Predicting Coupon Effects on Consumer Buying Behavior in Absence of a Control Group

Dr. Gerald Fahner, FICO
August 24, 2011

Coupon Marketing in Retail

» Coupon campaigns are increasingly targeted:
  - » Rich scanner data/customer transaction histories
  - » Data-driven coupon assignment, optimization, 1-on-1 targeting

» Business objectives include:
  - » High redemption, raising product awareness, encouraging brand switching, rewarding loyal customers, reducing attrition, …
  - » Incremental metrics for basket size, store visits, sales volume, revenue, …

» Importance of aligning objectives, targeting and measurement
  - » Otherwise may give discounts to the wrong customers

» For incremental metrics, “uplift modeling” is a logical approach
  - » Rank order customers from least to most influenced by a coupon

Agenda

» Modeling Uplift and Causal Inference

» Case Study

» Discussion

Concept of “Uplift”

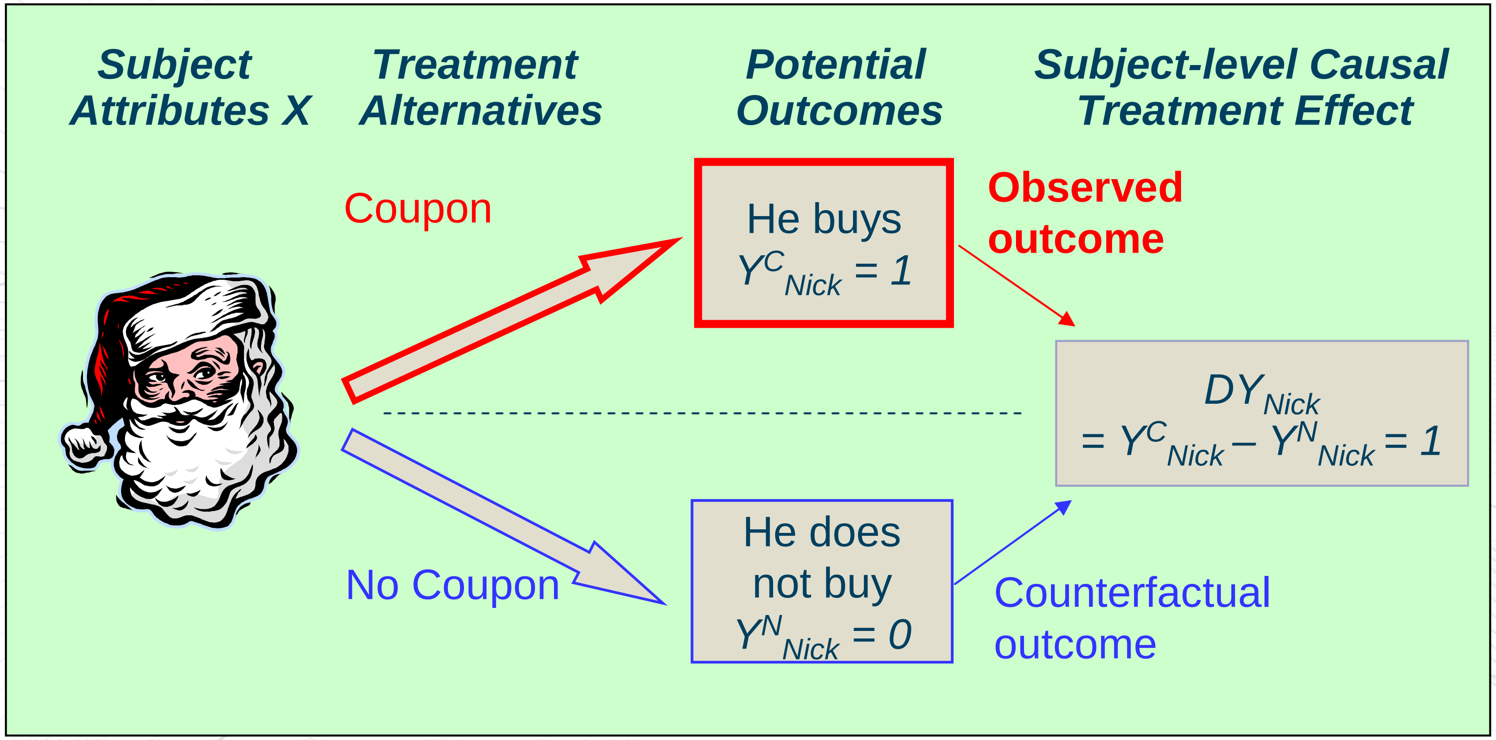
- Assume objective is to increase sales of a product
- An atomic marketing decision is whether to send Nick Smith a discount coupon. For this, we’re interested to predict differences such as:

Pr{Nick buys product | Send Nick a coupon}
– Pr {Nick buys product | Don’t send Nick a coupon}

- Idea of uplift modeling (Ratcliffe and Surry [1]) is to score subjects according to these differences
- With such a model we should be able to increment sales by targeting the top scorers

Relation to Causal Modeling

"Rubin Causal Model": A causal effect is a comparison of potential outcomes for alternative treatments defined on the same subject



Uplift modeling amounts to estimating $E[DY_i \mid X_i]$

Key difference to regression problem: DY_i is not observable

Conditions for Reliable Causal Inference Can Make Uplift Modeling a Treacherous Enterprise

- Estimation difficulties arise when conditions for causal inference regarding the historic treatment assignment process are violated:
  - Common Support
  - Unconfoundedness
  - Together: “Strongly ignorable treatment assignment” [2]

- Data from randomized experiments meet the conditions for data-driven causal inference
- Best practice for uplift modeling calls for experimental data
- But modelers often only have “business-as-usual” data where testing is very limited or non-existent [3]

Do Nonrandomized Data Close the Door on Uplift Modeling?

- » In fortunate situations, “natural experiments” come to rescue:
  - » Rubin and Waterman [4] estimated causal effects of pharmaceutical marketing on doctors’ prescriptions using techniques from observational studies to find doctors who can be most influenced
  - » Treatments were assigned subjectively (in haphazard ways) by a sales force
  - » This generated common support, while unconfoundedness assumption is questionable

- » What if historic treatments are assigned deterministically based on known attributes of the subjects?
  - » Rules-based
  - » Optimization-based

Agenda

» Modeling Uplift and Causal Inference

» Case Study

» Discussion

Coupon Campaign Process

Mail Matrix 10/20/2010*

	Ann	Bart	Cece	Dan	…
Ice Cream			X		
Toothpaste	X			X	
Lotion			X		
Beer				X	
…		X			…

» Customers were targeted with customized sets of product coupons on a campaign date

» Outcomes (product sales) were measured over several weeks thereafter

*Examples and figures for illustration purposes only

Coupon Campaign Process

Mail Matrix 10/20/2010*

	Ann	Bart	Cece	Dan	…
Ice Cream			X		
Toothpaste	X			X	
Lotion			X		
Beer				X	
…		X			…

- » Mail matrix was generated via constrained optimization
  - » Based on estimates of consumers' latent interests in products
  - » Based on some additional rules
  - » Constraints limited product volumes and coupons per customer
- » No randomized control groups were generated

*Examples and figures for illustration purposes only

Defining the Treatment Space

- Coupon assignments are highly customized and each customer receives multiple coupons
  - Tens of thousands of combinations in play
- To simplify, consider each product coupon as a binary treatment
  - Assume that effects of different coupons stand on their own

Approach

» Motivated by [4], estimate coupon effects on sales as follows:

Repeat for each coupon

1. <u>Propensity score-based matching</u>:

   Assess common support and generate matched pairs of similar customers who received coupon/no coupon

   Where successful, proceed to 2

2. <u>Nonparametric estimation of customer-level causal effects</u>:

   Differentiate observed outcomes between matched pairs to obtain DY_i ; $i = 1,\ldots,$ *# Matched customers*

3. <u>Regress DY_i on customer attributes</u>:

   Fit a model for $DYhat(X) = E[DY \mid X]$

Selection Bias in Observed Data*

» Optimization formulation prefers to assign coupons to customers with high latent interest scores for the couponed products

  » Customers who received a certain coupon tend to differ in certain respects from those who didn't receive it (selection bias)

Coupon: TOOTHPASTE*	**Means by Treatment Group (Observed Sample)**	
Treatment ***Buy-o-graphic attributes (X)***	***Coupon*** ***(N = 299,882)***	***No Coupon*** ***(N = 1,832,423)***
Units of shampoo purchased during 6 months prior to campaign date	1.23	0.71
Recency (months) of last toothpaste purchase as of campaign date	1.35	3.12
Gender	47% males	53% males
Total amount spent last 12 months	$887	$391
... many other attributes …	…	…

*Examples and figures for illustration purposes only

Greatly Mitigated Selection Bias in Matched Data Set

FICO™

Coupon: TOOTHPASTE*	**Means by Treatment Group (Matched Sample)**	
Treatment ***Buy-o-graphic attributes (X)***	***Coupon*** ***(N = 194,776)***	***No Coupon*** ***(N = 194,776)***
Units of shampoo purchased during 6 months prior to campaign date	0.97	1.03
Recency (months) of last toothpaste purchase as of campaign date	2.51	2.63
Gender	48% males	49% males
Total amount spent last 12 months	$612	$588
... many other attributes …	…	…

» Comparing average outcomes (i.e. Y = Sales of couponed product following the campaign date) across matched Coupon/No Coupon samples provides an aggregate measure of coupon success

» Average Treatment Effect of Coupon on Sales can be defined as:

$$ATE_{matched} = Mean_{matched}(Y \mid Coupon) - Mean_{matched}(Y \mid No\ Coupon)$$

*Examples and figures for illustration purposes only

Nonparametric Estimation of Customer-Level Causal Effects and Uplift Modeling

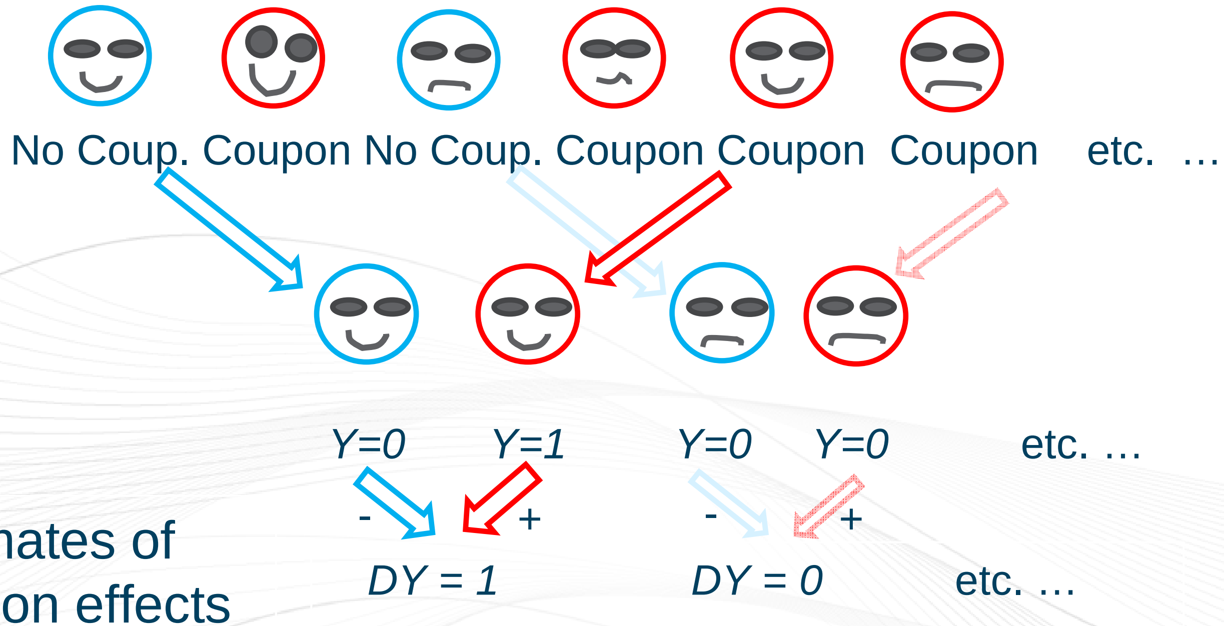
FICO

Customers and their treatments in observed sample

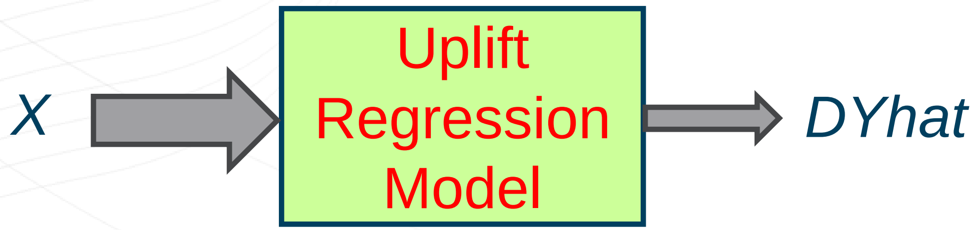
Customer pairs in matched sample

Observed outcomes

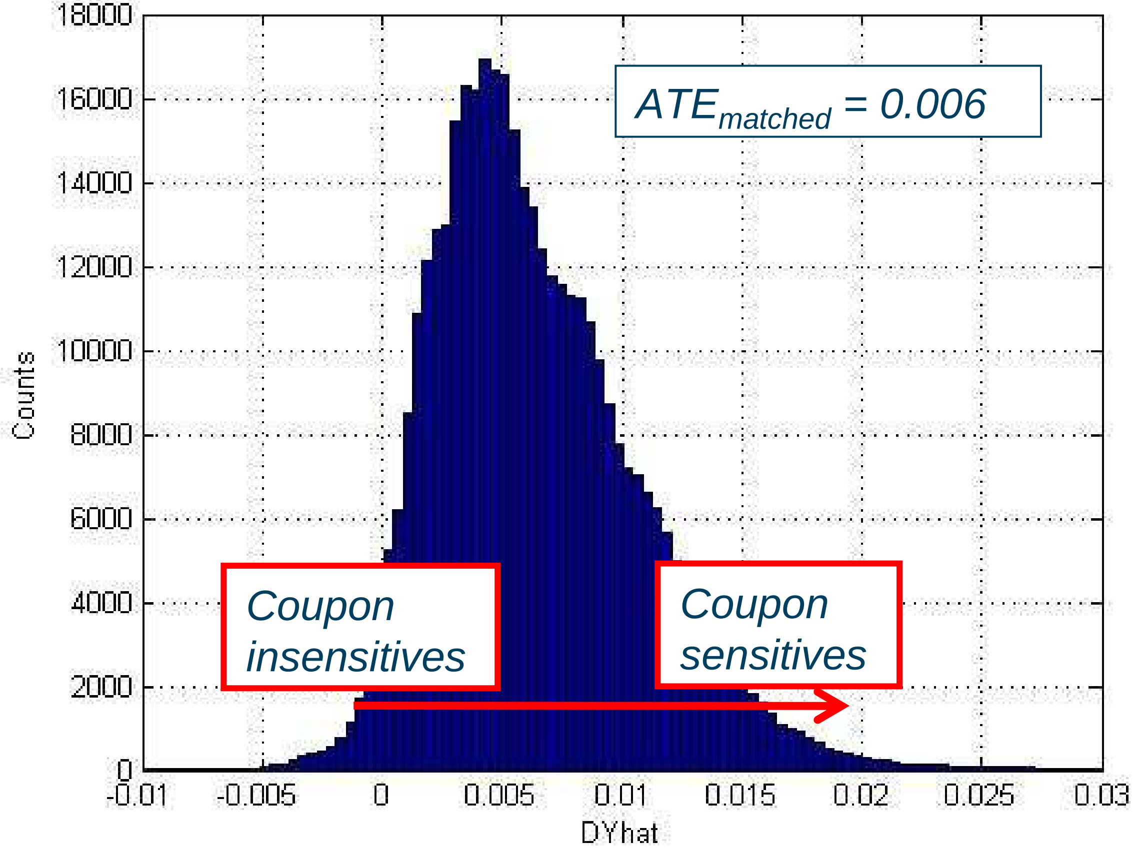
Nonparametric estimates of customer-level coupon effects



» Final step is to regress (smooth) *DY on X* (customer attribute vectors), resulting in a model for customer-level coupon effects *DYhat(X)*



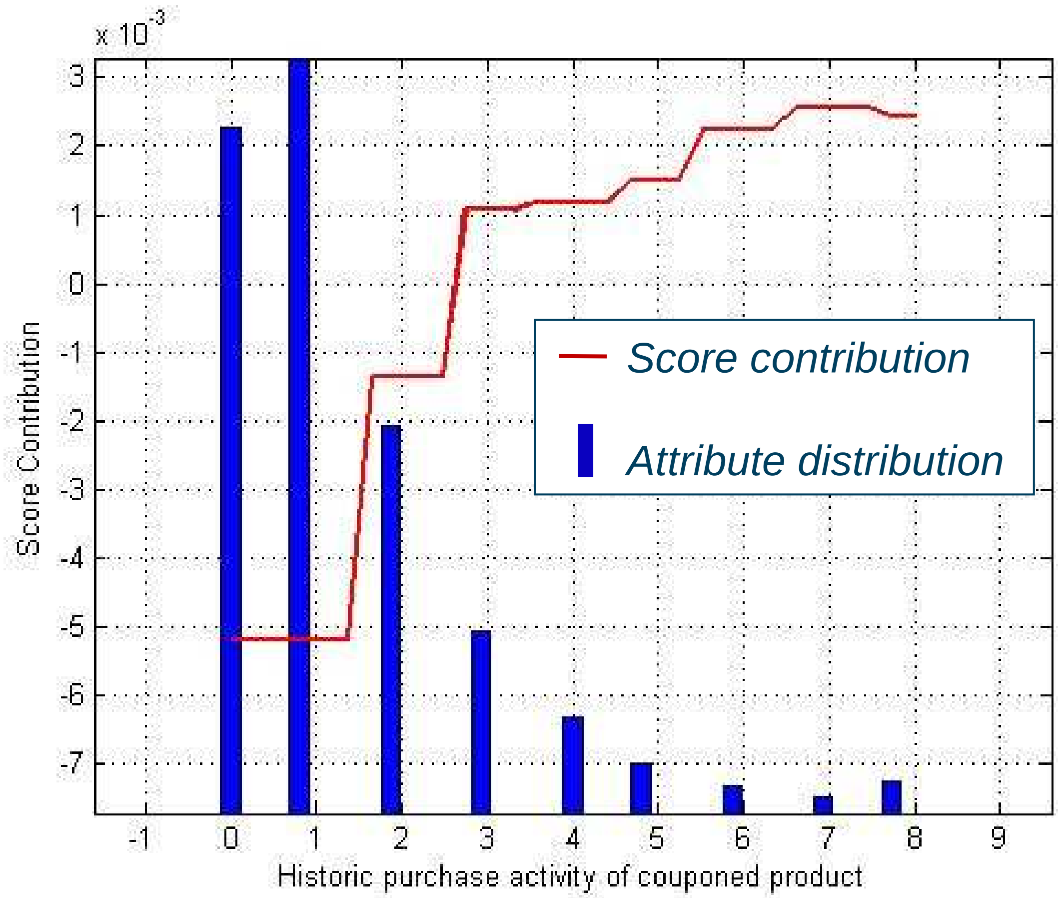
Smoothed Coupon Effects are Heterogeneous Across Matched Sample



» *ATE* indicates effectiveness of this coupon on the aggregate level

» Uplift model reveals substantial differences in coupon effects on sales across different customer types

Insights From Generalized Additive Uplift Model



» Moderate/frequent buyers of a certain product increase their buying substantially due to coupon. Seldom/non-buyers are hardly influenced

» Model-based prediction of uplift and insights gained from these relationships provide opportunities for targeting

Discussion

FICO

» As a transparent uplift modelling approach consider propensity score-based matching followed by nonparametric estimation followed by regression smoothing

» But no matter what the modelling approach is, uplift modelling with business-as-usual data is at the mercy of natural experiments

» What apparently helped in our retail project was the complex constrained coupon assignment which assigned similar customer types to different coupon treatments

» Similar modeling strategies can be applied to credit-related treatments such as loan pricing or fee discounts – best in conjunction with business-apt experiments!

References

[1] "Differential Response Analysis: Modeling True Response by Isolating the Effect of a Single Action", by N. J. Radcliffe and P. D. Surry. *Proceedings of Credit Scoring and Credit Control VI*. Credit Research Centre, University of Edinburgh Management School 1999.

[2] "The Central Role of the Propensity Score in Observational Studies for Causal Effects", by P. R. Rosenbaum and D. B. Rubin. *Biometrika*, Vol. 70, No. 1. (Apr., 1983), pp. 41-55.

[3] "Estimating Causal Effects of Credit Decisions", by G. Fahner, *International Journal of Forecasting* (2011). Article in press, http://caocao.myipcn.org/science/article/pii/S0169207011000082

[4] "Estimating Causal Effects of Marketing Interventions Using Propensity Score Methodology", by D. B. Rubin and R. P. Waterman. *Statistical Science, Vol. 21, No. 2*, (2006), pp. 206–222.

