Estimating Causal Effects of Credit Decisions Using Propensity Score Methodologies

Dr. Gerald Fahner, FICO
Aug. 26, 2009

FICO

Agenda

» Objective and Background

» Case Studies:

  » Risk Based Pricing

  » Credit Line Increase

» Discussion

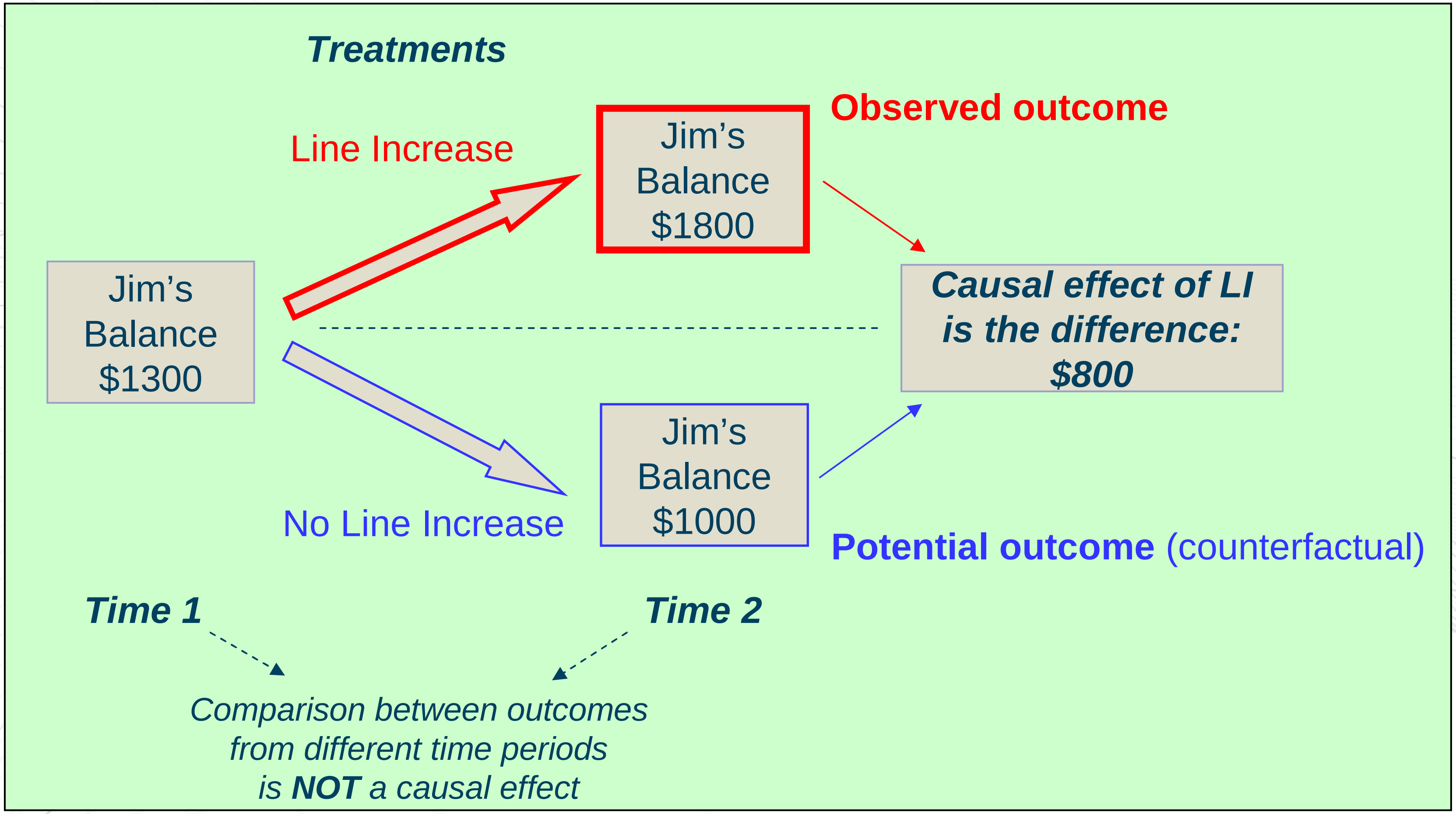
Objective

» Lenders and marketers want to determine the best credit decisions, offers, customer treatments

For this, they would like to understand the *causal effects* of their actions on future outcomes

Causal effect: "what will happen if…?"

- » Rubin's Causal Model
  - » Causal effects are a comparison of outcomes from two or more alternative treatments, with only one of the outcomes observed



Responses to causal effect estimation problem

- » Gold standard
  - » Randomized experiments generate maximally informative data to estimate causal effects of interest

Example*

	Nonsmokers	Cigarette Smokers	Cigar and Pipe Smokers
Mortality Rates per 1000 person-years	13.5	13.5	17.4

- » Treatments (self-selected, not randomly assigned)
  - » Nonsmoking
  - » Cigarette smoking
  - » Cigar and pipe smoking
- » Naïve conclusion
  - » Cigarette smoking is good for your health, especially relative to cigar and pipe smoking

    …may be misleading

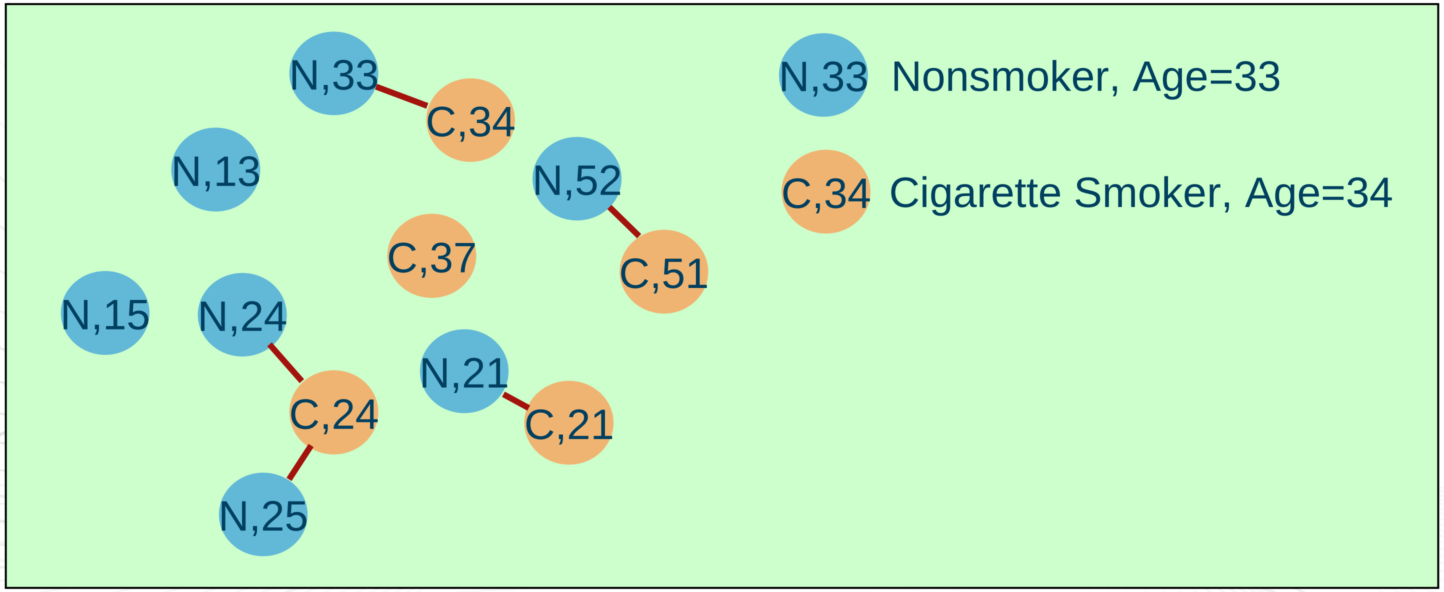
***Adapted from Rubin and Cochran**

Age is a confounding variable

	Nonsmokers	Cigarette Smokers	Cigar and Pipe Smokers
Mortality Rates per 1000 person-years	13.5	13.5	17.4
Average Age	57.0	53.2	59.7

- » Age is correlated with mortality and also with smoking habits
  - » Conclusions about effects of smoking should be adjusted for age
- » Adjustment approaches include
  - » Regression analysis
  - » Subclassification
  - » Matched sampling

Matched sampling on Age*



- For each nonsmoker, find a cigarette smoker of similar age if possible. Analogous for cigar and pipe smokers
- Age will be balanced between the matched samples
- Compare mortality rates based on matched samples only
  - Some samples may be discarded for estimation

* Many variants of matching, see e.g. D. Rubin's book, "*Matched Sampling for Causal Effects*"

Agenda

» Problem and Background

» Case Studies:

  » Risk Based Pricing

  » Credit Line Increase

» Discussion

Effect of risk-based pricing on Take Rate

Risk Premium	Base Rate	Base + 3%	Base + 5%
Take Rate (TR)	68%	37%	44%

- Naïve conclusions:
  - Changing from (Base+3%) to (Base+5%) increases TR by 7% …doesn't make sense, assuming consumers behave rational
  - Changing from Base Rate to (Base+3%) decreases TR by 31% …makes more sense, but estimate may nevertheless be biased

APR offer was determined in two steps:

1. Decide on Base Rate (not discussed here)
2. Decide on Risk Premium to add to Base Rate

Should adjust for many covariates

Risk Premium		Base Rate	Base + 3%	Base + 5%
Take Rate (TR)		68%	37%	44%
Application Score (mean)		172	140	134
Applicant Age (mean)		47	42	37
Channel:	Supermarket	30%	36%	52%
	Internet	16%	20%	1%
	Marketing	54%	44%	47%
Residence:	Owns	86%	68%	51%
	Rents	11%	29%	45%
	Other	3%	3%	4%
.. many other covariates		..	..	..

Data spanning across multiple marketing channel

Different marketing channels used different risk-based pricing strategies

Propensity score-based matching for causal effect estimation

FICO

Outline of process steps

I. Develop *propensity score(s)* between alternative treatments
   - » Will simplify matching when many covariates need adjustment

II. Generate *matched samples* based on propensity scores
   - » Matched data will resemble data from randomized experiments
   - » Useful for descriptive analysis and as basis for regression estimation

III. Estimate *causal effects* from matched samples
   - » Will use regression estimation techniques

Benefits

- » Reduced risk of confounding causal effect estimates
- » Clearer appreciation of what information is in the data and what isn't
- » More objective assessment of causal relationships

Step I: Develop propensity score(s)

» Propensity score models probability of individuals receiving treatment (as opposed to control), given covariates

$$\Pr\{T = 1 \mid X\} = p(X)$$

where:

$T : \{0,1\}$ treatment indicator (0 : control, 1 : treatment)

X : Covariates

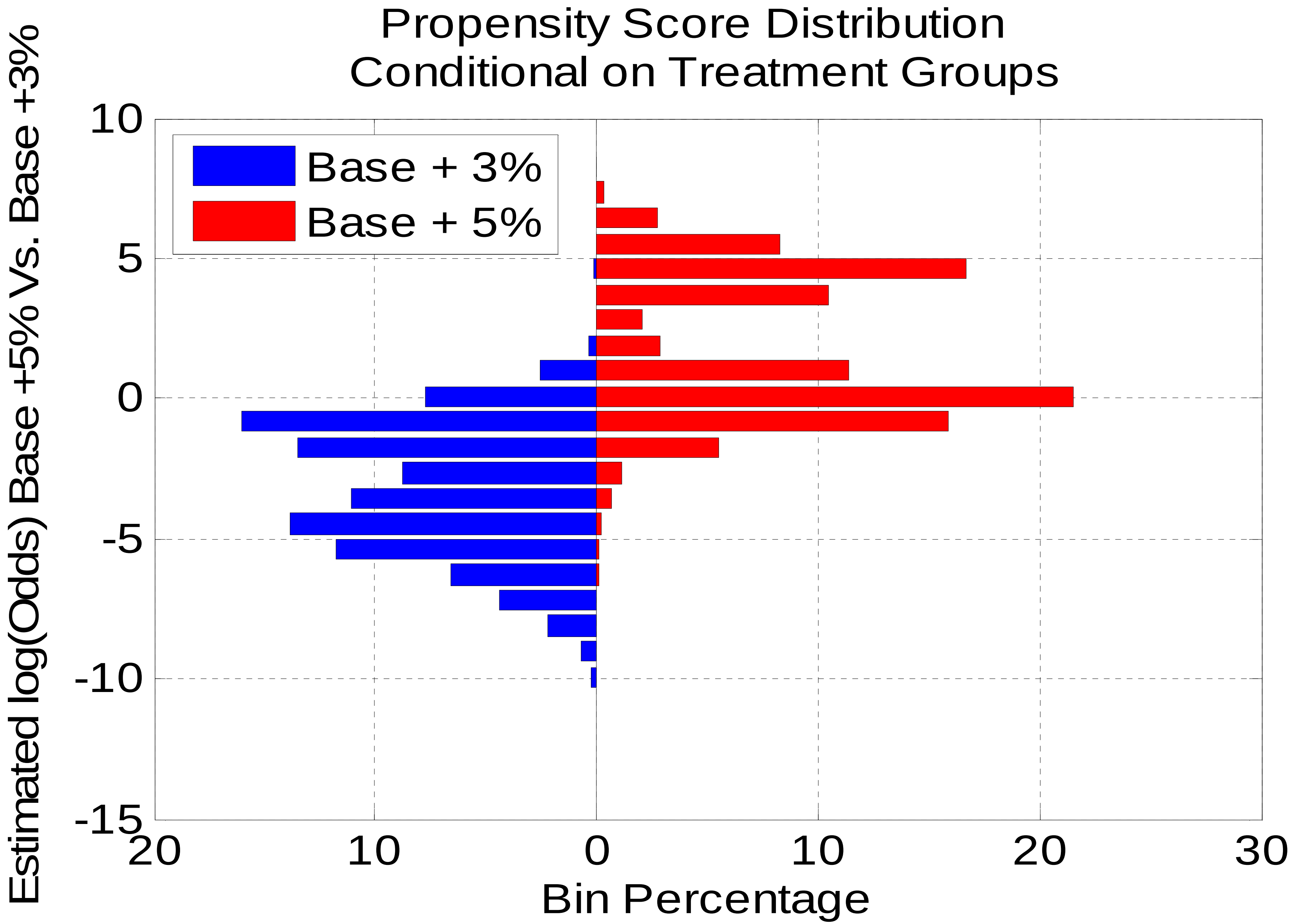
» Score is developed from observed data

» May estimate using logistic regression, scorecard or other generalized additive models, random forests, neural nets, etc.

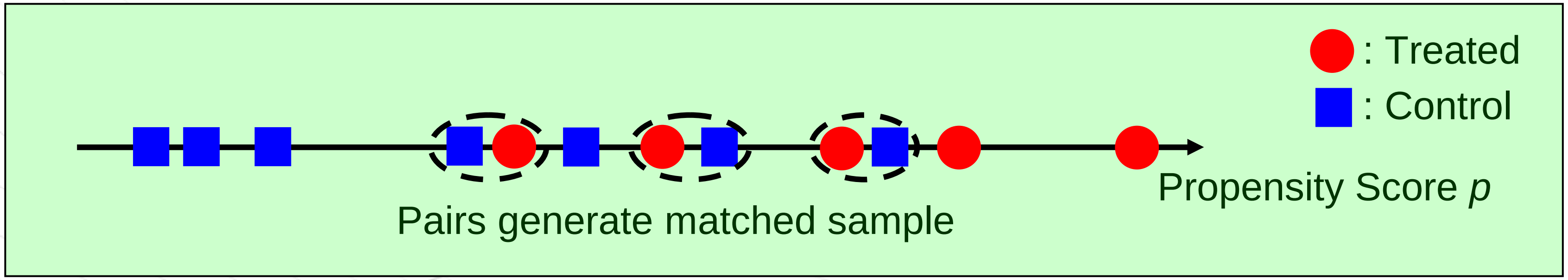
» Develop multiple propensity scores for various treatment dichotomies, such as:

» Control = Base Rate	Control = Base + 3%	
» Treatment = Base + 3%	Treatment = Base + 5%	etc.

Limited overlap between Base+3% and Base+5%



Step II: Matching on propensity score



- » Propensity score simplifies matching, by reducing all covariates to be matched, to a single covariate (*p*) to be matched
- » If the propensity score is reasonably developed, then, within any stratum of *p* that contains treated and controls, treated individuals are similar to controls in terms of their covariate distributions
- » Practical assessment of estimated propensity score
  - » Do the matching
  - » Covariates should be well balanced on the matched sample
  - » Diagnosis of balance is important. Seek to improve propensity score until satisfied with balance, *before* proceeding with estimation step

Risk-based pricing example before and after matching

FICO

	Observed sample N=11,837 Unbalanced covariates		*Matched sample* N=3,404 Balanced covariates	
APR Premium (treatment)	**Base + 3%**	**Base + 5%**	**Base + 3%**	**Base + 5%**
Application Score	140	134	137	137
Applicant Age	42	37	40	39
Channel: Supermarket	36%	52%	50%	51%
Internet	20%	1%	1%	1%
Marketing	44%	47%	49%	48%
.. many other covariates	..	..	..	..

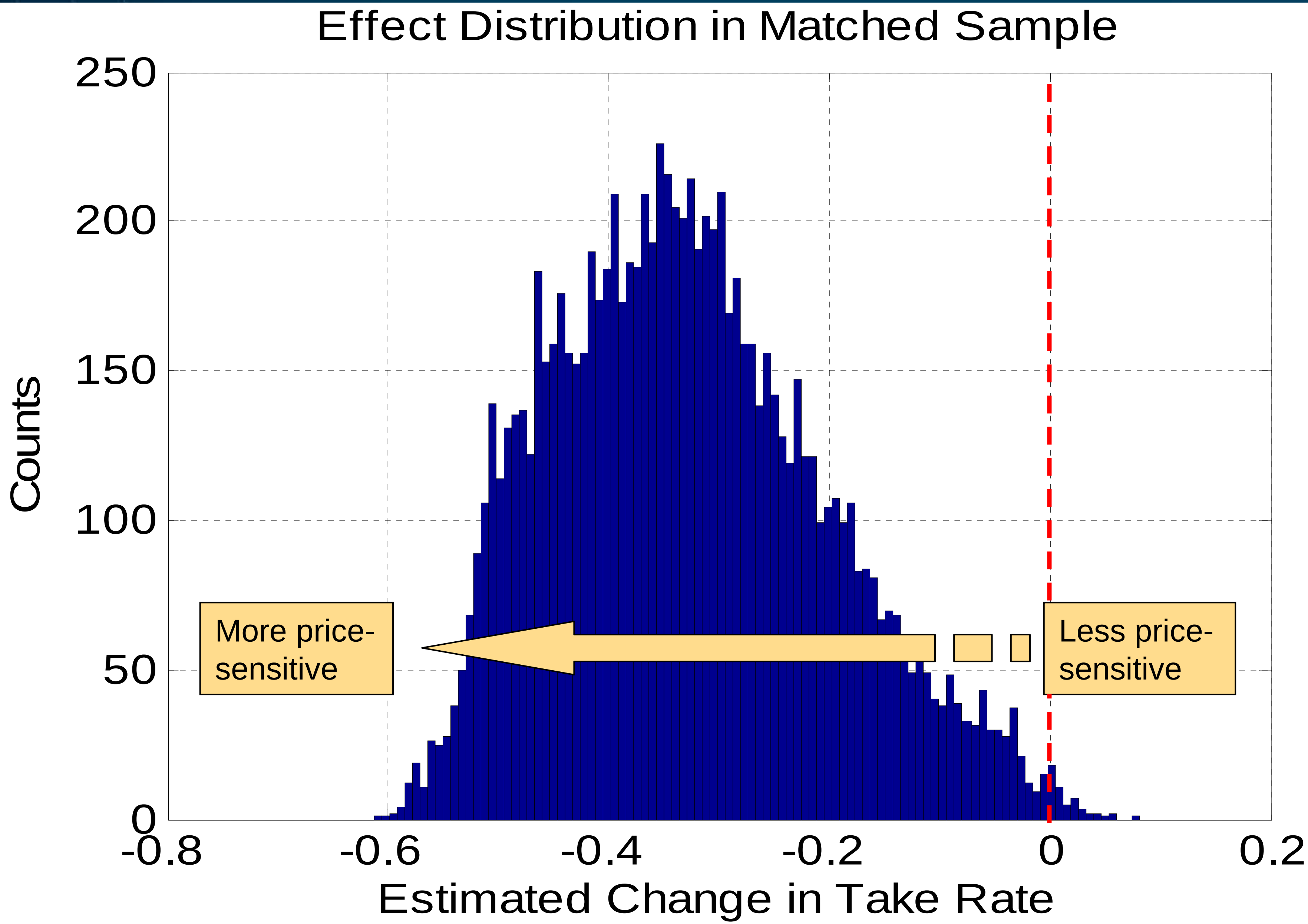
» Matched sample balances all covariates going into propensity score

» (Almost) as if data came from a randomized experiment

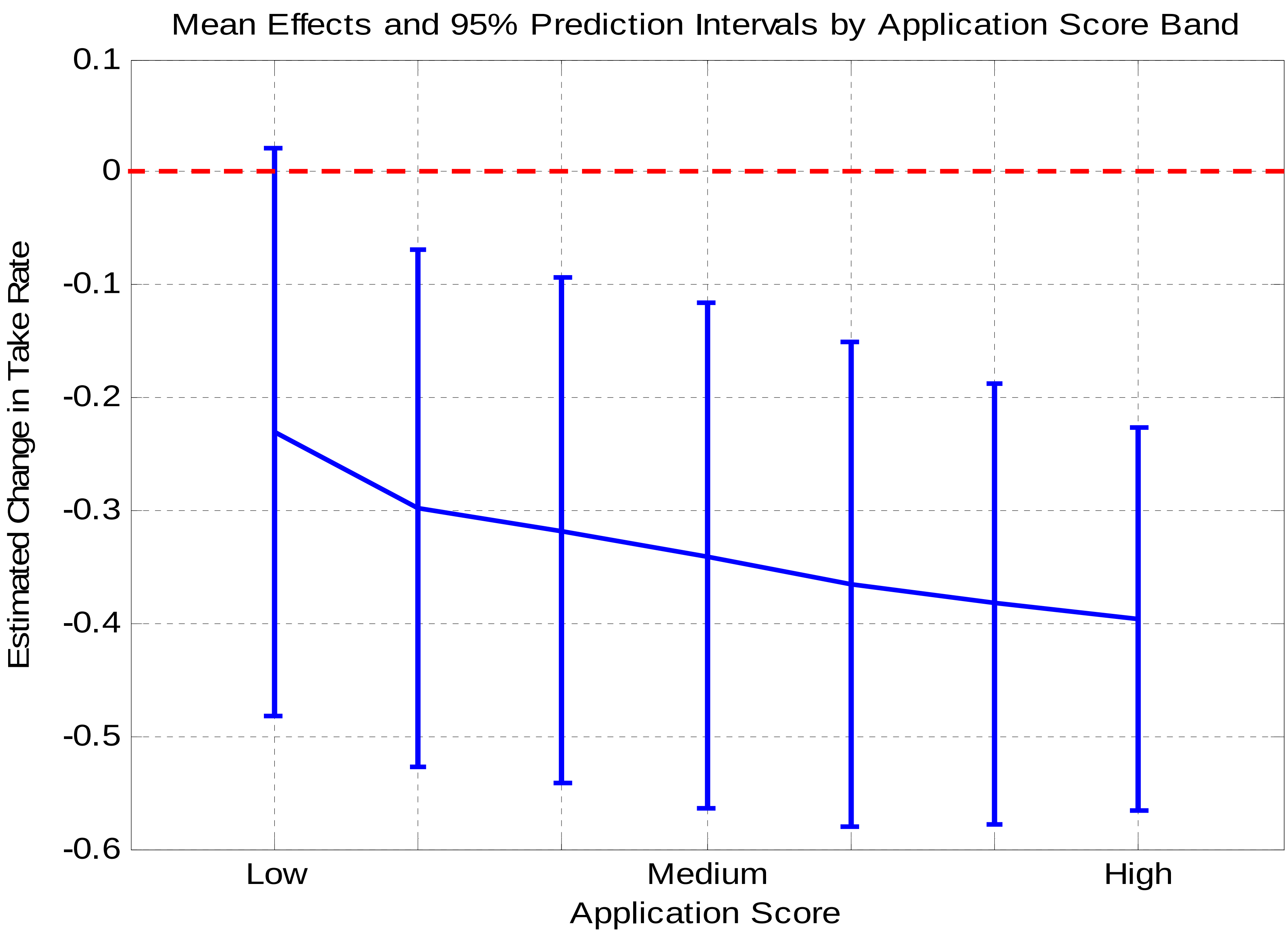
Step III: Regression analysis on matched sample

- Include treatment variable, along with covariates, in regression model (GLM, GAM, Scorecard, NN, …) to predict outcomes
  - Specify model depending on nature of data and analysis goals
- Models with additive treatment effect:
  - May improve estimate of average treatment effect
  - By further adjusting for covariate distributions
- Testing for treatment x covariate interactions:
  - Do treatment effects differ between individuals?
  - Infer segment-level or individual-level treatment effects

Effects of changing treatment from Base → Base + 3% are heterogeneously distributed



Price sensitivity (Base → Base + 3%)
By Risk band



Agenda

» Problem and Background

» Case Studies:

  » Risk Based Pricing

  » Credit Line Increase

» Discussion

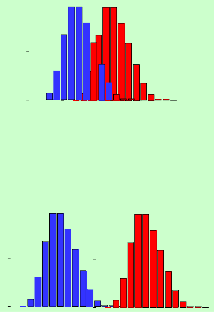
Line increases in a credit card portfolio

FICO

- Mostly deterministic line increase strategy (very limited testing)
  - Increase amounts in multiples of $500
- Analyzed all treatment-control pairs and generated matched samples, where possible
  - In tendency, overlap decreased with distance between treatments (with some exceptions). Example:

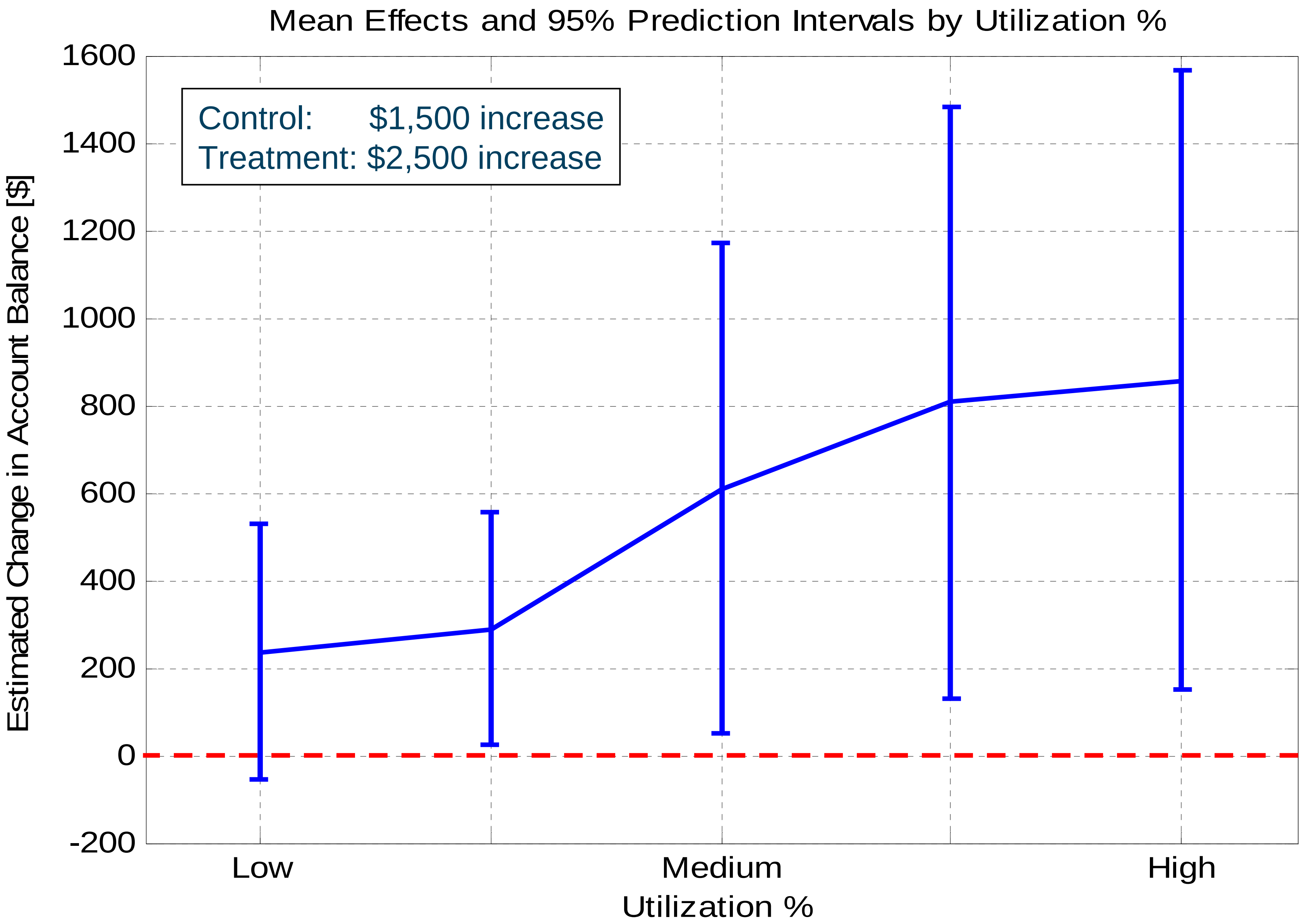
Control [$]	Treatment [$]	Overlap
1,000	1,500	23%
1,000	2,000	26%
1,000	2,500	13%
1,000	3,000	5%
1,000	3,500	0%

P'score distributions

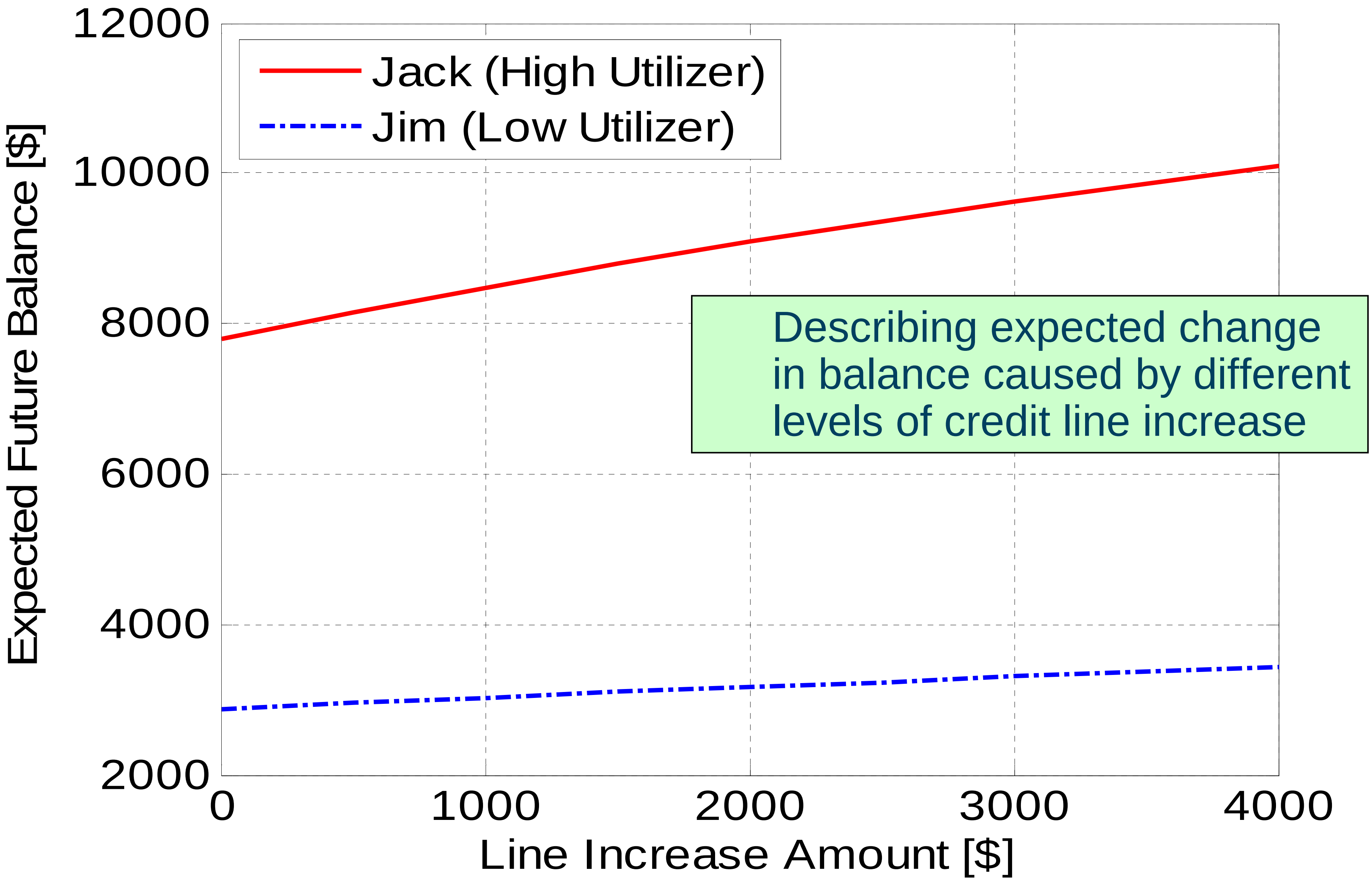


- Some comparisons struggled with too few matched samples. More aggressive testing would have been a boon for analysis
- But decent overlap across various treatment-control pairs, even with limited testing, is encouraging news for causal modeling

Effect of line increase on account balance
By Utilization band



Credit line increase-response curves



Discussion

- Selection bias in business-as-usual data poses analytic challenges for estimating causal effects
- Direct regression analysis of BAU data can yield unreliable results
- Benefit of matching prior to regression analysis
- No miracle cure exists if important confounders are not collected

- Mainstays of analytic success continue to include:
  - Purposeful design of experiments
  - Exhaustive collection of covariates associated with treatments or outcomes
  - Sniffer dog capabilities to sense lurking variables
  - Incorporation of domain expertise into the modeling process

Selected references

Landmark paper on properties and virtues of the propensity score:

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

Assembly of many papers from introductory to advanced, and good advice:

"*Matched Sampling for Causal Effects*", by Donald B. Rubin, Cambridge University Press, 2006.

For a leisurely introduction into propensity score ideas:

"Estimating Causal Effects from Large Data Sets Using Propensity Scores", by Donald B. Rubin, *Annals of Internal Medicine*, Volume 127(8S), October 1997, pp 757-763.

THANK YOU

Dr. Gerald Fahner, FICO

Relative distribution analysis over time

How treatment affects account balance distribution

FICO

Reference distribution (matched controls): No increase
Comparison distribution (matched treated): $1,000 increase

For each quantile of the reference distribution, *relative density* indicates enrichment (if >1), or depletion (if <1) of comparison distribution in this quantile.