Effective Methods for Measuring and Evaluating the Success of Newly Implemented Pricing Strategies

Leo Guelman, Ph.D., P.Stat.
Royal Bank of Canada
RBC Insurance

CAS 2015 Annual Meeting
Philadelphia, PA.
November 15 -18, 2015

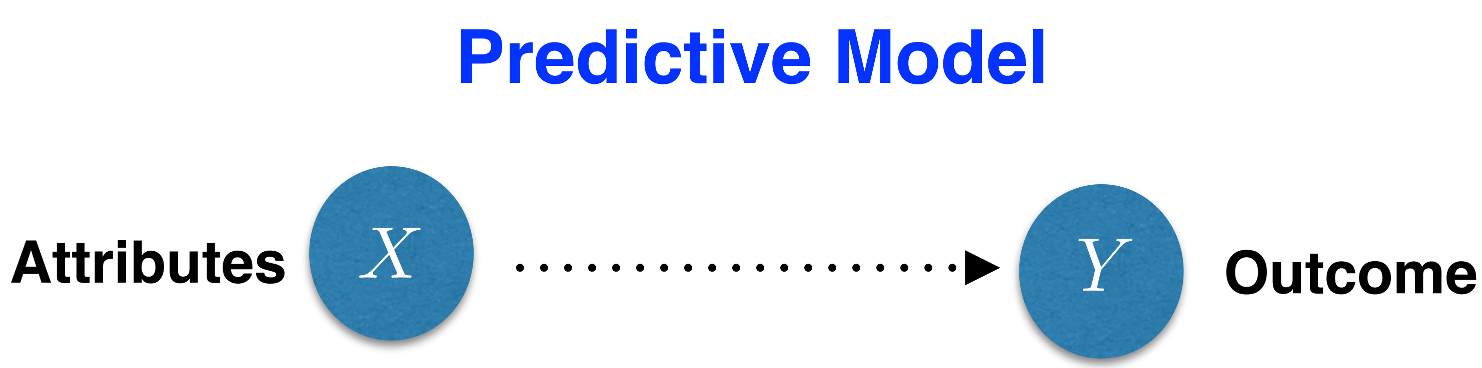
Agenda

- From Predictive to Causal Modeling.
- Causal Modeling with Experimental Data.
- Causal Modeling with Observational Data.
- Example: P&C Economic Price-Optimization.

From Predictive to Causal Modeling

Predictive Modeling has been established as a core strategic capability for many top insurers.

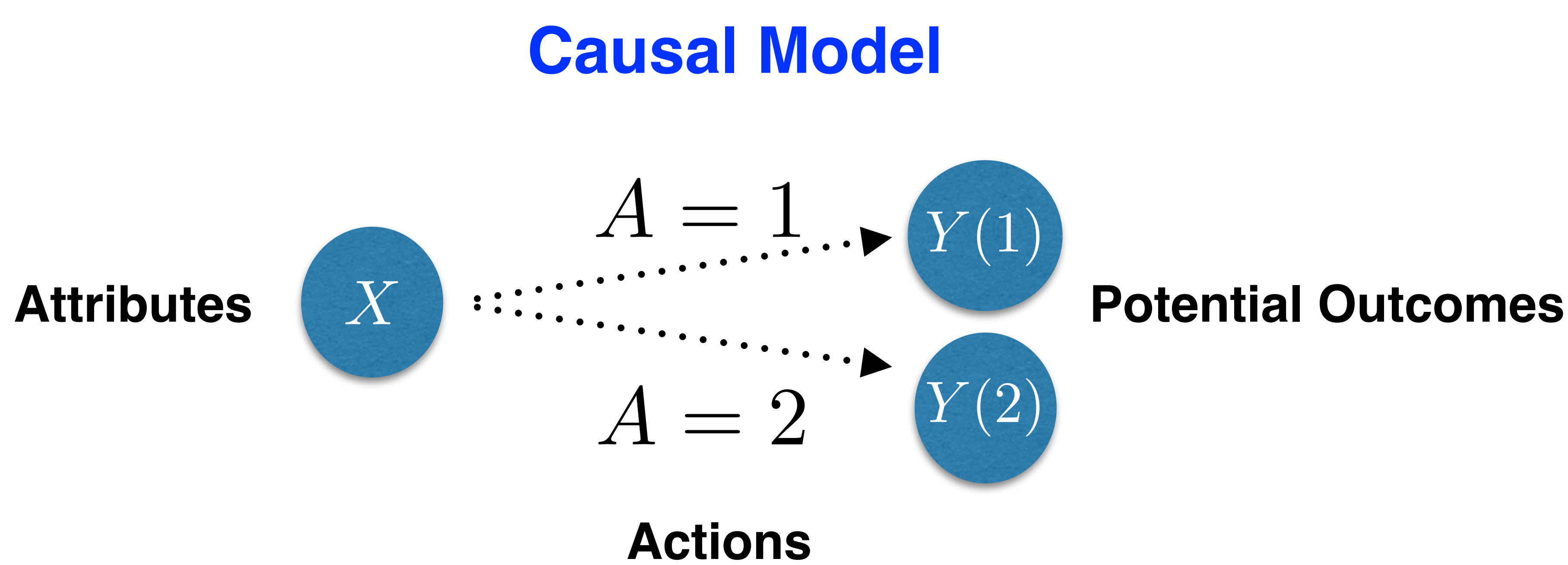
Goal: to predict an outcome variable using a collection of attributes under **"business as usual" conditions**.



From Predictive to Causal Modeling

Causal Modeling goes a step further relative to predictive modeling.

Goal: to predict an outcome under **changing conditions** — e.g., induced by alternative actions or "treatments".



Key Questions Faced by Decision Makers

Question	Estimand
Does the action affect the outcome?	**Average Treatment Effect (ATE)**
Does the action affect the outcome differently for different customer types?	**Subgroup Treatment Effect (STE)**
What is the impact of the action at the individual customer-level?	**Individual Treatment Effect (ITE)**

The **ITE** is an *unobserved* quantity, as a customer can never be observed simultaneously under more than one action — this is known as the **“fundamental problem of causal inference”**.

Causal Modeling with Experimental Data

- **Randomization**: the "gold standard" for scientific research.

  1. Randomly sample subjects from the population.

  2. Randomly assign subjects to treatment and control conditions.

  3. Estimate the ATE:

$$\text{ATE} = E[Y|\text{Action 1}] - E[Y|\text{Action 2}]$$

- As the sample size grows, the client attributes **X** will tend to be "balanced" between treatment and control groups — subjects become **"exchangeable"**.

- Even in an experimental setup, much can go wrong which requires statistical correction (Rubin, 2003).

Example: Price-Elasticity Estimation

- **Objective:** Estimate the impact of alternative rate increases on the portfolio's renewal rate.
- **Rate increase plays the role of the treatment** (e.g., 5% vs. 10%), and the response represents the **renewal outcome** (Y/N).
- Under randomized assignment of policyholders to rate changes, the **ATE (price-elasticity here)** can be computed straightforwardly.

Rate Change	+5%	+10%
N (policies)	10,000	10,000
Retained policies	9,200	8,700
Retention Rate	**92%**	**87%**

→ ATE = 87% - 92% = (5%)

Experimentation: Challenges and Threats

- Most insurance data come from “business as usual” conditions:
    - High costs associated with experimentation.
    - Legal and/or regulatory constraints.
    - Violation to ethical standards.
    - Lack of planning.
- In the absence of randomization, subjects are no longer exchangeable and thus direct comparisons can be misleading (Rosenbaum and Rubin, 1983).
- Assume the following extreme scenario (“**X**” indicates NO available data):

Rate Change	+5%	+10%
Age < 25 yrs	✓	✓
Age >= 25 yrs	✓	X

>= 25 yr old clients exposed to a 5% rate increase do not have a counterfactual under the 10% rate change group.

Regression-based Estimation: What Can Go Wrong?

- Need to adjust any difference in the outcome for differences in the client attributes under alternative treatments.

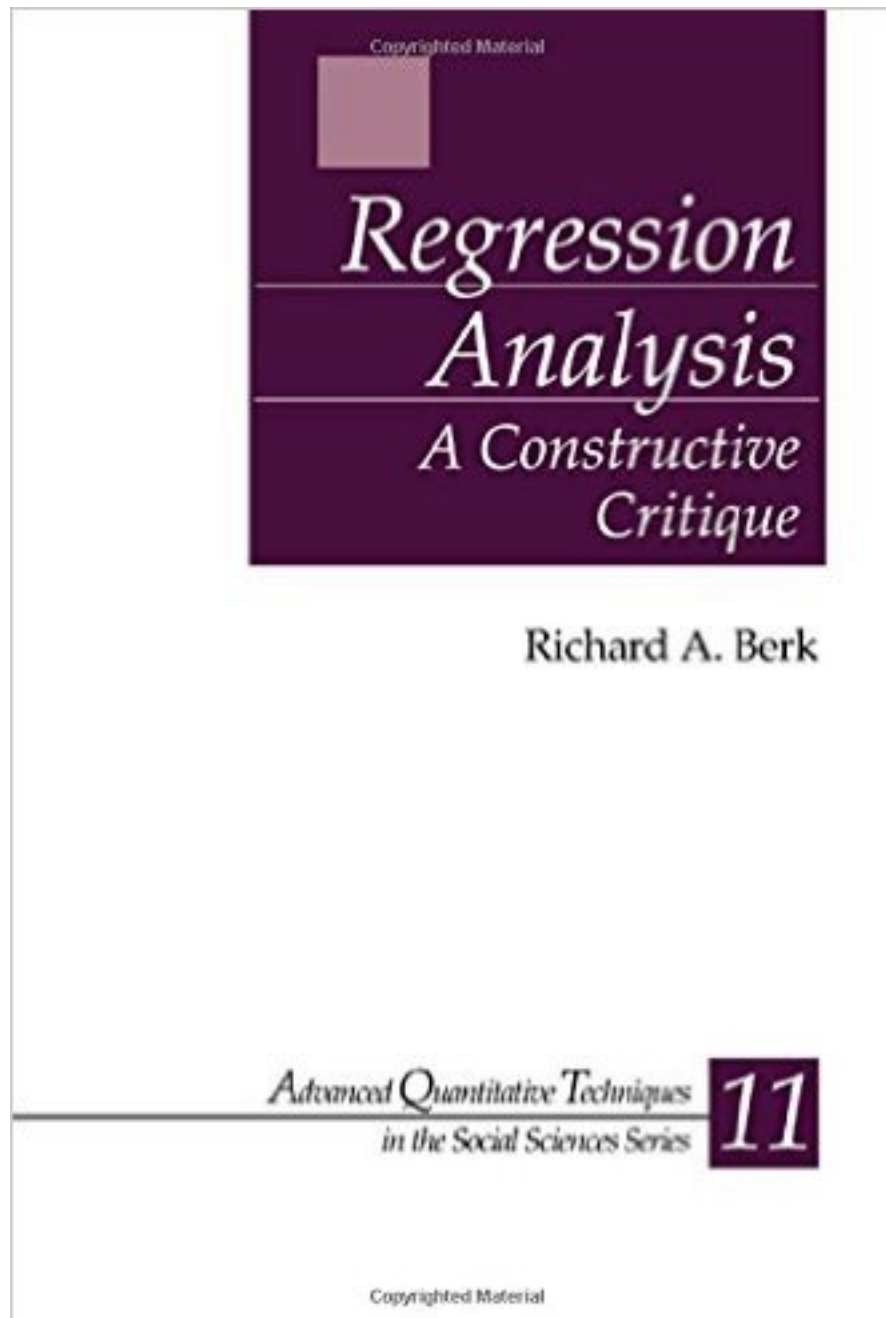
- The **standard regression approach** estimates:

$$Y = \alpha + \tau A + \beta X + \epsilon$$

where the slope τ of the treatment indicator is an estimator of the average treatment effect.

- In the absence of experimental data, **the standard regression approach is unreliable** (Berk, 2004):
    - **Regression-based methods mask non-overlap problems**, and they extrapolate inferences in regions of the predictors where certain treatment haven't been observed.
    - The **problem is worse with a large number of predictors**, as we cannot easily see non-overlap problems.
    - **Standard statistical software can be deceptive:** no warnings about potential non-overlap issues.

More about various strains of regression abuse…



Causal Inference with Observational Data

- An **observational study** attempts to draw inferences about the effect of treatments in the absence of experimental data (a.k.a. observational data).

- Rosenbaum and Rubin (1983) proposed **propensity score matching** as a method to remove the bias in the estimation of treatment effects from observational data.

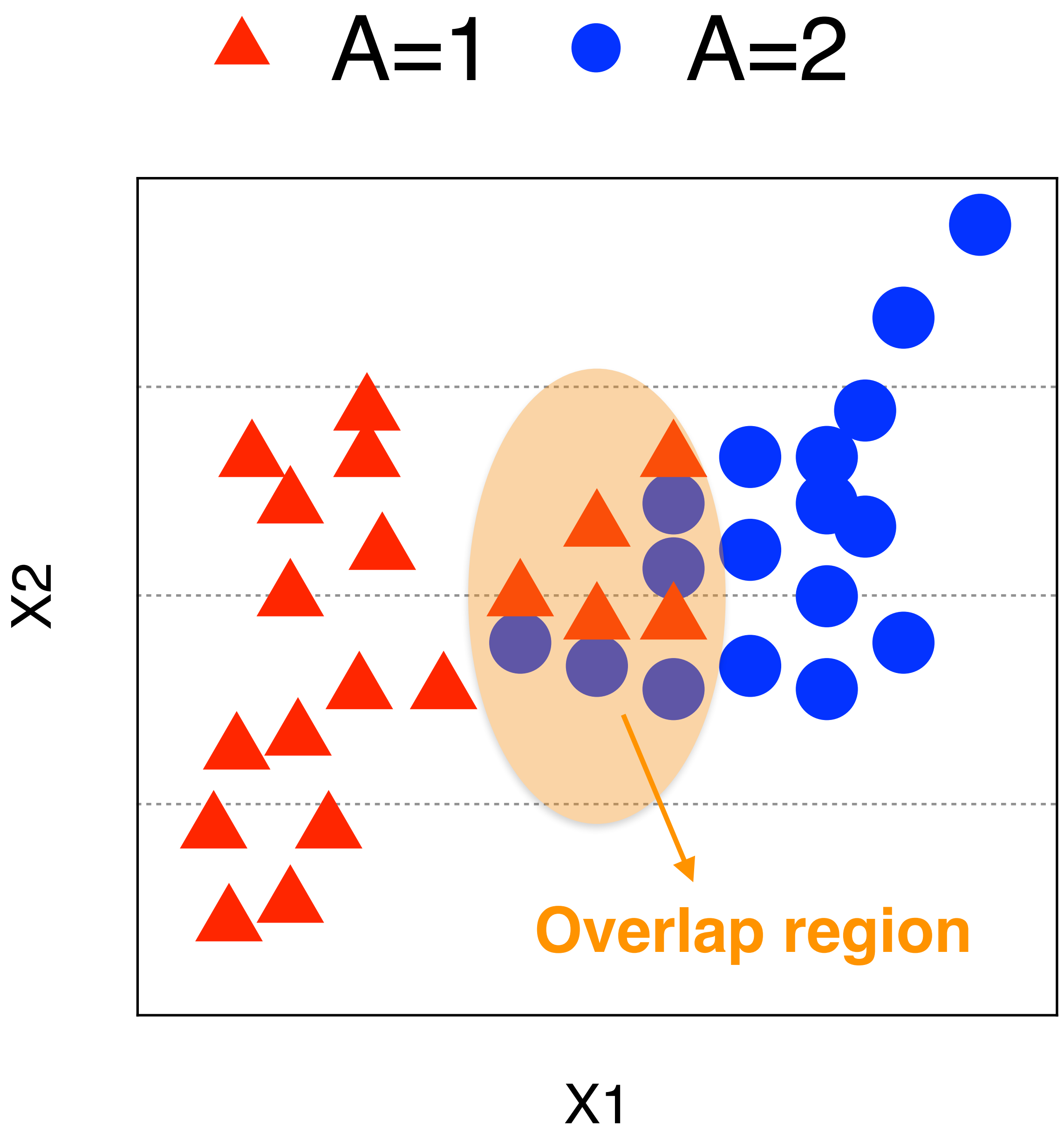
- These methods have become increasingly popular in a wide variety of fields (from economics to medicine).

- **Key concept:** Under certain data conditions, we can approximate a randomized experiment from observational data.

Key Data Conditions: **Common Support**

Common support (a.k.a. overlap) requires that similar customers were exposed to different actions.

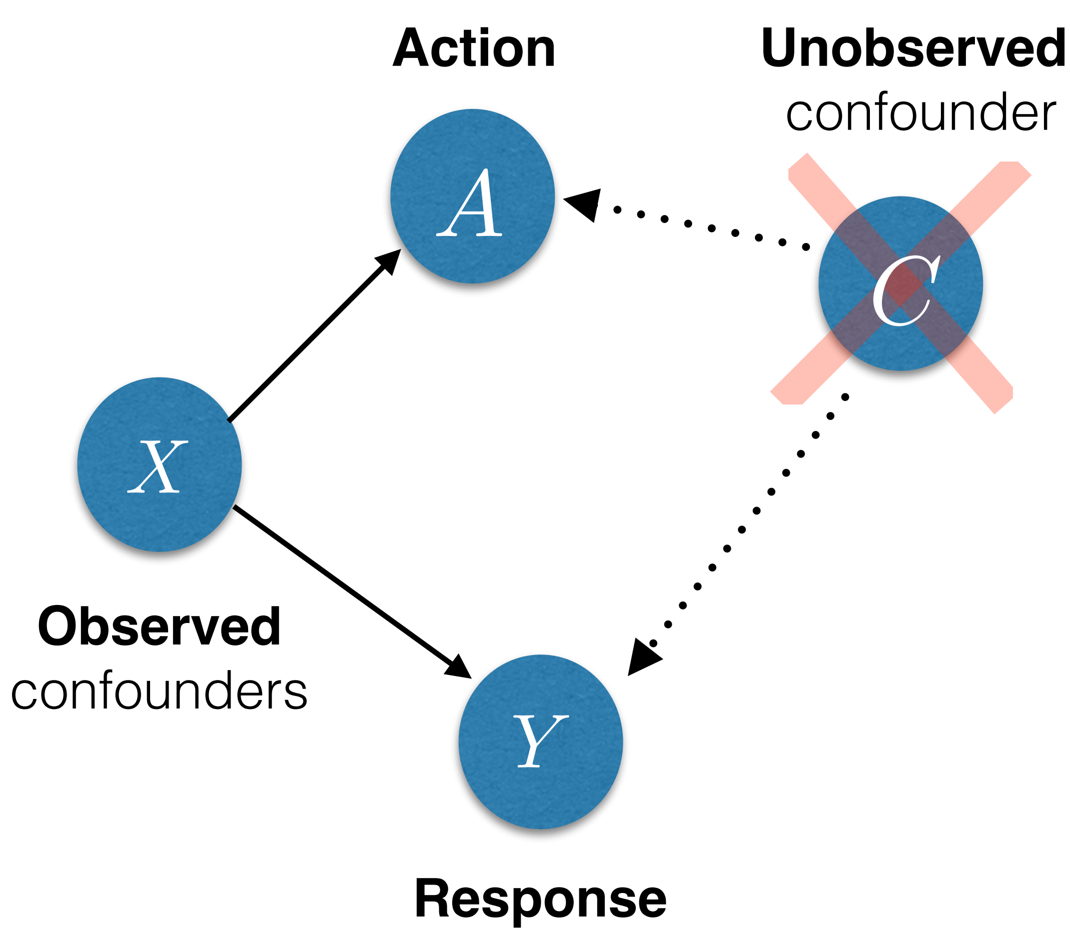
Figure: Distribution of customers colored by action.



- Estimates of treatment effects are only reliable within the overlap region.
- Outside the overlap region, causal effect estimates involves risky extrapolation.

Key Data Conditions: **Unconfoundedness**

Unconfoundedness requires that historic actions were entirely based on the observed attributes **X**.



- Confounders are variables associated with both the treatment and the outcome.
- Unobserved confounders will bias treatment effect estimates (omitted variable bias).
- Unconfoundedness is untestable and irreversible by statistical methods.

Propensity Score

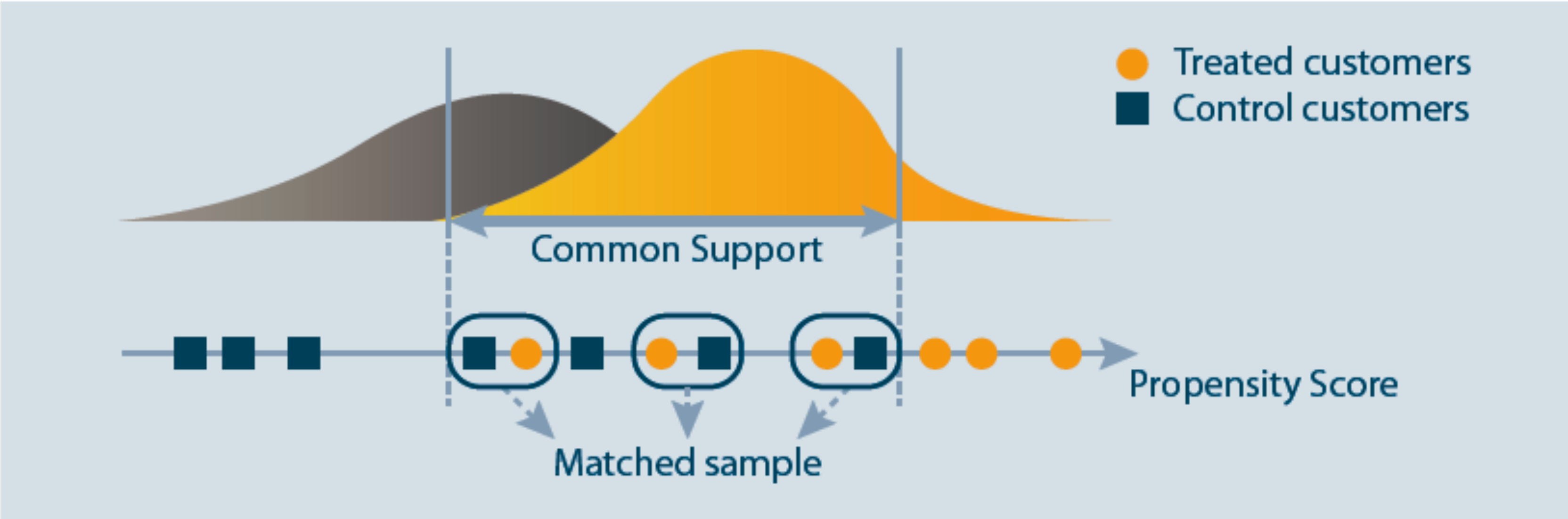
- In randomized experiments, subjects are assign to actions using some sort of **random mechanism**.
- In the absence of randomization (and assuming the key data conditions hold) subjects are assigned to actions on the basis of their attributes **X**.
- To approximate a randomized design from observational data **we need to understand the assignment mechanism**.
- This is answered by the **Propensity Score,** which is defined as the conditional probability of assignment to treatment given the attributes:

$$\pi(X) = \text{Prob}(A = 1|X).$$

Matching on the Propensity Score

- **Goal of matching**: Achieve balance on attributes between subjects exposed to different treatments.

- **The key idea:** Pair subjects that differ in the treatment they received, but have **approximately the same probability** of being assigned to the same treatment — i.e., the same propensity score.



Propensity Score — Balancing Property

What allows us to pair (match) subjects based ONLY on the propensity score?

Balancing Property

If we match subjects on the propensity score, the **distribution of attributes X** will be similar between treatment and control groups in the matched sample.

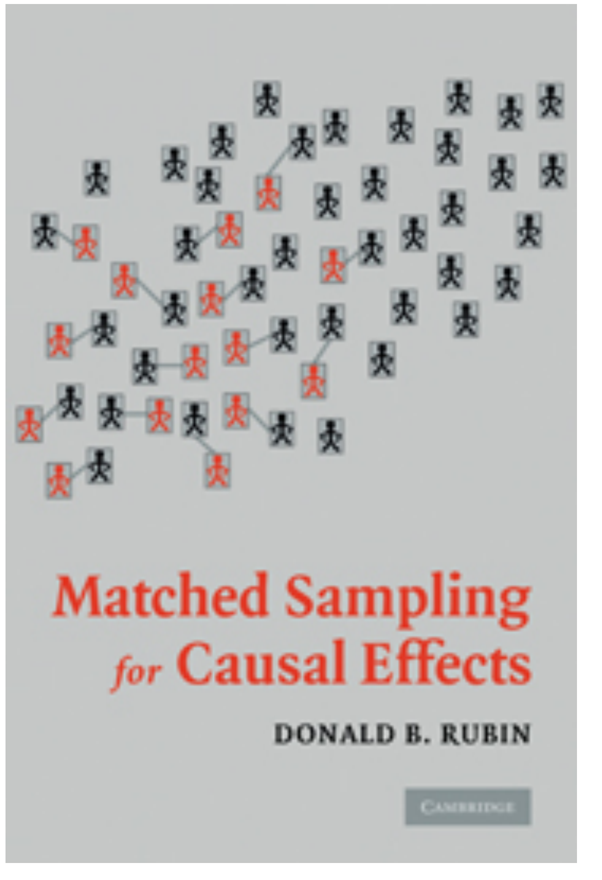
More on Matching…

Matching algorithms have many variants. There are 3 key choices:

1. The **definition of distance** between two subjects in terms of their attributes.

2. The choice of the **algorithm** used to form the matched pairs and make the distance "small" (**greedy vs. optimal matching**).

3. The **structure of the match** — i.e., the number of treated and control subjects that should be included in each match set.



Example

Propensity Score Matching for P&C Economic Price Optimization.

The Problem

1. Estimate the probability of policy renewal for each client under alternative rate changes (price-elasticity).

2. Used the estimates derived in 1 to determine the "optimal" rate change for each client.

Optimal *is defined as the set of rate changes (one for each client) that maximizes the company's profitability function subject to a desired overall (portfolio-level) retention rate.*

Data Challenges

- No access to experimental data from which to estimate the impact of alternative rating actions.
- Clients were historically exposed to rating actions based on:
    - A pricing modeling exercise.
    - Regulatory constraints.
    - Competitive analysis.
    - General business objectives.

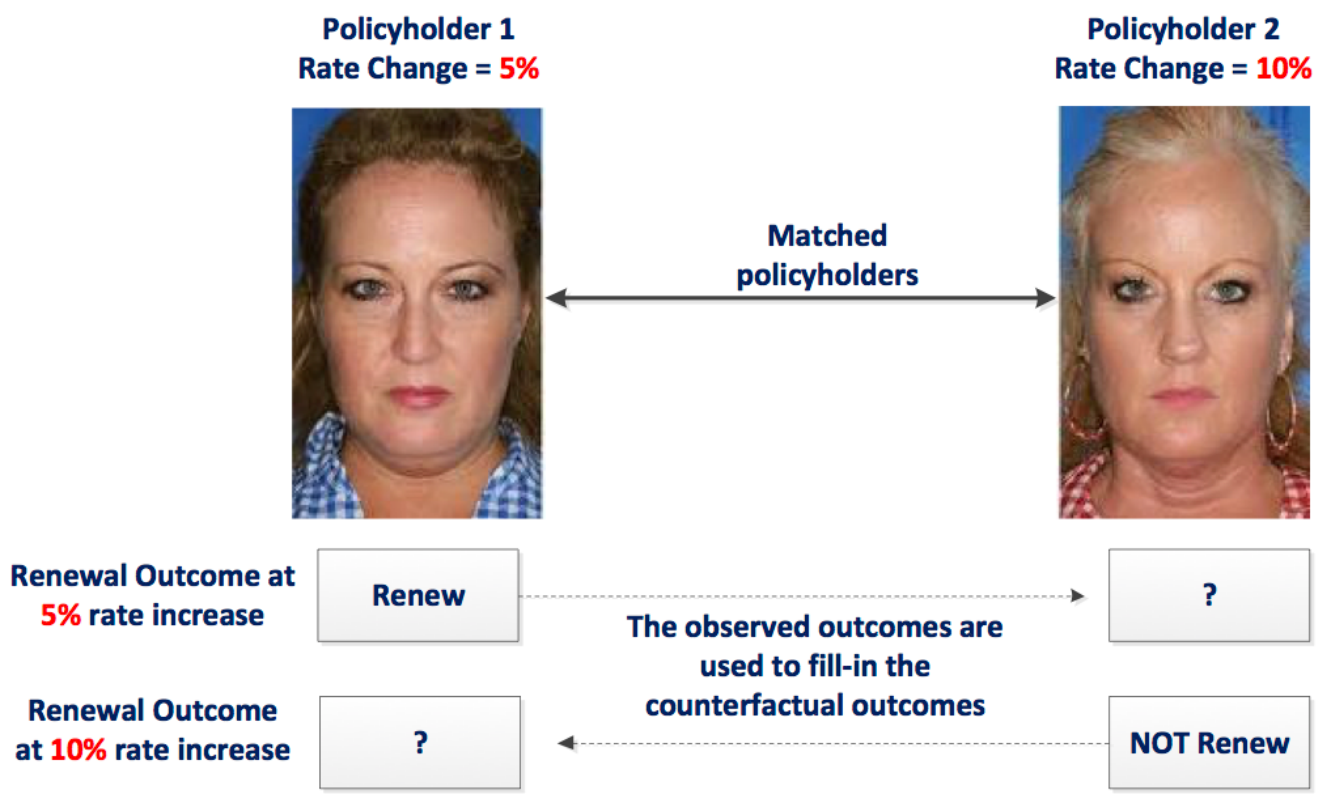
The starting point: Client-by-rate change table

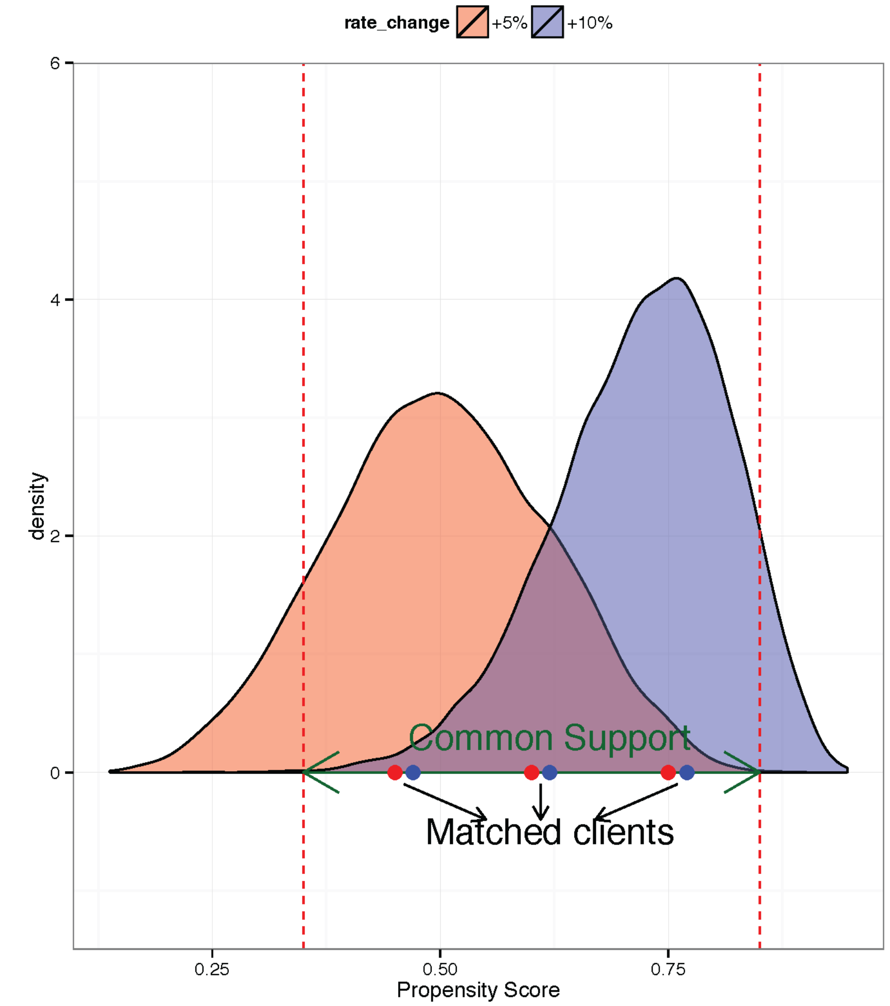
- The entries $r_{\ell a}$ below denote the observed renewal outcome $\in \{0,1\}$ of policyholder $\ell = \{1, \ldots, L\}$ when exposed to rate change level $A = a$.
- To simplify, the rate change is binned into five ordered values $A = \{1 < \ldots < 5\}$.
- Dots indicate counterfactual renewal outcomes, which are unobserved.
- The price elasticity estimation problem is equivalent to the problem of filling in the missing values in the client-by-rate change table with reliable estimates.

Table 1: **Client-by-Rate change table**

	Rate Change Level				
Client	Level 1	Level 2	Level 3	Level 4	Level 5
1	.	r_{12}	.	.	.
2	.	.	r_{23}	.	.
3	r_{31}	.	.	.	.
4	.	.	.	r_{44}	.
5	.	r_{52}	.	.	.
6	.	.	.	.	r_{65}
...	...	...	...	...	...
L	.	.	.	.	r_{L5}

Propensity Score Matching





- “**Clone policyholders**”: similar in terms of the relevant lapse predictors — i.e., about the same age, driving record, live in the same neighbourhood, etc.
- But exposed to different rate change levels.

- **Propensity score**: Probability of assignment to a 10% relative to a 5% rate increase.
- Distribution of propensity score is shown for each rate change group.
- Clients are matched only in the common support (overlap) region.

Filling the client-by-rate change table

Step 1: Replace the actual renewal outcomes with probability estimates: estimate $E[r_{\ell a}|X, A]$.

Client	Rate Change Level				
	Level 1	Level 2	Level 3	Level 4	Level 5
1	.	$\hat{r}_{12}$	.	.	.
2	.	.	$\hat{r}_{23}$	.	.
3	$\hat{r}_{31}$	.	.	.	.
...	...	...	...	...	...
L	.	.	.	.	$\hat{r}_{L5}$

Step 2: Infer the counterfactual renewal outcomes from the matched pairs (as far as the overlap situation permits).

Client	Rate Change Level				
	Level 1	Level 2	Level 3	Level 4	Level 5
1	$\hat{r}_{11}$	$\hat{r}_{12}$	$\hat{r}_{13}$	$\hat{r}_{14}$	$\hat{r}_{15}$
2	$\hat{r}_{21}$	$\hat{r}_{22}$	$\hat{r}_{23}$	$\hat{r}_{24}$	.
3	$\hat{r}_{31}$	.	.	.	.
...	...	...	...	...	...
L	$\hat{r}_{L1}$	$\hat{r}_{L2}$	$\hat{r}_{L3}$	$\hat{r}_{L4}$	$\hat{r}_{L5}$

Step 3: Develop a "global model" of the response: fit the observed + counterfactual renewal estimates on X and A.

Client	Rate Change Level				
	Level 1	Level 2	Level 3	Level 4	Level 5
1	$\hat{\hat{r}}_{11}$	$\hat{\hat{r}}_{12}$	$\hat{\hat{r}}_{13}$	$\hat{\hat{r}}_{14}$	$\hat{\hat{r}}_{15}$
2	$\hat{\hat{r}}_{21}$	$\hat{\hat{r}}_{22}$	$\hat{\hat{r}}_{23}$	$\hat{\hat{r}}_{24}$	$\hat{\hat{r}}_{25}$
3	$\hat{\hat{r}}_{31}$	$\hat{\hat{r}}_{32}$	$\hat{\hat{r}}_{33}$	$\hat{\hat{r}}_{34}$	$\hat{\hat{r}}_{35}$
...	...	...	...	...	...
L	$\hat{\hat{r}}_{L1}$	$\hat{\hat{r}}_{L2}$	$\hat{\hat{r}}_{L3}$	$\hat{\hat{r}}_{L4}$	$\hat{\hat{r}}_{L5}$

Price Optimization

Key Ideas

- Given the estimated renewal probability values, we can efficiently solve the price optimization problem.
- Find the rate change level for each client that maximizes the firm's profit function subject to portfolio-level retention constraints.
- The profit function considers: current premium, expected losses, and predicted renewal probabilities at the client-level.

The Integer Program

Maximize Profit

$$\underset{Z_{\ell a} \forall \ell \forall a}{\text{Max}} \sum_{\forall \ell} \sum_{\forall a} Z_{\ell a} \left[P_\ell (1 + RC_a)(1 - \hat{LR}_{\ell a})(1 - \hat{\hat{r}}_{\ell a}) \right]$$

Subject to retention constrains

$$\sum_{\forall a} Z_{\ell a} = 1 \quad \forall \ell$$

$$Z_{\ell a} \in \{0, 1\}$$

$$\sum_{\forall \ell} \sum_{a} Z_{\ell a} \hat{\hat{r}}_{\ell a} / L \leq \alpha .$$

Check-out our paper for more details…

Expert Systems with Applications 41 (2014) 387–396

ELSEVIER

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Expert Systems with Applications
An International Journal

A causal inference approach to measure price elasticity in Automobile Insurance

CrossMark

Leo Guelman [a,*], Montserrat Guillén [b]

[a] *Royal Bank of Canada, RBC Insurance, 6880 Financial Drive, Mississauga, Ontario L5N 7Y5, Canada*
[b] *Dept. Econometrics, Riskcenter, University of Barcelona, Diagonal 690, E-08034 Barcelona, Spain*

Key Takeaways

- Causal modeling is a more appropriate framework than predictive modeling when the objective is to assess the impact of alternative actions.
- When possible, planning a randomized experimental controlled design is the best approach to draw conclusions from the effect of alternative actions.
- As most insurance databases are derived from "business as usual conditions", inferences about treatment effects require special modeling considerations.
- Under certain data conditions, propensity score matching can be used to remove the bias in the estimation of treatment effects from observational data.
- At the very least, a visual display showing the common support region between subjects exposed to alternative actions can be useful to illustrate the extent of overlap.