut-off points to use



Oftentimes, our guest counts in TEST or CONTROL got too small after growing the tree further than 2-3 levels deep

Data available for modeling

Data available

- Modeling data sets were developed based on guest data from completed direct mail campaigns

- Generally speaking, each campaign had about 1MM test guests and 100-500K control guests
  - In addition, we had guest "randoms" (no selection criteria applied) of 75K test and control guests

- Our database allows us to associate a large percentage of in-store sales, nearly all of online sales, and fraction of online "cookie" browse behavior to our guests

Data available

Modeling datasets generally contained most or all these fields:

Purchase RFM

- In-store and online $, units, and trips for total store/site and 50+ product categories (summarized over time periods: 0-3, 4-6, 0-12, 13-24 months)

Browse RFM

- High-level online categories browsed and frequency in previous month

Demographics / other guest info

- Age, income, children in HH, distance to nearest store, active Target proprietary card, multi-channel guest, coupon redemption history, etc.

Contact history

- Number of coupons redeemed and average incentive values of direct mail, POS marketing, and email contacts received in last 0-3 months

Model development and selection

Model development

- For methods 1-4, the analytics team used SQL, SAS Enterprise Guide, and SAS Enterprise Miner to extract and prepare the data, plus develop our predictive models

- In general, we used regression, logistic regression, and decision trees to build our incremental sales models

Selecting the best model

- Models were built on a 50% train sample and cross-validated on a 50% validation sample
  - Later, models were applied to a similar "out of sample" direct mail campaign to see if results were consistent

- Lift metrics and charts were developed to determine if guests with top scores generated the highest incremental $/guest
  - Success meant that higher scores correlated with higher actual incremental $

Validating model results with lift charts

ISM score quartile	Test or Control	Number of Guests	Average ISM score	Actual Incr $ per Guest
top 20%	Test	15,000	$9.34	$8.06
top 20%	Control	7,500	$9.36	
20-40%	Test	15,000	$2.35	$2.41
20-40%	Control	7,500	$2.31	
40-60%	Test	15,000	$0.91	$0.04
40-60%	Control	7,500	$0.93	
60-80%	Test	15,000	-$0.45	-$1.07
60-80%	Control	7,500	-$0.47	
bottom 20%	Test	15,000	-$6.01	-$7.23
bottom 20%	Control	7,500	-$5.98	

Note: Fictional data for illustrative purposes only

GOOD!

- Higher ISM scores correlate with higher actual incr $/guest.
- Guests with top 20% ISM scores generate higher actual incr $ than do the next 20%.

Validating model results with lift charts

ISM score quartile	Test or Control	Number of Guests	Average ISM score	Actual Incr $ per Guest
top 20%	Test	15,000	$9.34	$1.01
top 20%	Control	7,500	$9.36	
20-40%	Test	15,000	$2.35	-$1.84
20-40%	Control	7,500	$2.31	
40-60%	Test	15,000	$0.91	$0.95
40-60%	Control	7,500	$0.93	
60-80%	Test	15,000	-$0.45	$6.31
60-80%	Control	7,500	-$0.47	
bottom 20%	Test	15,000	-$6.01	-$1.22
bottom 20%	Control	7,500	-$5.98	

Note: Fictional data for illustrative purposes only

BAD!

- Higher ISM scores do not correlate with higher actual incr $/guest.
- Guests with top 20% ISM scores do not generate higher actual incr $ than the next 20%.

Model Results

Model results

- Lift charts built on <u>in-sample validation</u> for more than 10 direct mail campaigns showed the following trends

(lift chart results ranked <u>best to worst</u>)

1. Two-stage ISM
2. Likelihood to shop ISM
3. Estimated spend ISM

Model results

- For in-sample results, the Two-stage ISM and Likelihood to Shop ISM generally showed:
  - Training sample: Fairly consistent rank ordering of the deciles and poor prediction of actual incremental sales values by decile
  - In-sample validation: Somewhat consistent rank ordering of the deciles and poor prediction of actual incremental sales
- However, out-of-sample validation results were generally poor
  - Out-of-sample validation was performed by applying the same ISM to a similar direct mail campaign during a different time period to see if results would still hold

Grocery direct mail: *TRAINING*

Results for "random" segment (0-12 month store purchasers)

Score rank	Average score INCR P(shop)	Average INCR $/guest
Top 10%	77%	**$6.85**
11-20%	70%	**$2.63**
21-30%	66%	**$4.78**
31-40%	61%	**$0.58**
41-50%	53%	**-$0.64**
51-60%	43%	**-$0.85**
61-70%	37%	**$4.51**
71-80%	33%	**-$0.40**
81-90%	29%	**-$5.64**
Bottom 10%	20%	**-$1.70**

- Rank ordering of INCR $/guest isn't perfect, but directionally OK
- More negatives in lower deciles, higher values in the top deciles
- Surprisingly high results in the 61-70th percentile

Note: Results are fictional, but are directionally consistent with actual results

Grocery direct mail: *In-sample VALIDATION*

Results for "random" segment (0-12 month store purchasers)

Score rank	Average score INCR P(shop)	Average INCR $/guest
Top 10%	77%	**$5.60**
11-20%	70%	**$1.55**
21-30%	66%	**$2.70**
31-40%	61%	**$1.45**
41-50%	53%	**-$2.56**
51-60%	43%	**-$0.48**
61-70%	37%	**-$0.38**
71-80%	33%	**-$1.16**
81-90%	29%	**-$0.38**
Bottom 10%	20%	**-$1.76**

- Similar to previous lift chart, the rank ordering of INCR $/guest isn't perfect, but OK
- All negatives in lower 60%, and all positive values in the top 40%
- Very low results in 41-50th percentile

Note: Results are fictional, but are directionally consistent with actual results

Grocery direct mail: *Out-of-sample VALIDATION*

Results for "random" segment (0-12 month store purchasers)

Score rank	Average score INCR P(shop)	Average INCR $/guest
Top 10%	75%	**$1.60**
11-20%	68%	**-$0.69**
21-30%	67%	**-$3.65**
31-40%	62%	**-$0.07**
41-50%	53%	**-$0.26**
51-60%	46%	**-$0.22**
61-70%	34%	**$0.26**
71-80%	27%	**$0.78**
81-90%	19%	**$1.43**
Bottom 10%	20%	**$1.45**

- Results significantly worse than previous in-sample lift charts
- While the top 10% performs well, the model does not hold rank order
- Because of these results, we are unwilling to use the model for actual guest selection

Note: Results are fictional, but are directionally consistent with actual results

Hypotheses: Why our modeling didn’t work

Hypotheses

- Data availability: *We don't have the right data.*
  - Economic indicators (e.g. consumer confidence index)
  - Share of wallet (e.g. % of grocery $ spent at Target)
  - Attitudinal information (e.g. price-sensitive guest)

- Modeling methodology: *We're not using the right method.*
  - Key assumptions are being overlooked
  - Continue work on developing the "decision tree" method

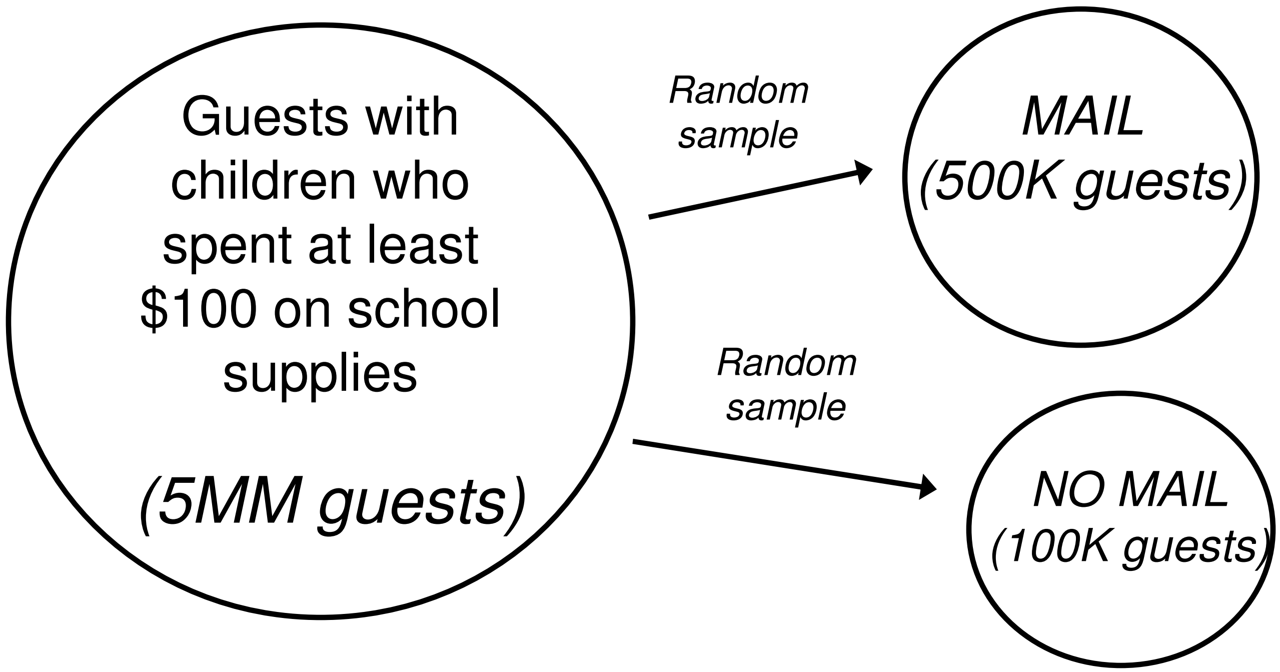
- Problem too complex: *Cannot predict with any more accuracy.*
  - Even with a more data and a robust methodology, perhaps the problem cannot be practically solved

Questions?

Appendix

How we design campaigns

- The marketing team gives the campaign execution team guest selection criteria for each of the segments



- After determining which guests fit the criteria, a random sample of guests is placed into a MAIL (i.e. test) segment, and another in a NO MAIL (i.e. control) segment

Model development

- The team tried several variations on modeling techniques as well as variable transformations:
  1. Building segmented models (e.g. high spenders vs. low)
  2. Using de-seasonalized independent and dependent variables
  3. Inference models for variables with missing values
  4. Modeled on variations of the target variable
     - Promo period $ spent
     - Guest shopped in promo period (Yes / No)
     - Percentiles of promo period $ spent
     - Transformations of promo period $ spent (log, sqrt)
     - Two-stage modeling: P(shopping) * estimated $ spent

Influential variables in ISMs

Predictor variables commonly found in our incremental sales models:

- 0-12 month guest spend and trips to Target stores
- Whether or not a guest shops online at Target.com
- Spend, trip, or browse variables to related product categories (e.g. Essentials/grocery spend for TargetMail campaigns)
- Change in guest spend and/or trips 0-6 vs. 7-12 months
- Days since last purchase date, i.e. recency
- Number of coupons redeemed in past 0-12 months
- RedCard holder status and Target.com online shopper status
- Tender type used in transactions
- Occasionally, a demographic variable (e.g. guest age) or a contact history variable (e.g. number of POS offers last month)