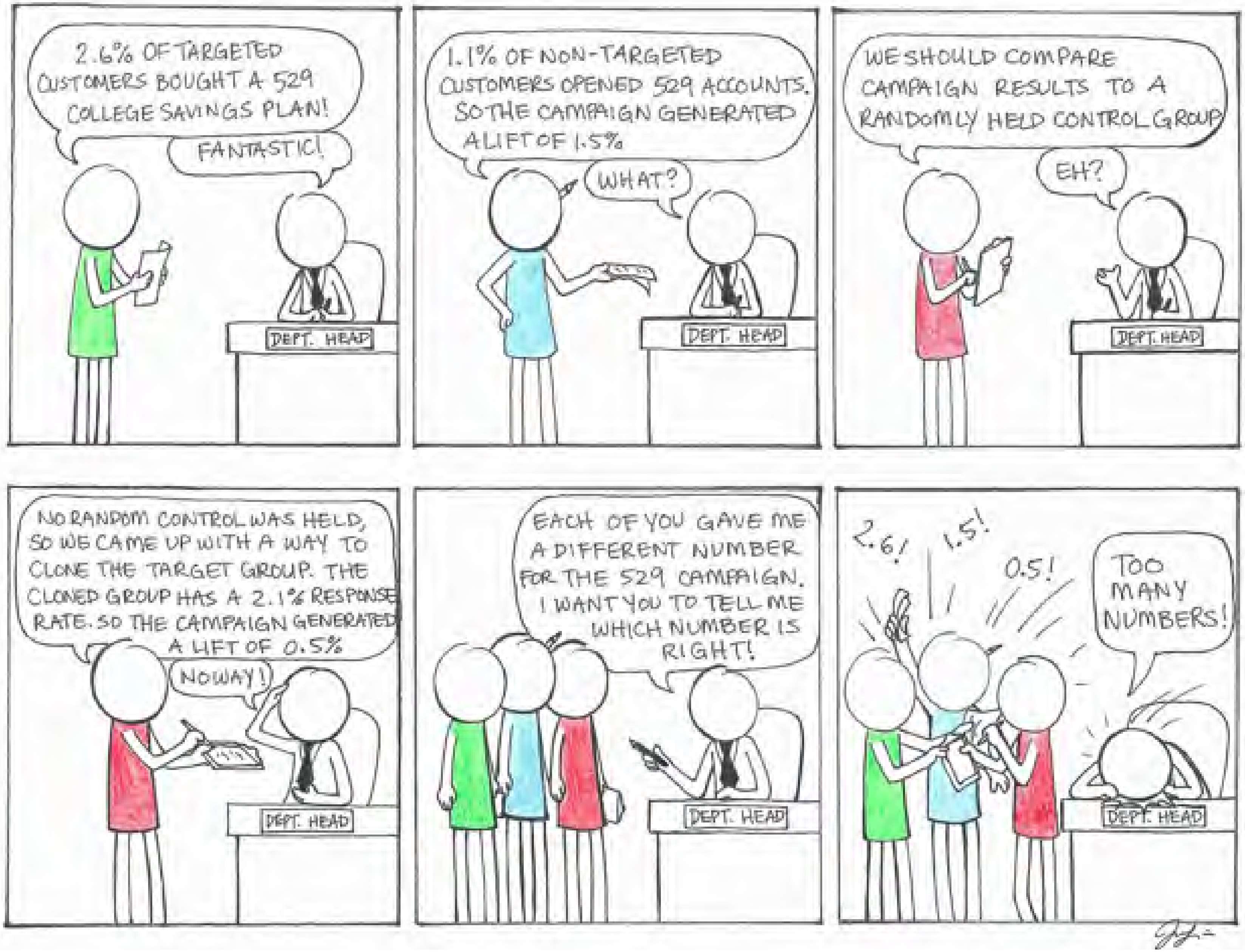
True-Lift Modeling: Mining for the Most Truly Responsive Customers and Prospects

Kathleen Kane
Jane Zheng
Victor Lo[1]
Alex Arias-Vargas

Fidelity Investments

[1]Also with Bentley University

Predictive Analytics World
New York City
October 19-20 2011


2.6% OF TARGETED CUSTOMERS BOUGHT A 529 COLLEGE SAVINGS PLAN!
FANTASTIC!
DEPT. HEAD
1.1% OF NON-TARGETED CUSTOMERS OPENED 529 ACCOUNTS. SO THE CAMPAIGN GENERATED A LIFT OF 1.5%
WHAT?
DEPT. HEAD
WE SHOULD COMPARE CAMPAIGN RESULTS TO A RANDOMLY HELD CONTROL GROUP
EH?
DEPT. HEAD
NO RANDOM CONTROL WAS HELD, SO WE CAME UP WITH A WAY TO CLONE THE TARGET GROUP. THE CLONED GROUP HAS A 2.1% RESPONSE RATE. SO THE CAMPAIGN GENERATED A LIFT OF 0.5%
NO WAY!
DEPT. HEAD
EACH OF YOU GAVE ME A DIFFERENT NUMBER FOR THE 529 CAMPAIGN. I WANT YOU TO TELL ME WHICH NUMBER IS RIGHT!
DEPT. HEAD
2.6!
1.5!
0.5!
TOO MANY NUMBERS!
DEPT. HEAD

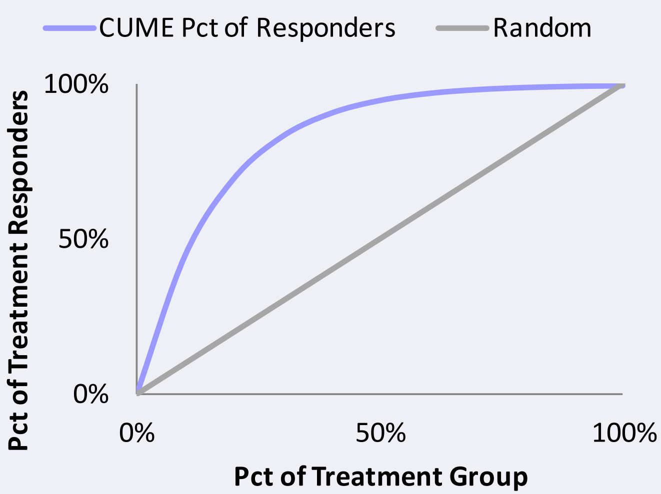
Stop spending direct marketing dollars on customers who would purchase anyway!

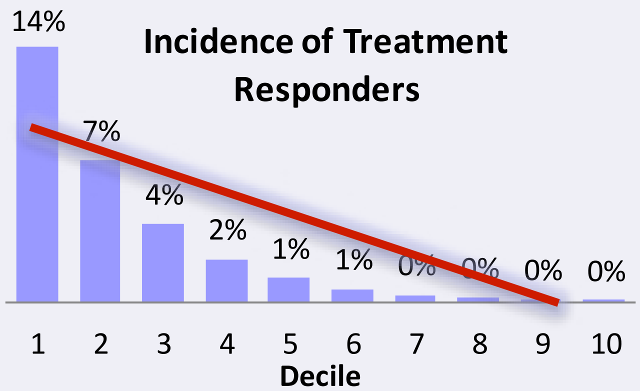
- True-lift modeling can identify:
  - which customers will purchase without receiving a marketing contact
  - which customers need a direct marketing nudge to make a purchase
  - which customers have a negative reaction to marketing (and purchase less if contacted)
- This discussion will describe:
  - the basic requirements needed to succeed with true-lift modeling
  - scenarios where this modeling method is most applicable
  - the pros and cons of various approaches to true-lift modeling

Outline

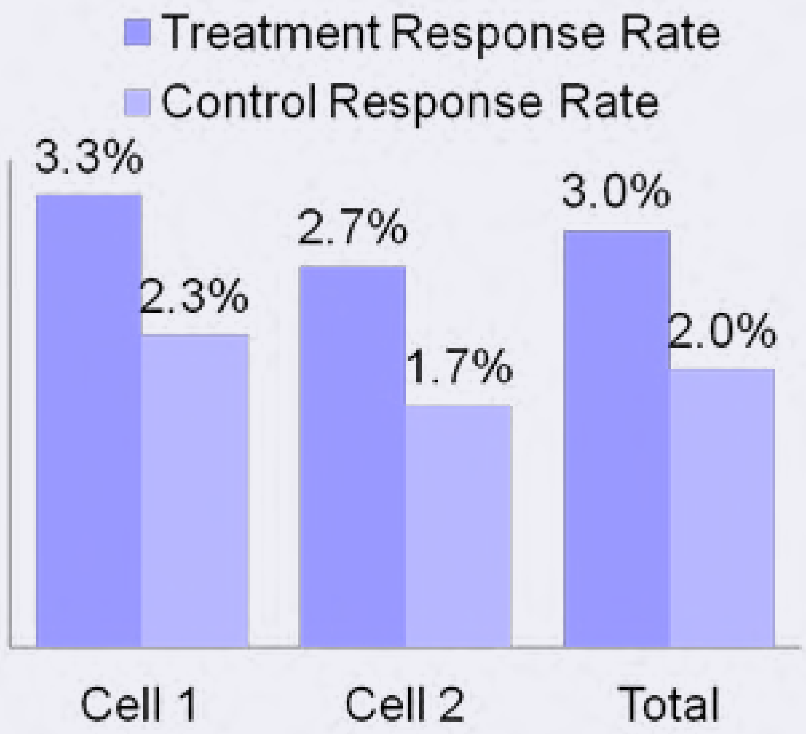
- Why do we need true-lift modeling? 10min
- What are the methods of true-lift modeling? 10min
- What is the context where true-lift modeling is most necessary & useful? 10min
- Questions 10min

What's wrong with this picture?



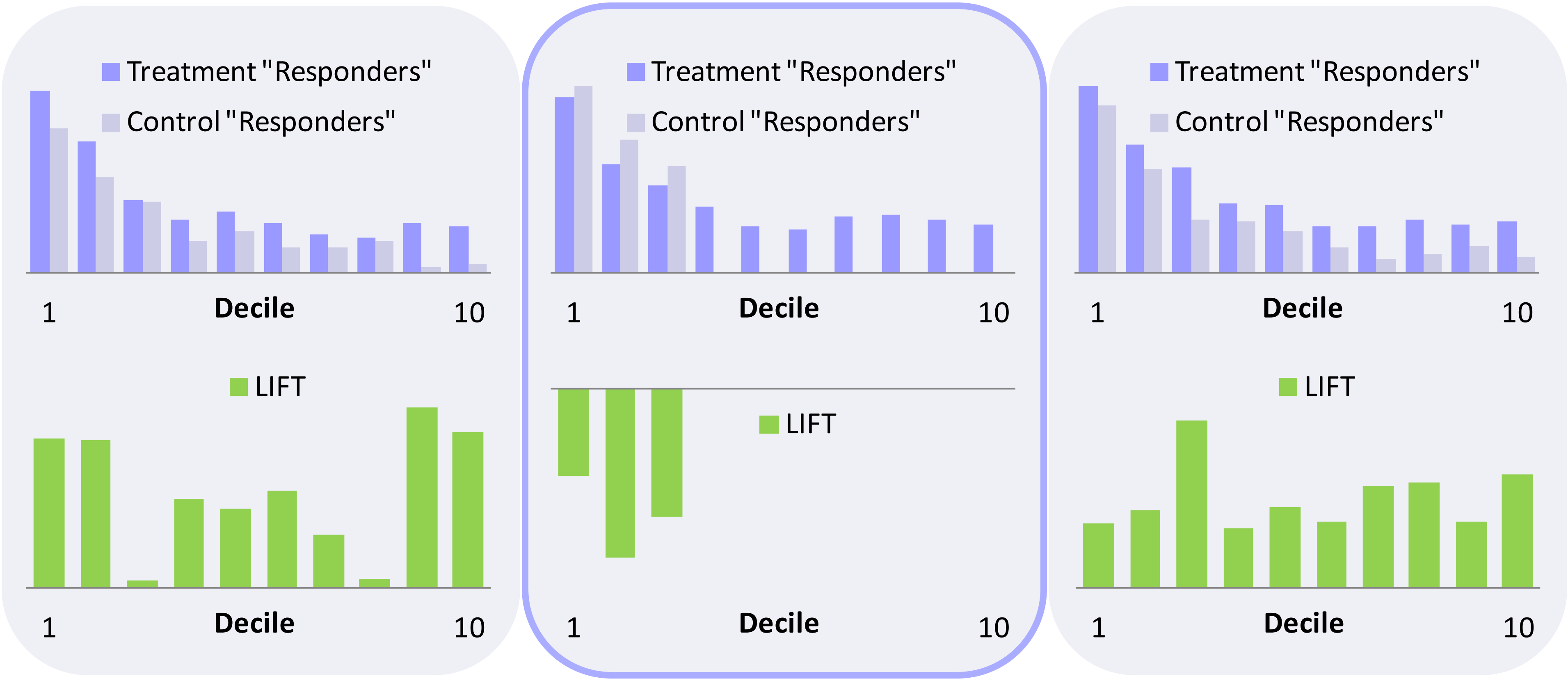


A successful response model



A successful marketing campaign

Measuring response models by lift over control



Simulated data

Why do we need true-lift modeling?

Look-alike model = find people who will take Action A

=P(A)

Standard response model = find people who will take Action A after receiving a treatment

=P(A | Treatment)

True-lift model = find people who will take Action A ***<u>only</u>*** after receiving a treatment

=P(A | Treatment)
– P(A | no Treatment)

- Standard response models often behave more like Look-alike models than like True-lift models
- Why spend marketing $$$ on people who would do Action A anyway?

Why do we need true-lift modeling?

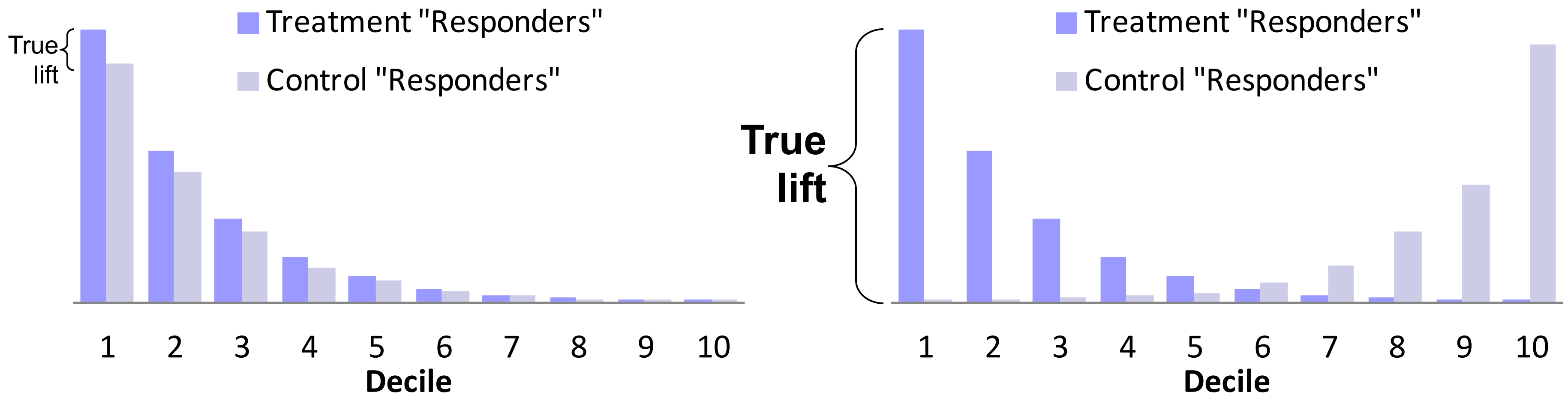
Look-alike model = find people who will take Action A

True-lift model = find people who will take Action A ***only*** after receiving a treatment

- **When is a look-alike model good enough?**
  - **Responders can only take Action A if they receive one unique marketing contact**
    - **Single channel**
    - **Single contact**
    - **No other way to take Action A**
- **When is a true-lift model needed?**
  - **Responders have many opportunities to take Action A**
    - **Multiple channels**
    - **Multiple contacts**

The True-Lift model objective

- Maximize the Treatment responders while minimizing the control "responders"



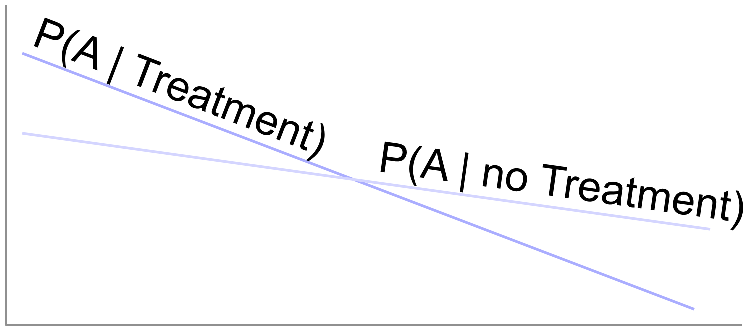
A standard response model **A True-Lift response model**

Simulated data

True-Lift model solutions

A. Difference of two models: Treatment – Control
B. Two sequential models: Treatment Actual – Control Prediction
C. Binned & Averaged dependent variable

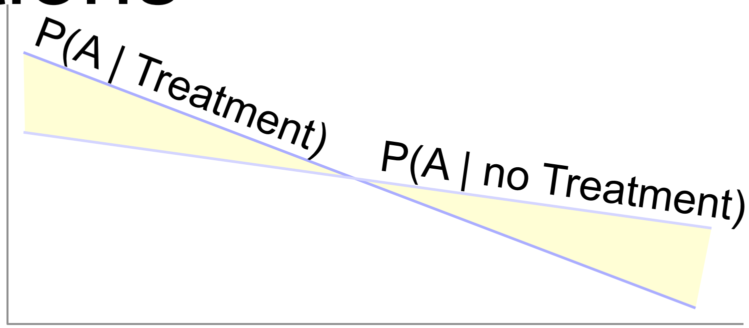
Solution A1: Difference of two models: Treatment - Control



- Model 1 predicts P(A | Treatment)
  - Dependent variable = Action A
  - Model Population = Treatment Group
- Model 2 predicts P(A | no Treatment)
  - Dependent variable = Action A
  - Model Population = Control Group
- Final prediction of lift =

  Model 1 Score – Model 2 Score
- Pros: simple concept, familiar execution (x2)
- Cons: indirectly models true-lift, the difference may be only noise, 2x the work, scales may not compare, 2x the error, variable reduction done on indirect dependent vars

Solution A2: Single combined model using Treatment interactions



- Model population = Treatment & Control together
- Dependent variable = Action A
- Independent variables are attributes x,y,z:
- Conceptually:

Lift

P(Action A) = P(A | not Treated) + P(A | *only* if Treated)

= {some coefficients} * {x,y,z} + 0/1 treatment flag * {some coefficients } * {x,y,z}

During model development, the interaction flag is 0 for control records and 1 for treatment records

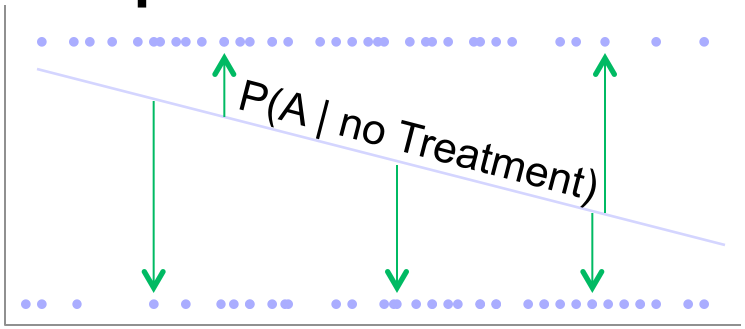
- Final prediction of lift = difference of two scores

  = Prob(response if Treated) – Prob(response if not Treated)

  = score with interaction flag set to 1 – score with interaction flag set to 0
- Pros: combined model minimizes compounded errors
- Cons: indirectly models true-lift; large number of independent terms; collinearity of terms; reduction needed; adding two model scores may compound errors

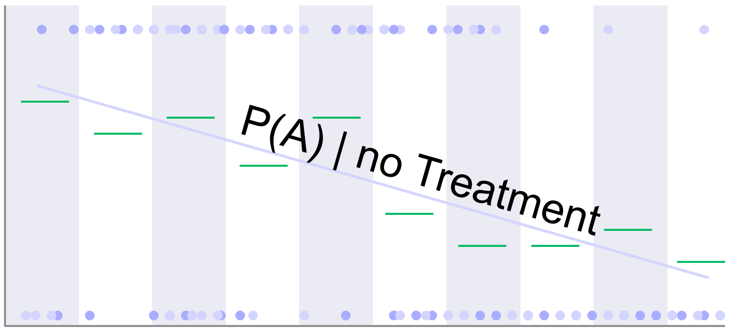
Lo, Victor S.Y. The True Lift Model - A Novel Data Mining Approach to Response Modeling in Database Marketing. SIGKDD Explorations, Volume 4, Issue 2, Dec 2002, p.78-86.

Solution B: Two sequential models: Treatment actual – Control prediction



- Model 1 predicts P(A | no Treatment)
  - Dependent variable = Action A
  - Model Population = Control Group
- Model 2 predicts P(A | Treatment) – P(A | no Treatment)
  - Dependent variable = Action A – Model 1 Score
  - Model Population = Treatment Group
- Final prediction of lift = Model 2 Score
- Pros: more directly models true-lift; identifies variables that are directly correlated with true-lift (some of which are drivers of lift)
- Cons: the Model 2 dependent variable contains Model 1 errors; 2x the work, Model 1 scores and Action A should (but might not) share the same scale

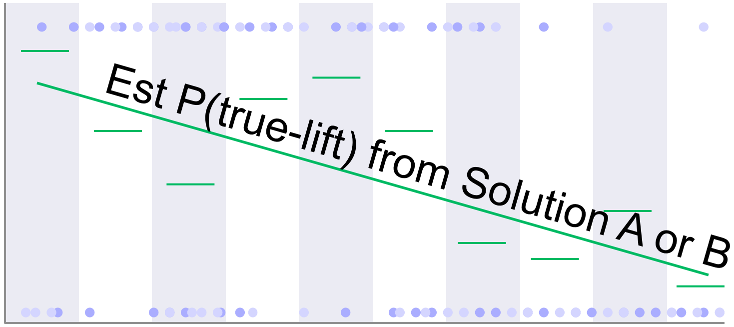
Solution C1: Binned & averaged dependent variable



- Model 1 predicts P(A | no Treatment)
  - Dependent variable = Action A
  - Model Population = Control Group
- Create N bins for Treatment & Control population together, ranked by Model 1 score (control "response")
- Calculate dependent variable value for each BIN:

  Treatment response rate – Control response rate
- [Could stop here, using the bin average lift as the predicted lift, or continue with]:
- Model 2 predicts actual average lift of each bin
  - Dependent variable = Average lift within each bin
  - Model Population = Treatment Group
- Final prediction of lift = Model 2 Score
- Pros: directly models true-lift; identifies variables that are directly correlated with true-lift (some of which are drivers of true-lift)
- Cons: 2X the work; the approach requires variation in average lift across bins (which might not exist); control response needs to be correlated to true-lift response

Solution C2: Solution A or B + binned & averaged dependent variable



- Complete Solution A or B first to rank-order observations by estimated lift
- Use Solution A/B model score to rank and bin the observations: create N bins for Treatment & Control population together, ranked by Solution A/B score
- Calculate dependent variable value for each BIN:

  Treatment response rate – Control response rate
- [Could stop here, using the bin average lift as the predicted lift, or continue with]:
- Model 3 predicts actual average lift of each bin
  - Dependent variable = Average lift within each bin
  - Model Population = Treatment Group
- Final prediction of lift = Model 3 Score
- Pros: directly models true-lift; this approach is more likely to maximize the variation in average lift across bins; identifies variables that are directly correlated with "lift" (some of which are drivers of lift)
- Cons: 3X the work

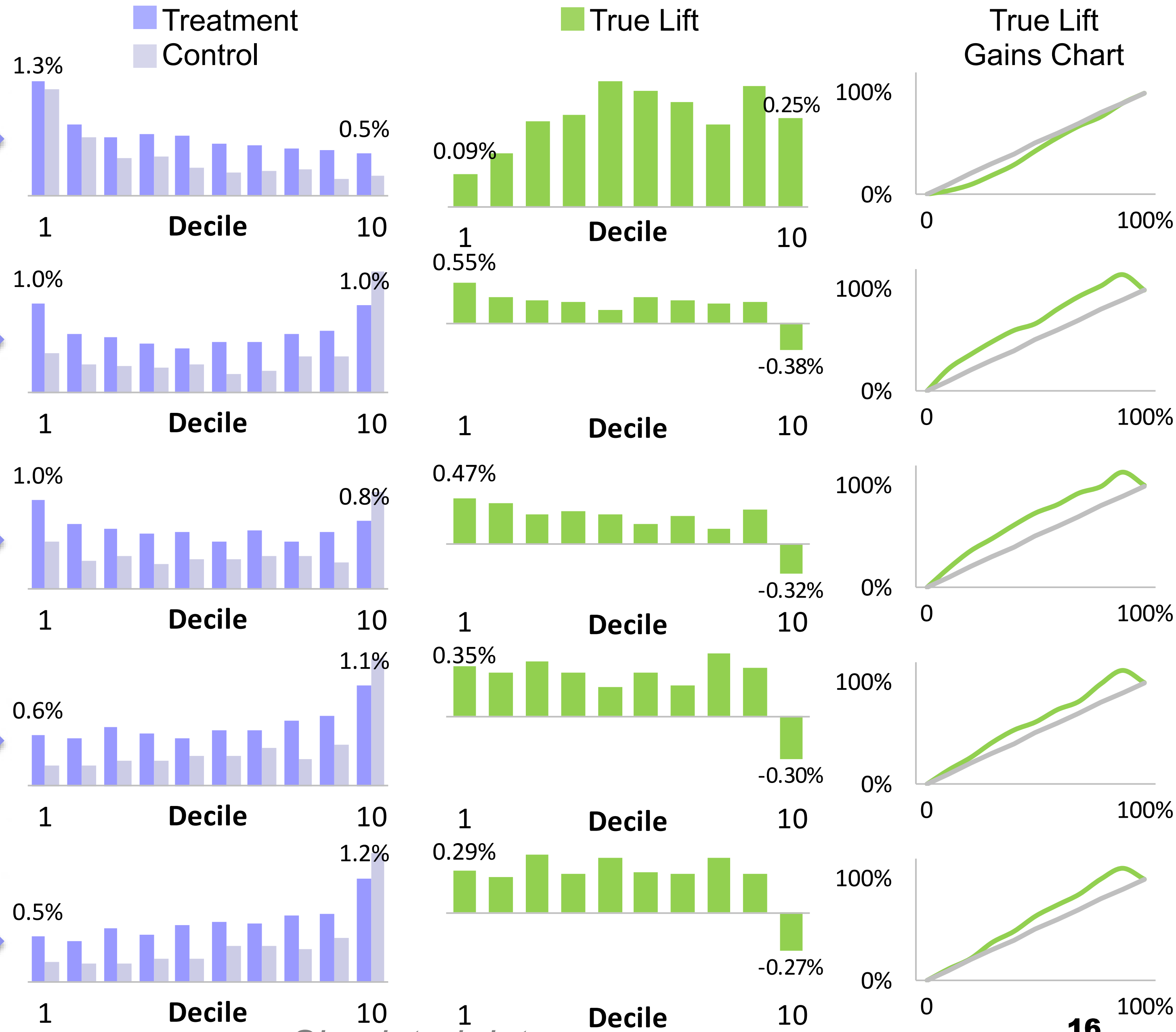
Standard response model

Solution A2: Single combined model with interactions

Solution B: Depvar = Treatment actual – Control prediction

Solution C1: Ranked & binned by Control model

Solution C2: Ranking & binned by Lift model



Simulated data

Other solutions, variations & applications

- Decision trees
- Clustering / K-nearest neighbor
- Bootstrapping
- Optimization

- Personalized medicine
- Other marketing situations (how to separate very similar groups who act differently)

Ideal conditions for true-lift modeling

- A randomized control group is withheld!
- Treatment does not cause all “responses”
- “Response” is not correlated to “lift” (i.e., response model is not good enough)
- Lift-to-noise ratio is large enough
- If overall lift is near 0, then you need pockets of both negative lift and positive lift
- Repeated campaigns, or at least test campaign precedes rollout

Q&A

Stop spending direct marketing dollars on customers who would purchase anyway!

- True-lift modeling can identify:
  - which customers will purchase without receiving a marketing contact
  - which customers need a direct marketing nudge to make a purchase
  - which customers have a negative reaction to marketing (and purchase less if contacted)
- This discussion will describe:
  - the basic requirements needed to succeed with true-lift modeling
  - scenarios where this modeling method is most applicable
  - the pros and cons of various approaches to true-lift modeling

Glossary

- "Response" = taking the desired action (Action A); might have done Action A whether treated or not
- True-lift = taking the desired action (Action A) *only* in response to the Treatment; would not have done Action A if not treated (aka uplift, net lift, incremental lift)