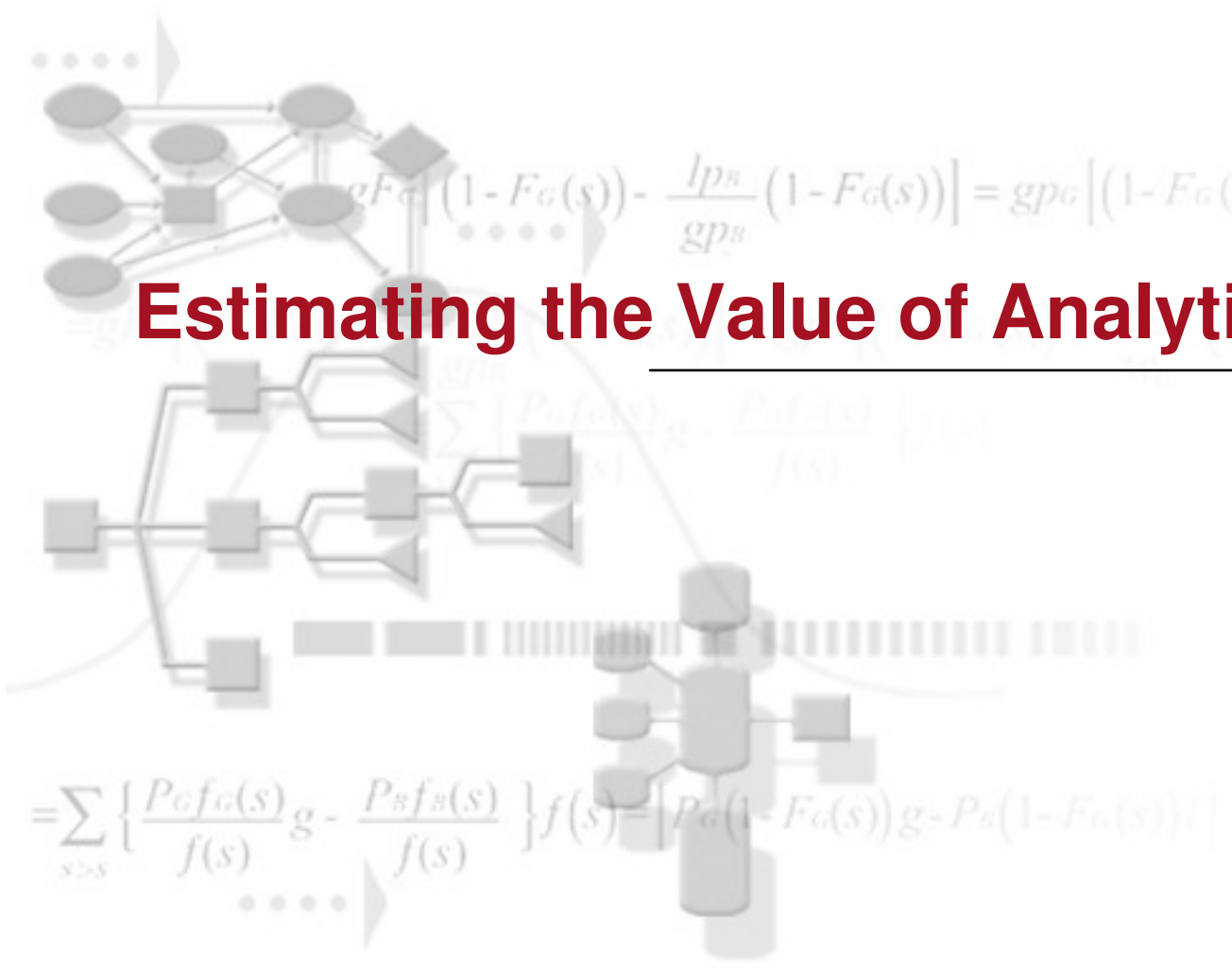


# Estimating the Value of Analytic Innovations

Dr. Gerald Fahner

Fair Isaac Corporation



## Agenda

- ▶ Analytic Value Chain Simulation
- ▶ Applications
  - ▶ Transaction Scoring
  - ▶ Reject Inference

# Problem

---



- ▶ Estimate commercial potential of analytic innovations (new data sources, score development technology), *before* deploying them in the field

- ▶ Developed scientific framework of

### **ANALYTIC VALUE CHAIN SIMULATION**

*(Patent Pending\*)*

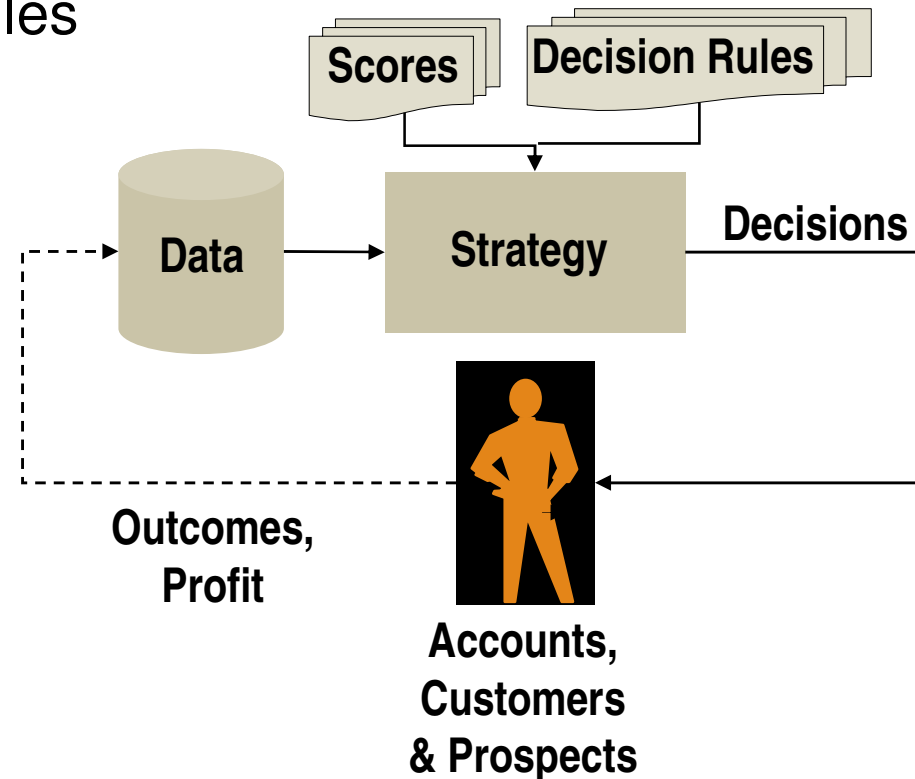
to tackle this question for various innovations

*(\* Pending Patent covers broader simulation methodologies and applications than those presented here*

# Profit From Data-Driven Operation



- ▶ Depends on:
  - ▶ Information in the data
  - ▶ Quality of scores/models
  - ▶ Choice of decision rules
  - ▶ Consumer reactions to decisions



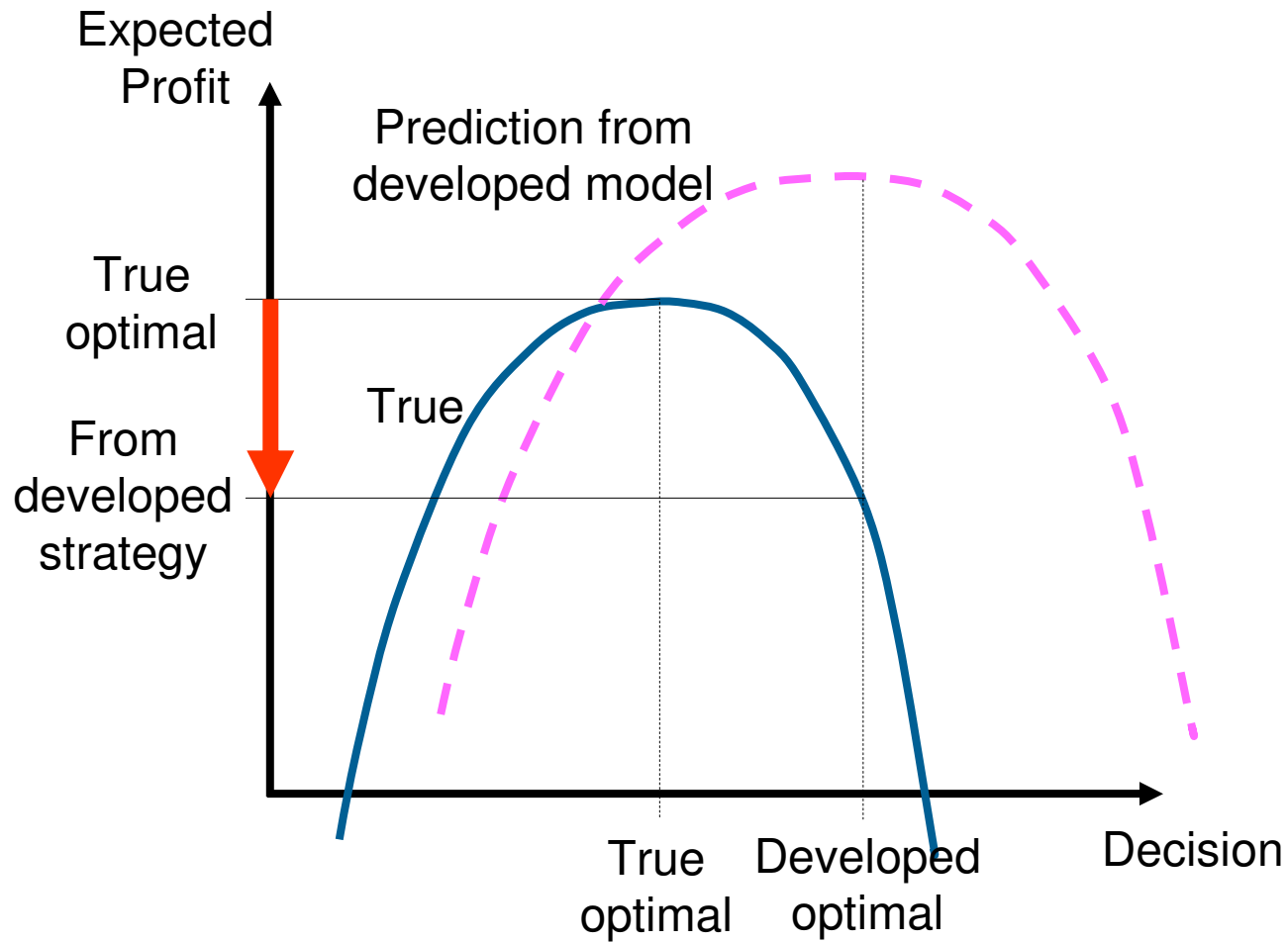
# Model-Based Optimization Framework

---



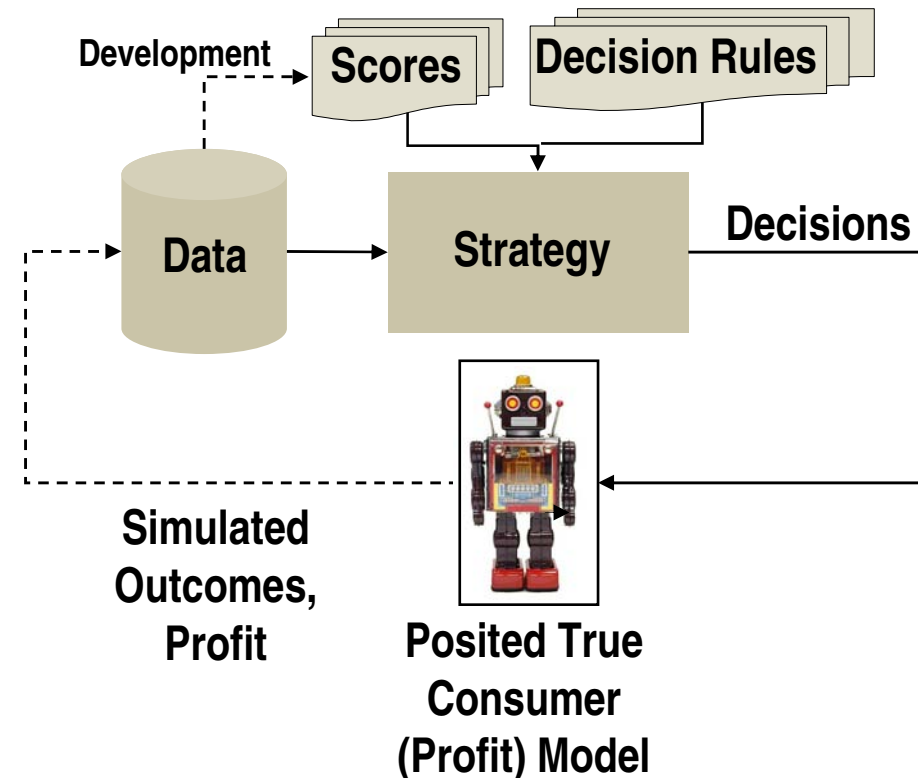
- ▶ Profit (or any business utility function) is modeled as a function of consumer data/scores and decisions
- ▶ Decision rules (=strategy) are determined through function optimization
- ▶ This creates a value chain:
  - ▶ *More Informative Data → More Accurate Models → More Optimal Strategies → Higher Profit*

# Inaccurate Models Reduce Profit



# Posit a “True Model” for Profit

- ▶ Used as a benchmark in simulations to gauge performance of developed scores and strategies
- ▶ Guided by best data sources, deep data analysis, and domain expertise
- ▶ Uncertainties about true model typically exist and can be taken into account



# Analytic Value Chain Simulation



- ▶ Calculates opportunity of an analytic innovation

$$\text{Opportunity} = \text{Profit with Innovation} - \text{Profit without Innovation}$$

- ▶ Elements of simulation:
  - ▶ Obtain data for model development
  - ▶ Develop profit models, (i) with and (ii) without innovation
  - ▶ Optimize two strategies based on the two profit models
  - ▶ Calculate profits from the strategies, using posited model

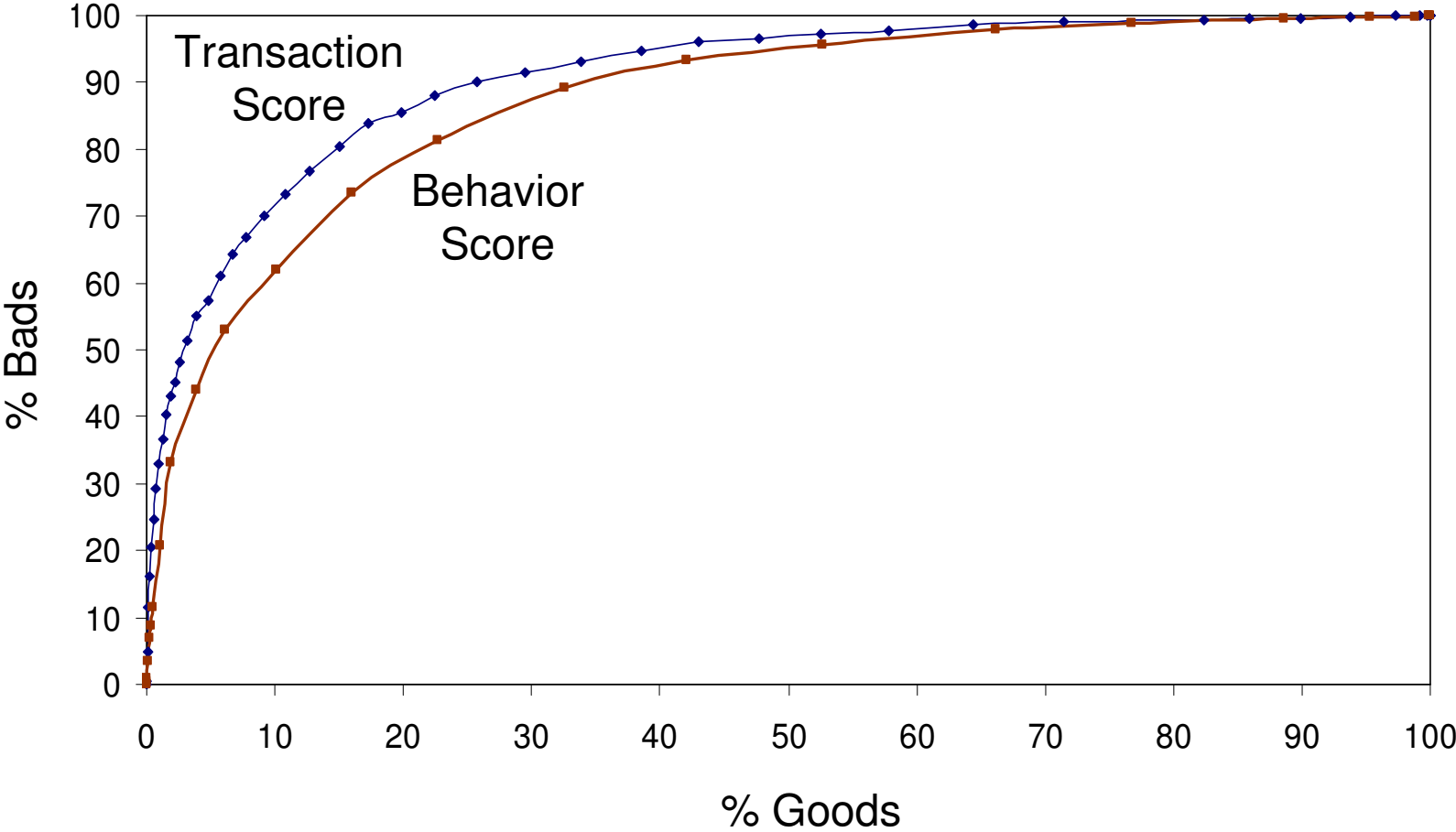
## Agenda

- ▶ Analytic Value Chain Simulation
- ▶ Applications
  - ▶ Transaction Scoring
  - ▶ Reject Inference

# Transaction Score Innovation Improves Risk Prediction (Credit Card Portfolio)



### ROC Curve Comparison



# Opportunity of Transaction Risk Score for Making Credit Line Increase Decisions



- ▶ We posited a profit model that depends on line increase amount and account information; in particular, we posited a dependency of future  $pD$  on transaction data (strong evidence exists)
- ▶ Result of Analytic Value Chain Simulation:

Mean Expected Profit per account		Mean Opportunity (of using the Trx. Score) per account
Developed model uses no Trx. Score	Developed model uses Trx. Score	
\$44*	\$50*	\$6*

- ▶ There are however uncertainties about the posited profit model
  - ▶ Uncertain future  $pD$  to Risk Score calibration, also uncertain calibration parameters for revenue-, loss-, and attrition

\*Example figures – not calibrated to a particular portfolio. Assumed profit measured over one year, one credit line decision

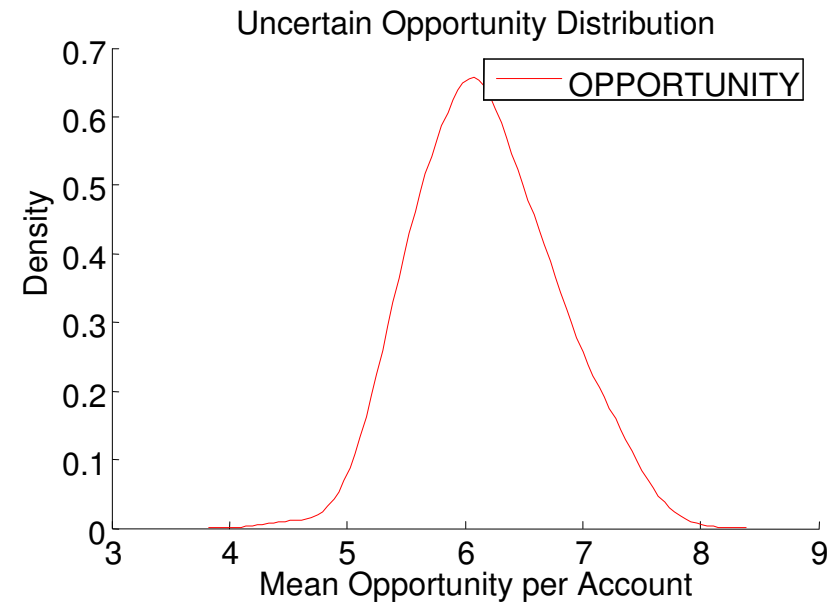
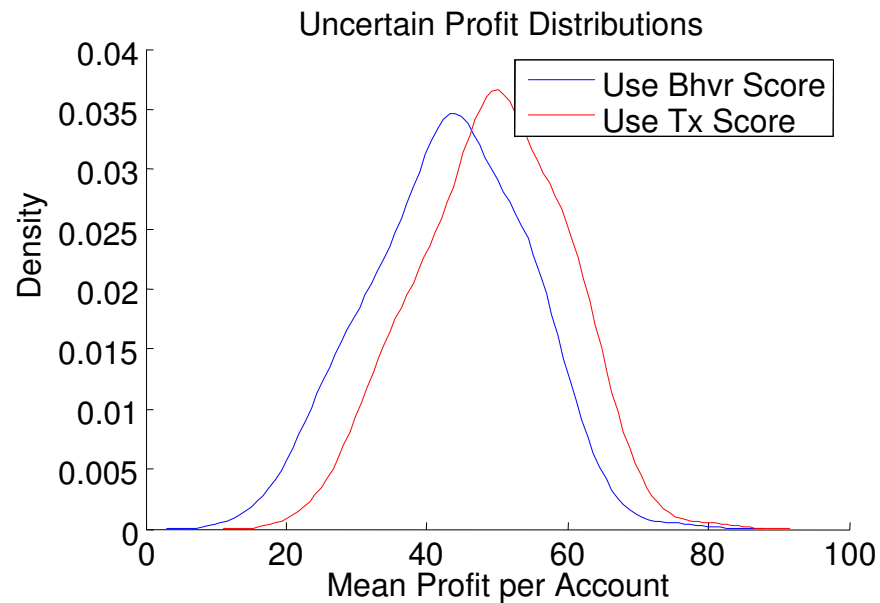
## Dealing with Uncertainty

---



- ▶ Positing a model to be “true” is a strong assertion
- ▶ Assumptions should be challenged
- ▶ Stress testing and/or Monte-Carlo risk analysis:
  - ▶ Repeat Analytic Value Chain Simulation multiple times, while positing a variety of likely true models

# Uncertain Profit- and Opportunity Profiles



- ▶ Significant uncertainty around profit - mainly driven by uncertainty about future  $pD$
- ▶ Transaction Score opportunity holds up robustly

## Agenda

- ▶ Analytic Value Chain Simulation
- ▶ Applications
  - ▶ Transaction Scoring
  - ▶ **Reject Inference**

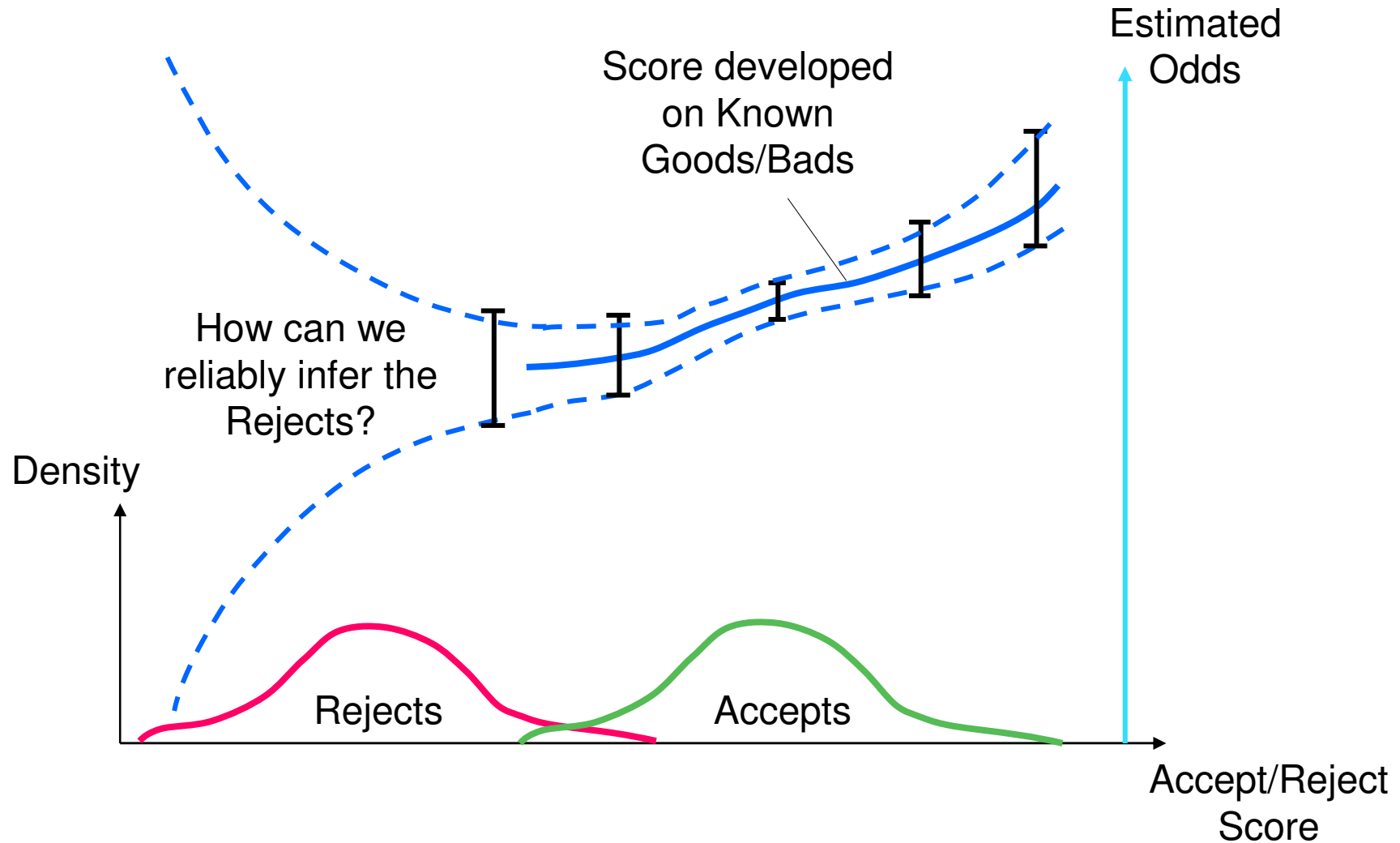
# Challenges of Reject Inference

---

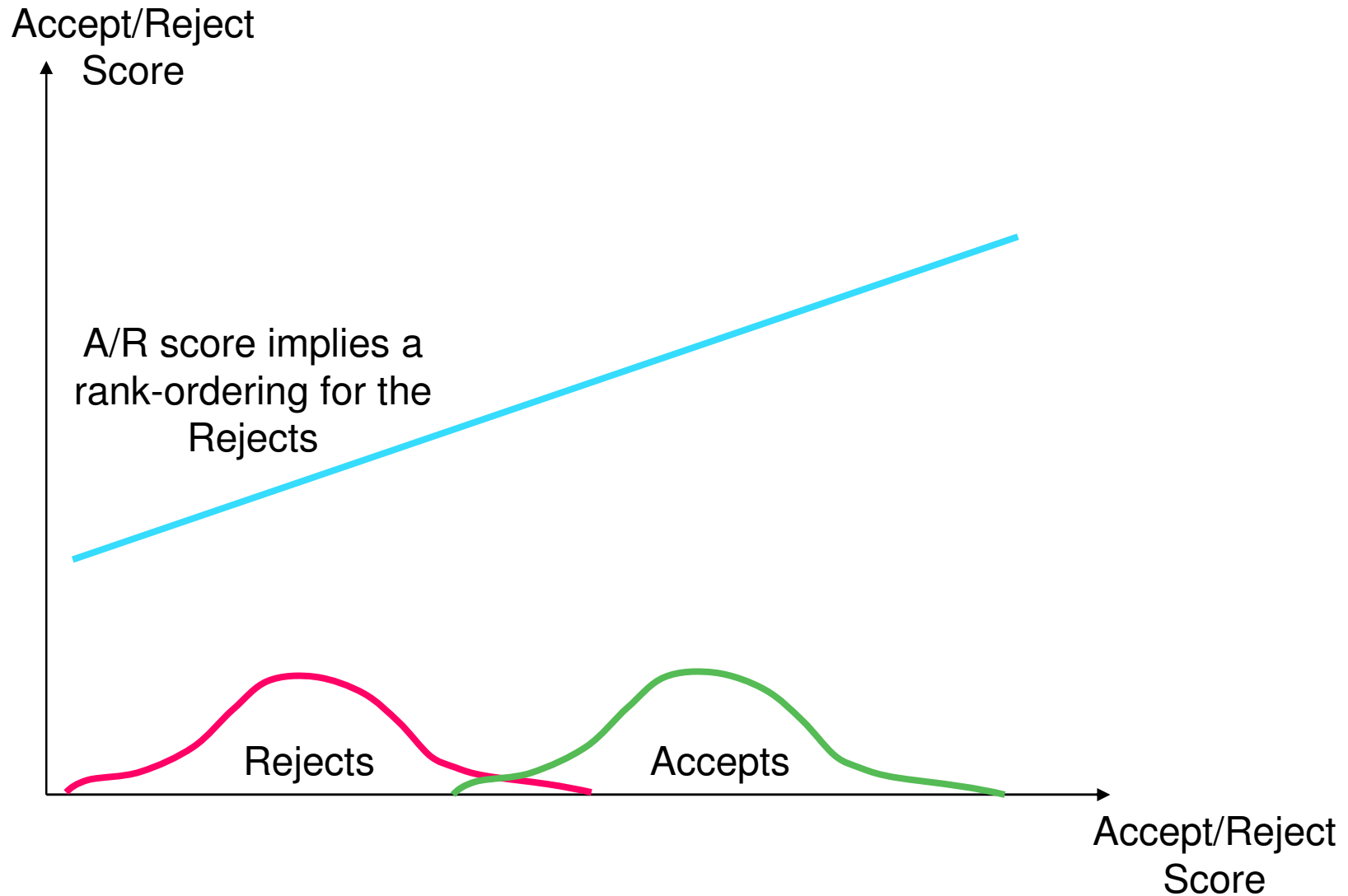


- ▶ “Cherry picking” - omitted variables can lead to biased models
- ▶ Extrapolation - Rejects may form a very different population from Accepts
- ▶ In the following, assume that all variables that mattered for the historic decision are available for risk score development, thus extrapolation is the remaining challenge

# Extrapolation Challenge



# Can Memory from the Historic Screen Help?



## “Dual Score” Reject Inference Innovation

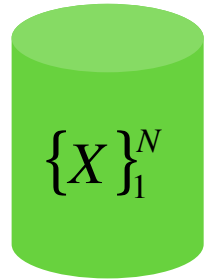


- ▶ Uses Accept/Reject (A/R) score to bias extrapolation of Known Good/Bad score into the Rejects
- ▶ Success relies on assumption that A/R score is a somewhat rational rank ordering instrument for risk within the Rejects

# Simulation Setup



Application data set



Posit logistic regression scorecard  $R$

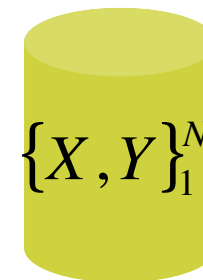
Specify historic screen scorecard  $Q$  and cutoff

$$Y = \begin{cases} \text{Bernoulli}(p_G(X)) & \text{if } Q \geq \text{historic cutoff} \\ \text{Missing} & \text{otherwise} \end{cases}$$

## Stress Test Historic Screen Scenarios

Characteristics	Their Weight Patterns		
	Posited Scorecard $R$	Mildly irrational scorecard $Q1$	Very irrational scorecard $Q2$
CB score			
Mts. since last TL opened			
Avg. mts. on file			
# Satisfactory ratings			
Applicant age			

Development data set



Score- and strategy development

## Structure of Profit Model



- ▶ Expected profit:

$$E_{profit} = \begin{cases} p_G(X; \beta)g - (1 - p_G(X; \beta))l & \text{if Accepted} \\ 0 & \text{if Rejected} \end{cases}$$

$X$  = Application Data

$p_G$  = Probability of good outcome

$l$  = Loss in case of bad loan

$g$  = Gain in case of good loan

- ▶  $l, g$ , scorecard for  $p_G$  are posited
- ▶ Developed scores estimate  $p_G$  from development data set

# Structure of Optimal Strategy



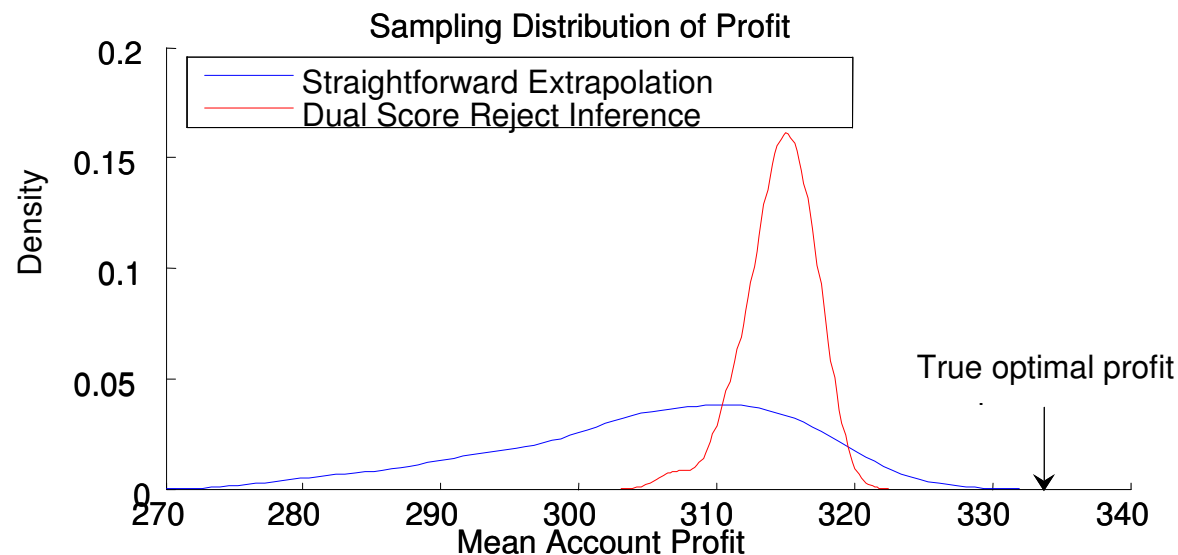
$$Decision = \begin{cases} \text{"Accept"} & \text{if } Eprofit > 0 \\ \text{"Reject"} & \text{otherwise} \end{cases}$$

- ▶ If plug in posited *Eprofit*:
  - obtain true optimal strategy and its associated true optimal profit
- ▶ If plug in developed models - obtained by:
  1. Monotonicity-constrained (otherwise straightforward) extrapolation
  2. Dual Score reject inference
  - obtain the respective developed optimal strategies and their expected profits

# Dual Score Reject Inference Can Outperform Straightforward Extrapolation



- ▶ To get a sense of variation, we repeated Analytic Value Chain Simulation 40 times, by re-sampling development data from the fixed posited model



- ▶ Two-fold benefit of Dual Score Reject Inference:
  - ▶ Increased expected profit
  - ▶ Reduced sampling variation of profit

# Under Which Conditions Does Dual Score Inference Help?



- ▶ Historic screen needs to be somewhat rational AND new acceptance rate needs to be comparable or higher than historic acceptance rate (portfolio expansion)
- ▶ Dual Score method gracefully degrades to straightforward extrapolation performance if the historic screen is less rational, or in the case of portfolio contraction

# Summary



- ▶ Analytic Value Chain:

Data → Models → Decisions → *Profit*

- ▶ Simulation determines sensitivity of profit with respect to improvements of data and modeling techniques
- ▶ Results of simulation can guide investment decisions to put improved analytics into use
- ▶ Simulation of value chain has been extended to understand the value of “*closing the loop*” (see conference proceedings)