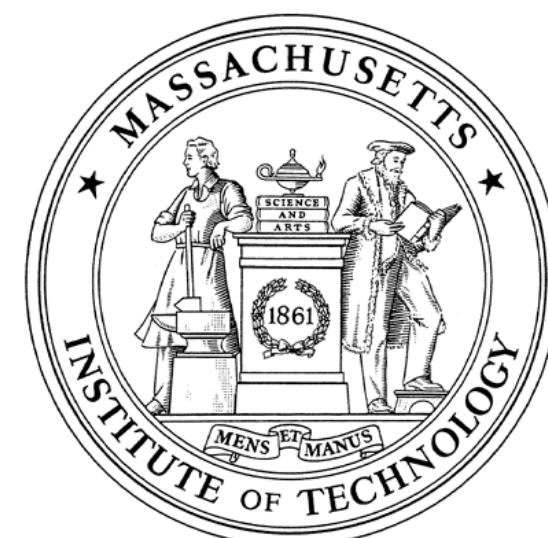# *From Data to Innovations*

David Simchi-Levi

MIT Data Science Lab: Executive Summary

Theoretically Elegant & Practically Relevant Research

Supply Chain Resiliency	Price Optimization	Personalized Offering	Inventory, Transportation & Procurement	Online Resources Allocation	Supply Chain Digitization
Ford	Rue la la	RYANAIR	BlueYonder	IBM	accenture
accenture	GROUPON	accenture	THE HOME DEPOT	Alibaba.com	ABInBev
Sigma alimentos / CAMPOFRIO FOOD GROUP	ORACLE	XL Group Insurance Reinsurance	accenture		zalando lounge
DENSO	Coppel	ERICSSON	MANGO		
Schneider Electric	zalando				
	starwood Hotels and Resorts				
	B·TECH				

Strategic intent: Develop solutions to leading edge problems for lab partners through research that brings together data, modeling, and analysis to improve business performance

Cross-industry: Oil / Gas, Retail, Financial Services, Government, Insurance, Airlines, Industrial Equipment, Software

Global footprint: NA, EU, Asia, LA

Data Driven Business Transformation

- **Supply Chain Resiliency**
  - **Ford Case Study**

- **Online Pricing for Fashion Retailers**
  - **Zalando Case Study**

- **Conclusions**

The Science of Supply Chain Resiliency

Harvard Business Review

JANUARY–FEBRUARY 2014
REPRINT

From Superstorms to Factory Fires

Managing Unpredictable Supply-Chain Disruptions

by David Simchi-Levi, William Schmidt, and Yehua Wei

Harvard Business Review

Economy

How Coronavirus Could Impact the Global Supply Chain by Mid-March

by Pierre Haren and David Simchi-Levi

February 28, 2020

Harvard Business Review — Crisis Management | We Need a Stress Test for Critical Supply Chains

Crisis Management

We Need a Stress Test for Critical Supply Chains

by David Simchi-Levi and Edith Simchi-Levi

April 28, 2020

Abstract Aerial Art/Getty Images

Harvard Business Review — Operations And Supply Chain Management | How the War in Ukraine Is Further

How the War in Ukraine Is Further Disrupting Global Supply Chains

by David Simchi-Levi and Pierre Haren

March 17, 2022

HBR Staff/Niitaya Singhaseri/Getty Images

4

Excellence in the Supply Chain Resiliency

1 **Resiliency is a philosophy (like LEAN) not a tactic (like JIT)**

In the current disruptive world resilient supply chain wins business and growth

2 **Balance Cost, Performance, Resiliency**

It is not choosing one over another. Mitigate risks with cost benefit analysis

3 **Leverage Digitization**

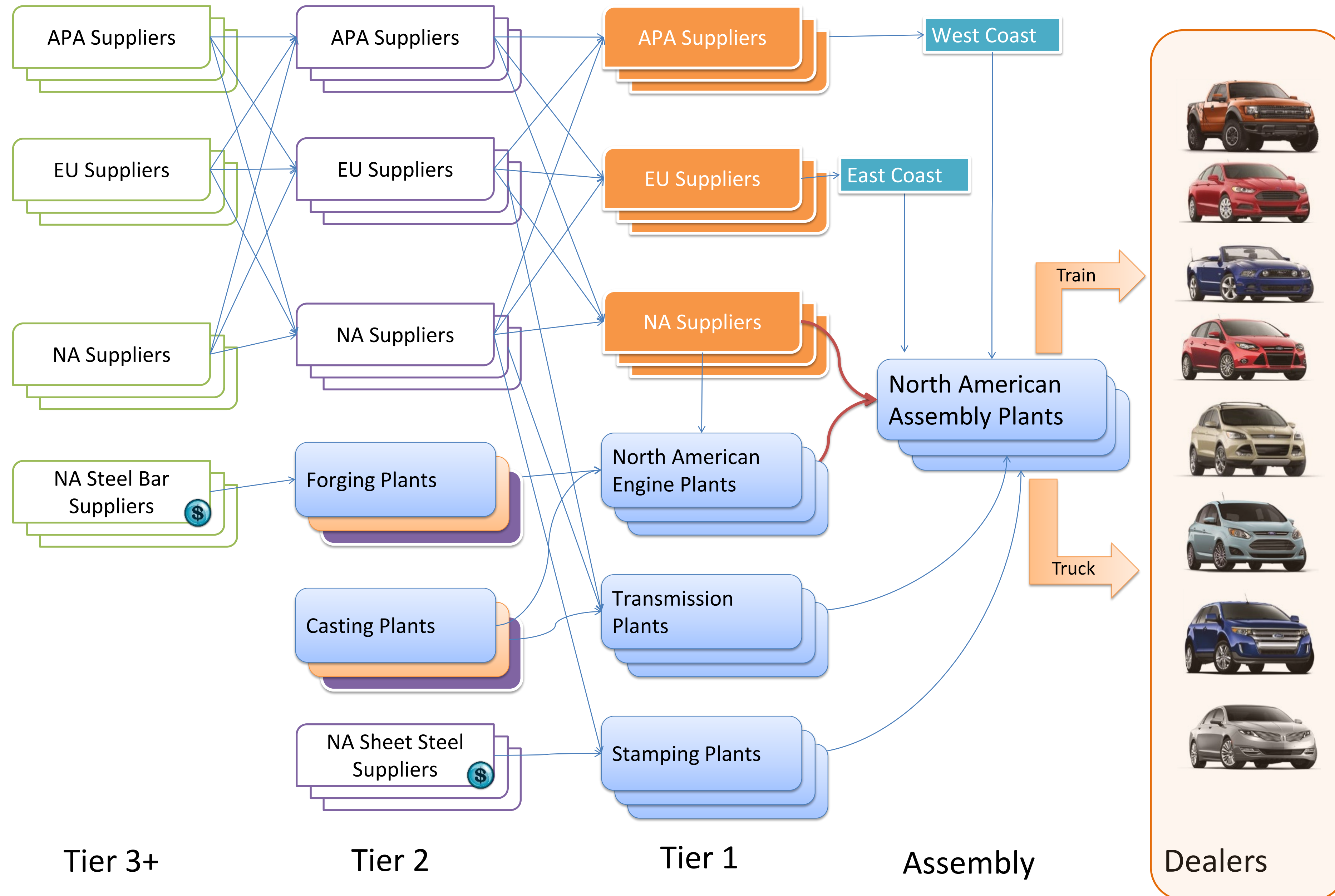Build a Digital Twin of your Supply Chain and continuously reassess risks (TTR, TTS, PI)

4 **Risk is hidden in unexpected places**

Continuously reassess risk

Ford Supply Chain

APA Suppliers

EU Suppliers

NA Suppliers

NA Steel Bar Suppliers

APA Suppliers

EU Suppliers

NA Suppliers

Forging Plants

Casting Plants

NA Sheet Steel Suppliers

APA Suppliers

EU Suppliers

NA Suppliers

North American Engine Plants

Transmission Plants

Stamping Plants

West Coast

East Coast

North American Assembly Plants

Train

Truck

Dealers

Tier 3+

Tier 2

Tier 1

Assembly

Supply Chain Resiliency: Ford Implementation

Steel Bar Suppliers

Raw Chemical Suppliers

TTR =2 Weeks

PI = $400M

Sheet Steel Suppliers

Contract Manufacturers

Assembly Suppliers

Stamping Plants

Engine Plants

Assembly Plants
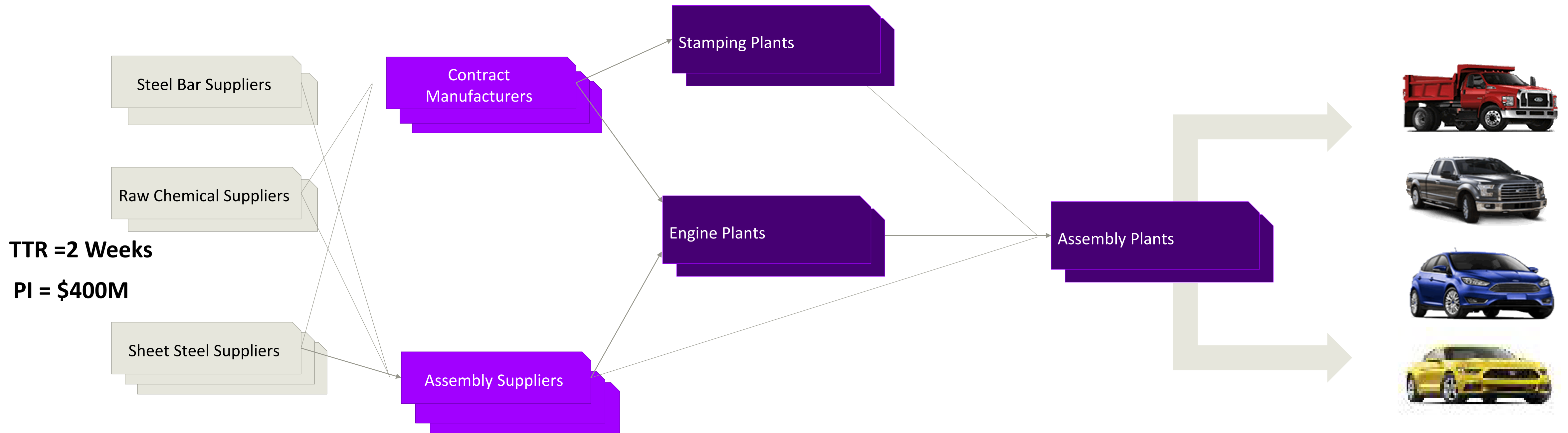
Time-To-Recover (TTR): The time it takes to recover to full functionality after a disruption

Performance Impact (PI): Impact of a disruption for the duration of TTR on a performance measure

Ford Data: Risk Exposure by Supplier Site

Insight 1: Hidden Risk

Total Spend at Supplier Site

Performance Impact (Lost Profits, $000)

0   50,000   100,000   150,000   200,000   250,000   300,000   350,000

Insight 2: Risky suppliers are everywhere
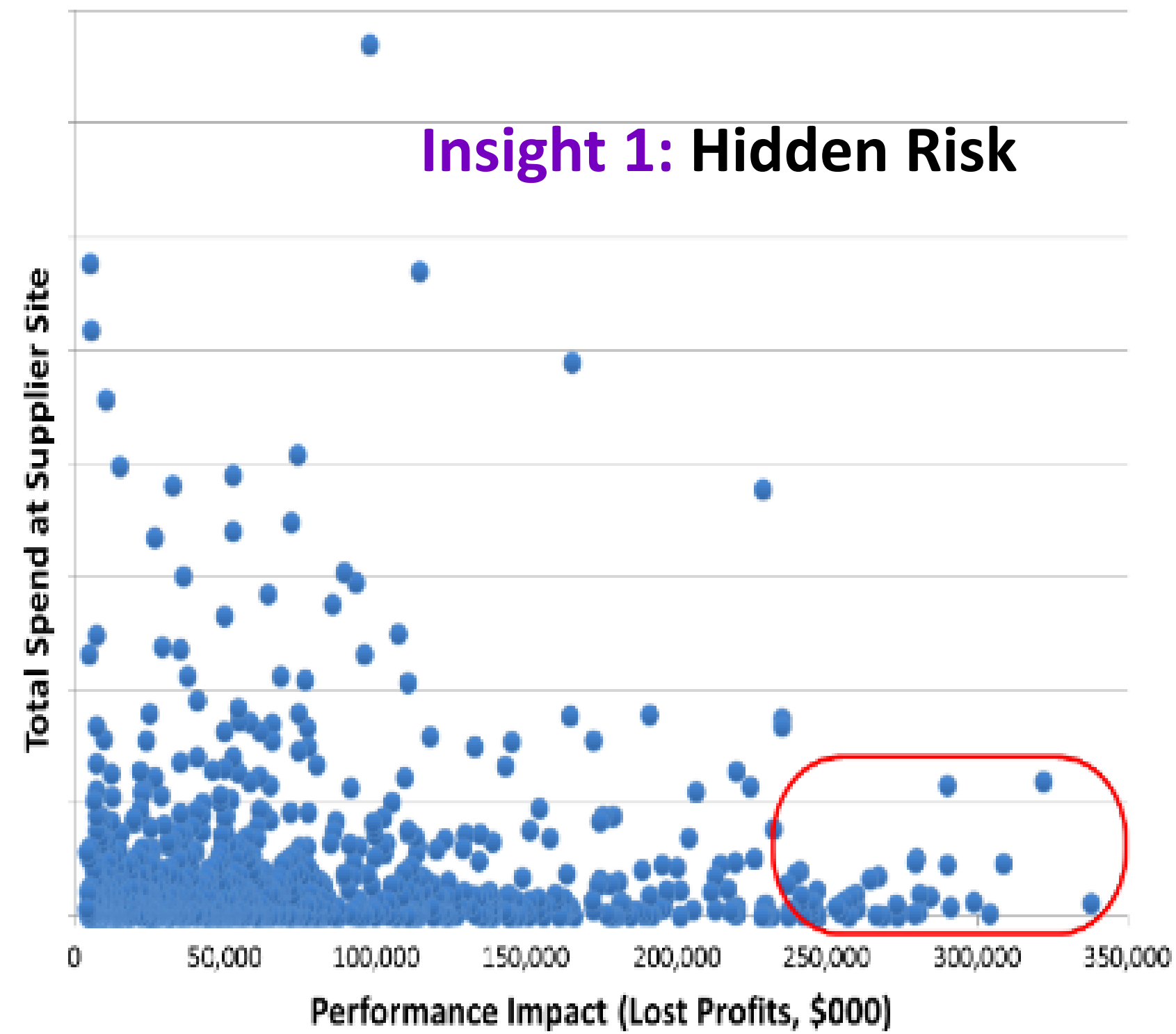
Time-to-Survive across all Ford Tier 1 suppliers
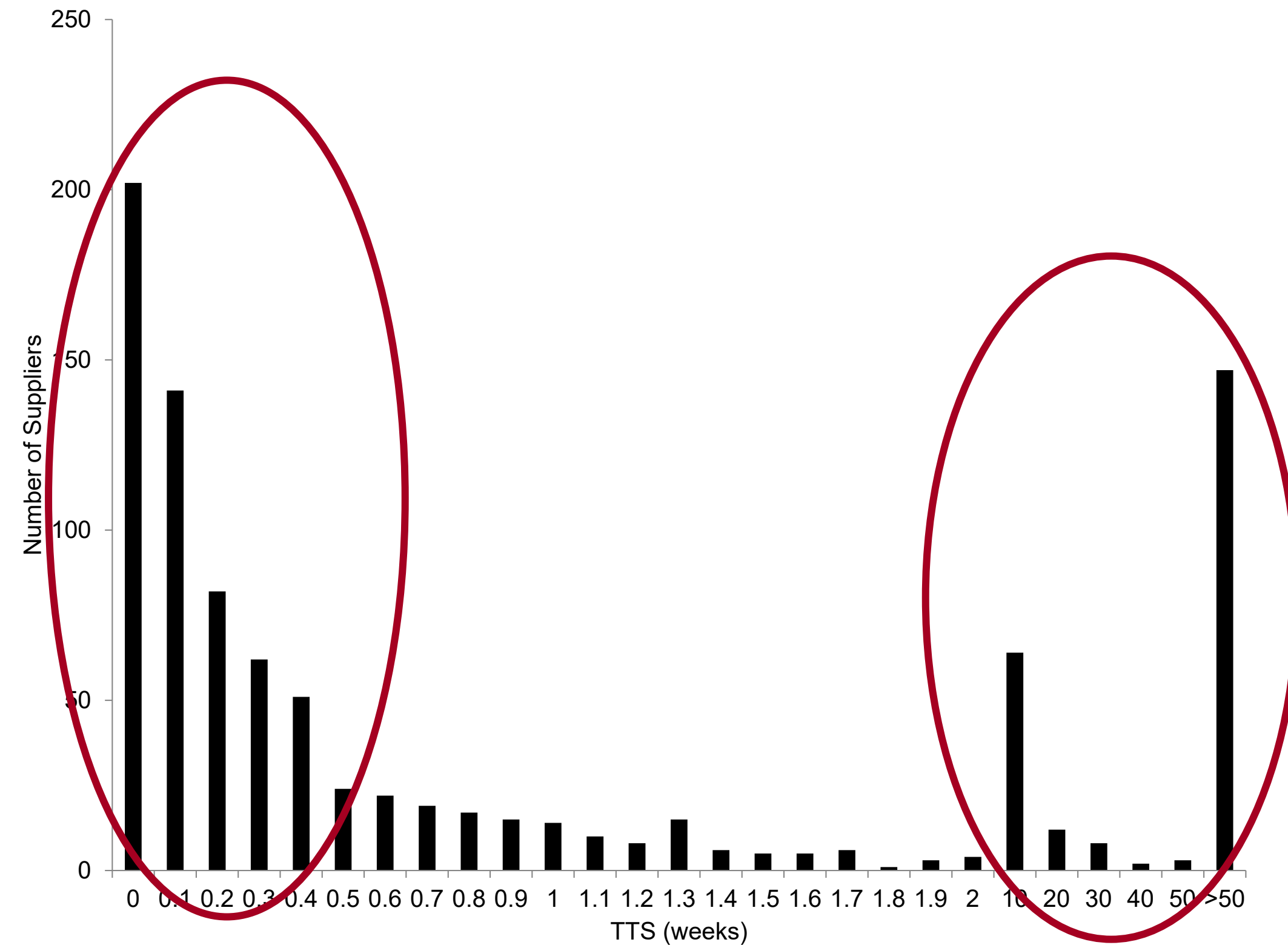
Insight 3: Facilities with short TTS values will lead to immediate performance deterioration; facilities with long TTS values provide saving opportunities

Time-to-Survive (TTS): The maximum duration that the supply chain can match supply with demand after a disruption

Supplier Sites Segmentation
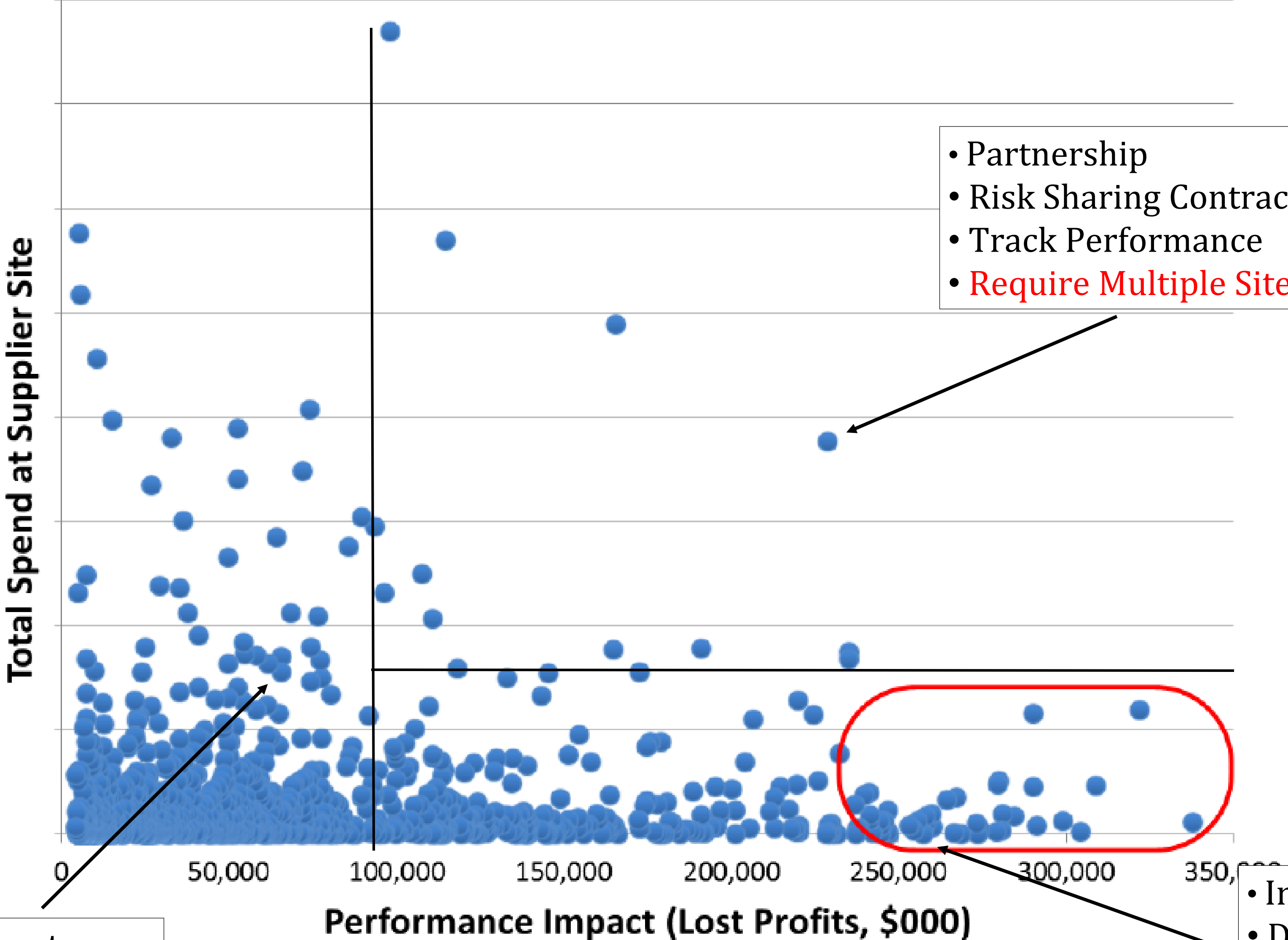
- Partnership
- Risk Sharing Contracts
- Track Performance
- Require Multiple Sites

- Long Term Contracts
- Track Inventory

- Inventory
- Dual Sourcing
- New Product Design

Total Spend at Supplier Site

Performance Impact (Lost Profits, $000)

0   50,000   100,000   150,000   200,000   250,000   300,000   350,000

Go Beyond Just Risk 'Identification' and Invest in a Digital Twin

Understand the financial impact on the business

Look around the corner to identify the risk

Continue to reassess risk (TTR, TTS, PI) – leverage digitalization

In a 2018 interview for Automotive News with Hau Thai-Tang, Ford Exec VP for product Development and Purchasing:

- Ford prides itself on having detailed information about each supplier and which one is susceptible to disruption including calculation of recovery time and financial impact.
- "If you can do that analysis across your value chain, you can learn how to prioritize and manage risk," says Thai-Tang.

Data Driven Business Transformation

- **Supply Chain Resiliency**
  - **Ford Case Study**
- **Online Pricing for Fashion Retailers**
  - **Zalando Case Study**
- **Conclusions**

Easy ways to pay
More choices, less chore

Where would you like to start?

| Women | Men | Kids |

Women's Clothing

Clothing

Dresses
T-shirts & tops
Shirts & Blouses
Knitwear & Cardigans
Sweatshirts & Hoodies
Jackets & Blazers
Coats
Jeans
Trousers
Shorts
Jumpsuits
Skirts
Underwear
Nightwear & Loungewear
Socks & Tights
Swimwear
Sportswear

Sort by   Size   Brand   Colour   Sustainability   Price   Shop

Material   Multipack   Licensed characters   Pattern   Show all filters

349,580 items

Sponsored

Deal
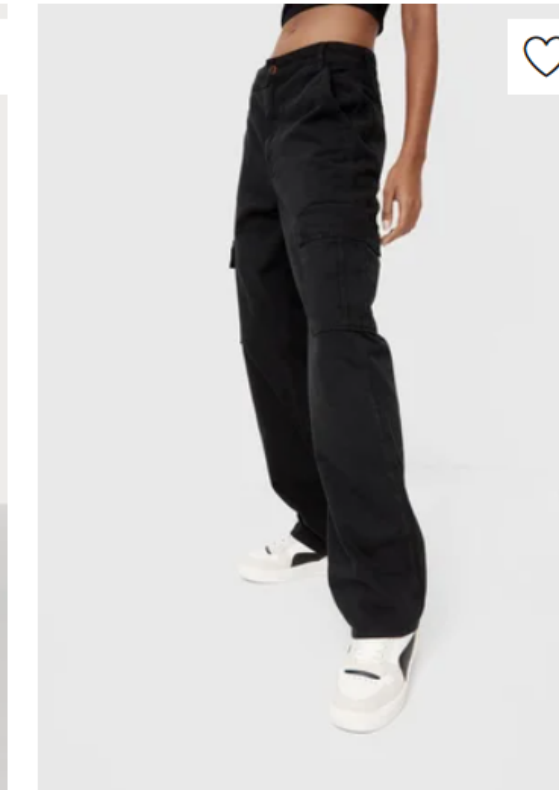
Kids' Clothing

Clothing

T-shirts & tops
Outerwear
Suits & Blazers
Dresses
Sweatshirts & Knitwear
Jeans
Trousers
Shorts
Onepieces & Sets
Skirts
Sportswear
Underwear & Nightwear
Socks
Swimwear
Baby gifts

Sort by   Gender 1   Age 2   Size   Brand   Colour   Sustainability

Price   Shop   Material   Multipack   Show all filters

41,237 items

Sponsored   Sponsored   Sponsored

Deal   Sustainability   New   Deal

Men's Designer Fashion on Sale

Sale
Clothing
Shoes
Sport
Accessories
Designer
  Clothing
    T-shirts & Polos
    Shirts
    Knitwear & Cardigans
    Sweatshirts & Hoodies
    Jackets & Coats
    Jeans
    Trousers
    Shorts
    Suits
  Shoes
  Bags
  Accessories

Sort by   Price 1   Size   Brand   Colour   Sustainability   Savings

Material   Pattern   Length   Sleeve length   Show all filters

3,869 items

Deal   Deal   Deal
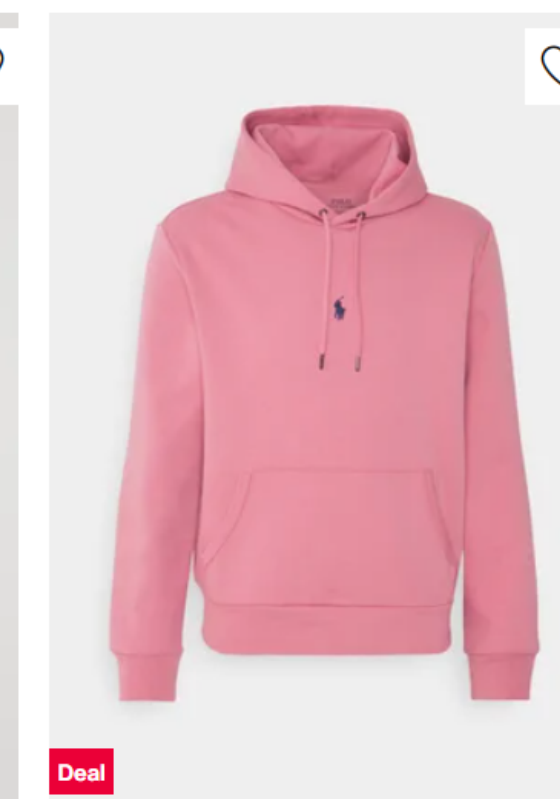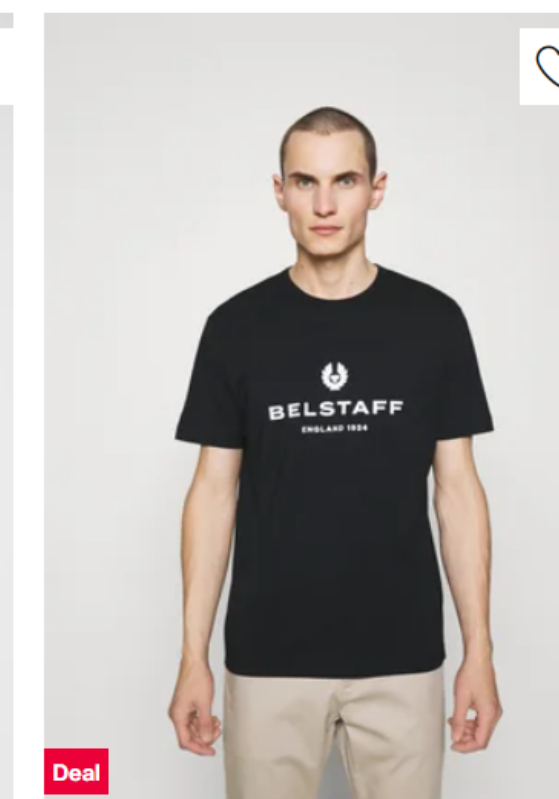
Zalando by the Numbers

- 48 million active customers in 23 countries

- 1.4M articles (SKUs)

- 5,800 brands

- 2021 revenue of 10 billion Euro

- Five primary value proposition: Accessibility, Price, Customization, Innovation, and Brand.


SKU= Unique color, multiple sizes

Zalando Pricing Challenges

- Large Scale
  - Planning period of 40 weeks
- Single SKU Characteristics and Requirements
  - Max-Min Discounts; Max Upwards/Downwards Discount Step
  - Long tail distribution of products
- Emphasize Customer Experience
  - Stock Hedging: Must satisfy demand whenever there's available inventory
  - Size Availability: Matching supply and demand is difficult when inventory is low
- Global Business constraints
  - Average discount for specific product categories
  - Revenue targets by country and by product category

-34%

basically you
Jumper - moosgrün

32,99 € 49,99 €

Global Optimization Challenges

- Single-SKU discount optimization problem is already non-trivial
- Global business constraints tie SKUs together

	Country 1	Country 2	...	Country 23
Category 1				
Category 2				
Category 3				
Category 4				

→ Category Target 1
→ Category Target 2
→ Category Target 3
→ Category Target 4

↓ Country Target 1   ↓ Country Target 2   ↓ Country Target 23

-34%

basically you
Jumper - moosgrün
32,99 € 49,99 €

Zalando Pricing Challenges

- Large Scale
  - Planning period of 40 weeks

- Single SKU Characteristics and Requirements
  - Max-Min Discounts; Max Upwards/Downwards Discount Step
  - Long tail distribution of products

- Emphasize Customer Experience
  - Stock Hedging: Must satisfy demand whenever there's available inventory
  - Size Availability: Matching supply and demand is difficult when inventory is low

- Global Business constraints
  - Average discount for specific product categories
  - Revenue targets by country and by product category

- **Business Goals**
  - **Optimize discounts to maximize profit while satisfying all business requirements**
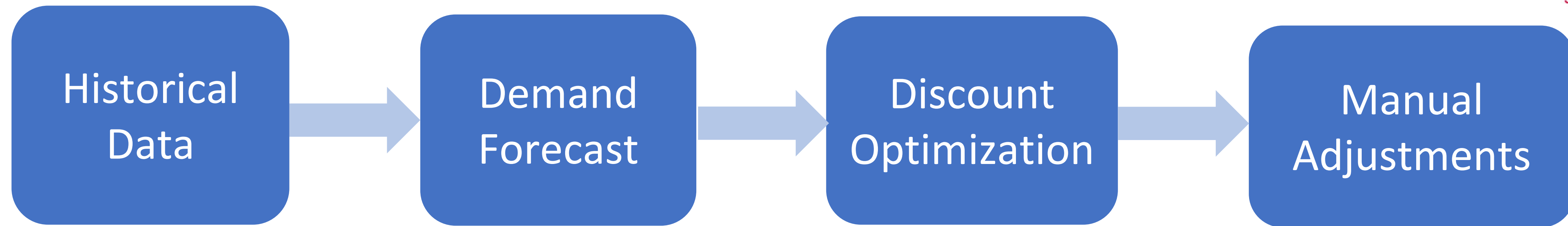
-34%

basically you
Jumper - moosgrün

32,99 € 49,99 €

Zalando Old Process


```
┌──────────────┐     ┌──────────────┐     ┌──────────────┐     ┌──────────────┐
│  Historical  │ ──▶ │    Demand    │ ──▶ │   Discount   │ ──▶ │    Manual    │
│     Data     │     │   Forecast   │     │ Optimization │     │ Adjustments  │
└──────────────┘     └──────────────┘     └──────────────┘     └──────────────┘
```

- **Historical sales, prices**
- **Stock level**
- **Product Features**

- **Country-Week-SKU-Discount level**
- **23 Countries, 40 weeks**
- **Time Series Analysis**

- **Singe SKU Model**
- **Goal: Max profit**
- **Gradient-based Optimization**
- **Cannot satisfy Global Requirements**

- **Rely on business experience**
- **Satisfy global business targets**

The New Approach

- **Goal:** Automate the pricing process through the sales season, so that profit is maximized while satisfying all business requirements.

Historical Data	Demand Forecast	Aggregated Model	Category Level Optimization

- Historical sales, prices
- Stock level
- Product Features

- Country-Week-SKU-Discount level
- 23 Countries, 40 weeks
- Neural Network

- Clustering to aggregate SKUs
- Optimization Techniques
- Decouple by product category

- Piecewise Linear Approximation
- MIP using Cutting Plane
- Lower Bound for Approx. Accuracy

basically you
Jumper - moosgrün
32,99 € 49,99 €

-34%

Demand Forecast Model

- Demand forecast by SKU, country, week and discount level
- Unique feature: high return rates

Performance Comparison for Demand Forecast Models

Models	1st Week Error	Seasonal Error	Bias
Time Series Model	0 (calibrated)	0 (calibrated)	14%
LSTM	-30%	+9%	12%
Transformer	-30%	-9%	-3%

Aggregation Framework

1. **Clustering:** group SKUs with similar features (e.g. demand, price, inventory...) for each category

2. **Aggregation:** create a dummy SKU (e.g. sum demand, weighted average price...) for each cluster

3. **Optimization:** Solve MIP with dummy SKUs to obtain category-country specific targets.

Advantages

- Reduce problem size to allow fast computing

- Decompose the optimization problem into smaller, detailed models by categories
- Develop high level business insights

Aggregation Framework

- We approximate the problem by aggregating the SKUs, and solving a higher-level problem, to obtain category-specific targets

	Country 1	Country 2	...	Country 23
Category 1				
Category 2				
Category 3				
Category 4				

Cluster SKUs in each category into 50 groups

Category Target 1
Category Target 2
Category Target 3
Category Target 4

Country Target 1
Country Target 2
Country Target 23
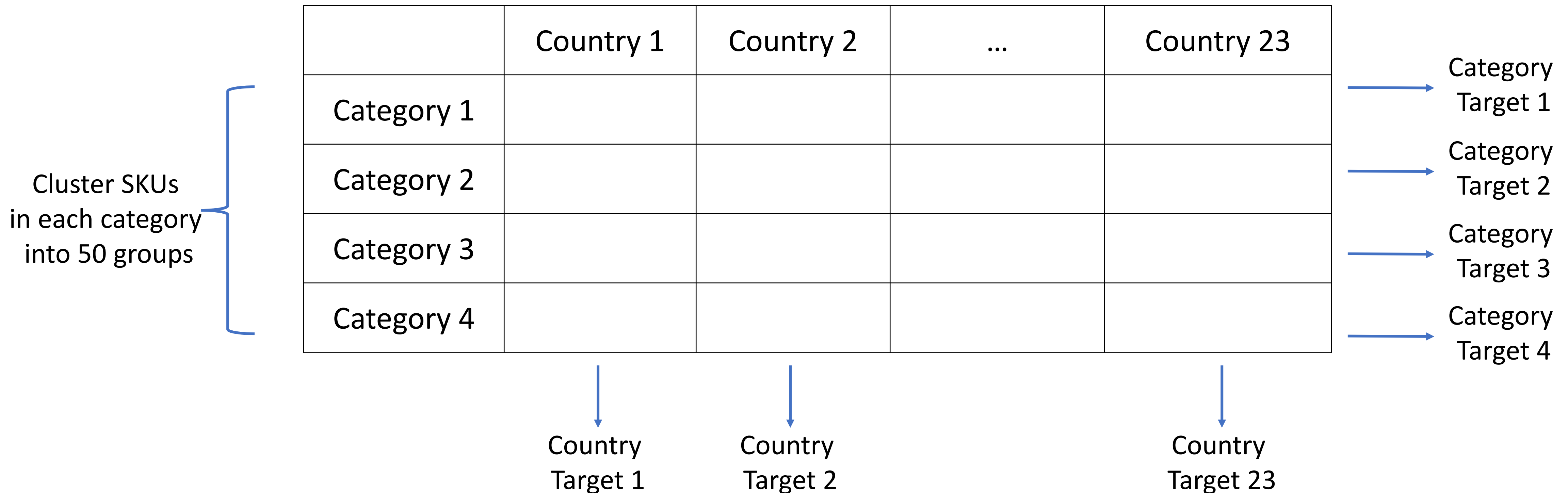
Aggregation Framework

- We approximate the problem by aggregating the SKUs, and solving a higher-level problem, to obtain category-specific targets

- We can then decompose the problem into a model for each category
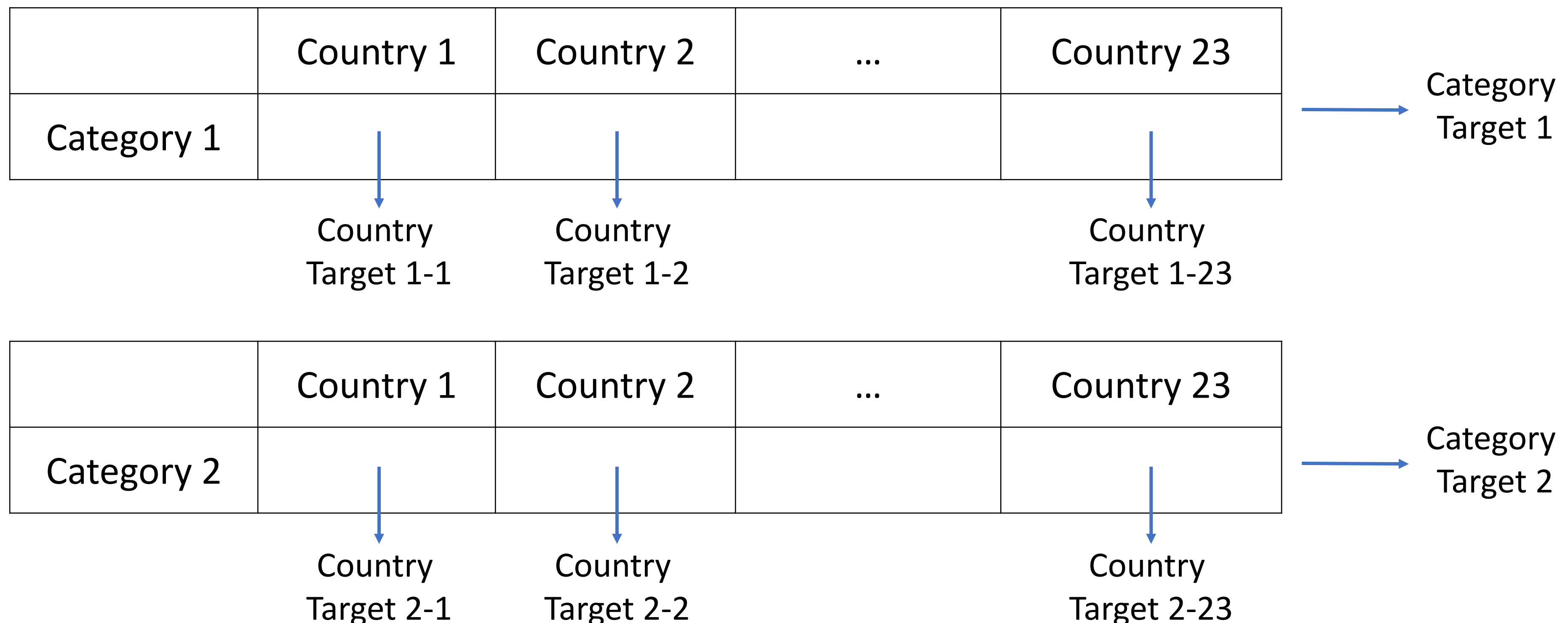
	Country 1	Country 2	...	Country 23
Category 1				

Category Target 1

Country Target 1-1   Country Target 1-2   Country Target 1-23

	Country 1	Country 2	...	Country 23
Category 2				

Category Target 2

Country Target 2-1   Country Target 2-2   Country Target 2-23

Aggregation Accuracy

- We compare the optimal discount distribution between
  - the benchmark (directly solving the MIP), and
  - the aggregated model (solving the MIP with inputs from the aggregated model)
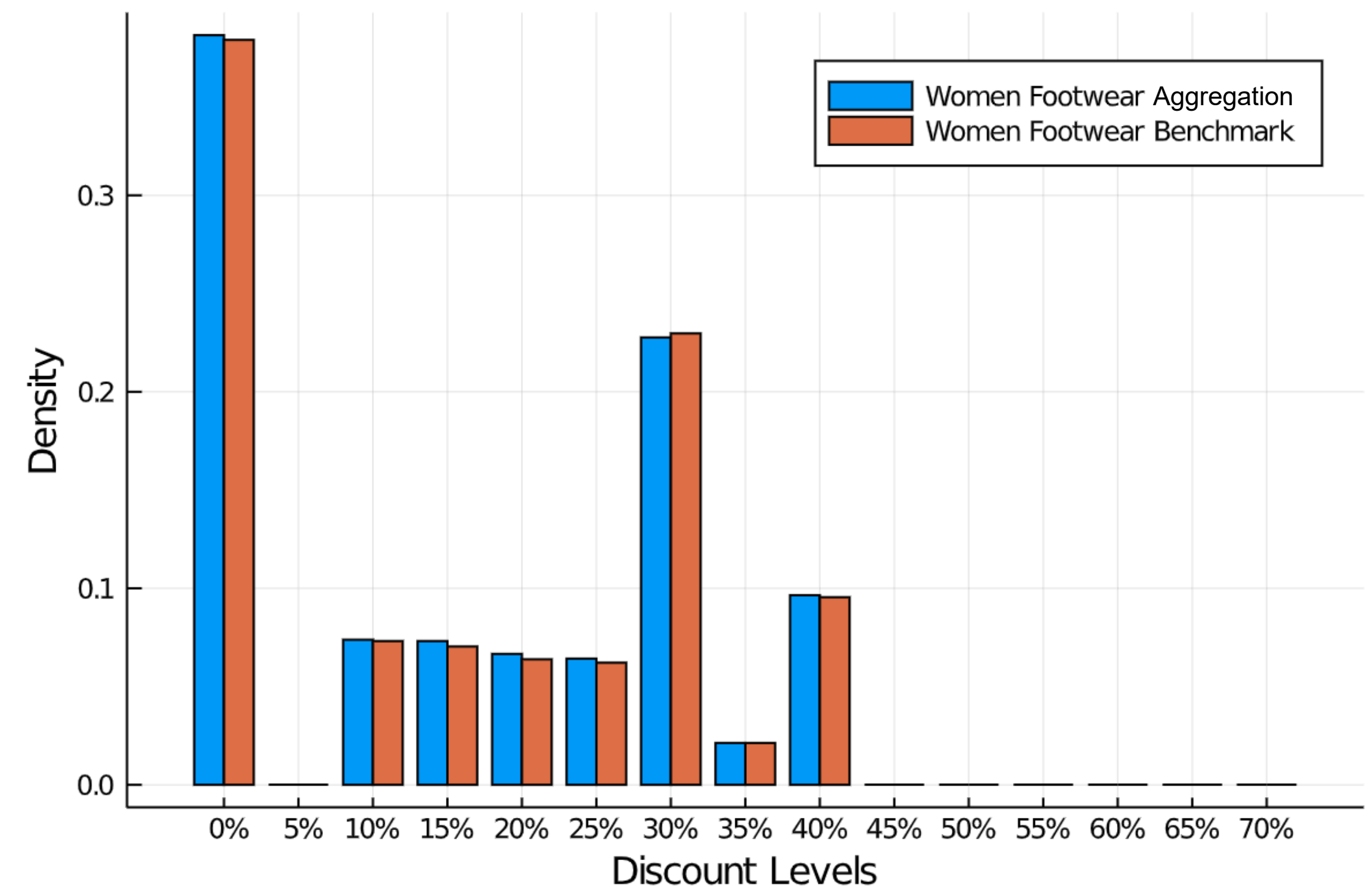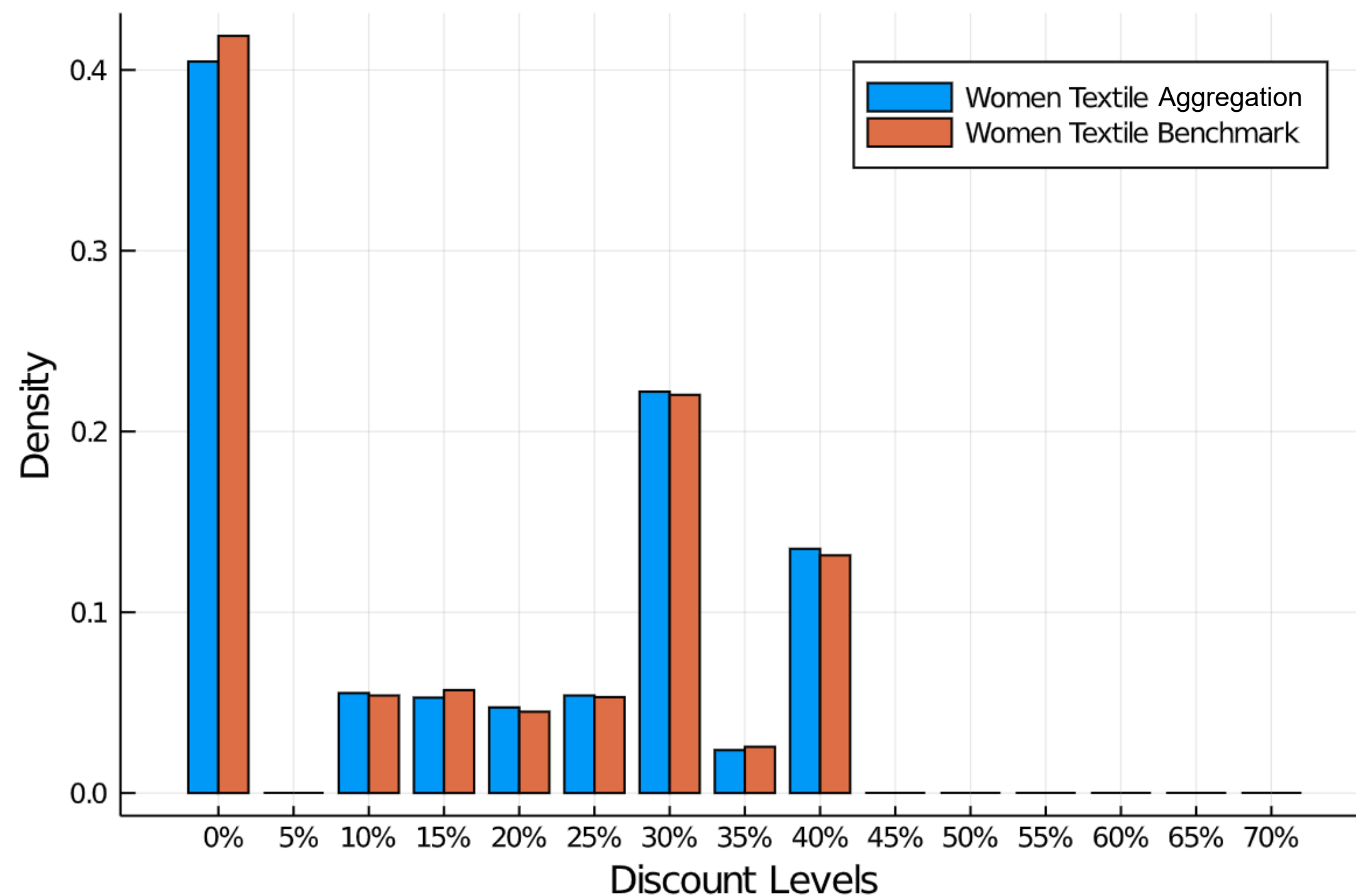
Aggregation Accuracy

- We compare the optimal discount distribution between
  - the benchmark (directly solving the MIP), and
  - the aggregated model (solving the MIP with inputs from the aggregated model)
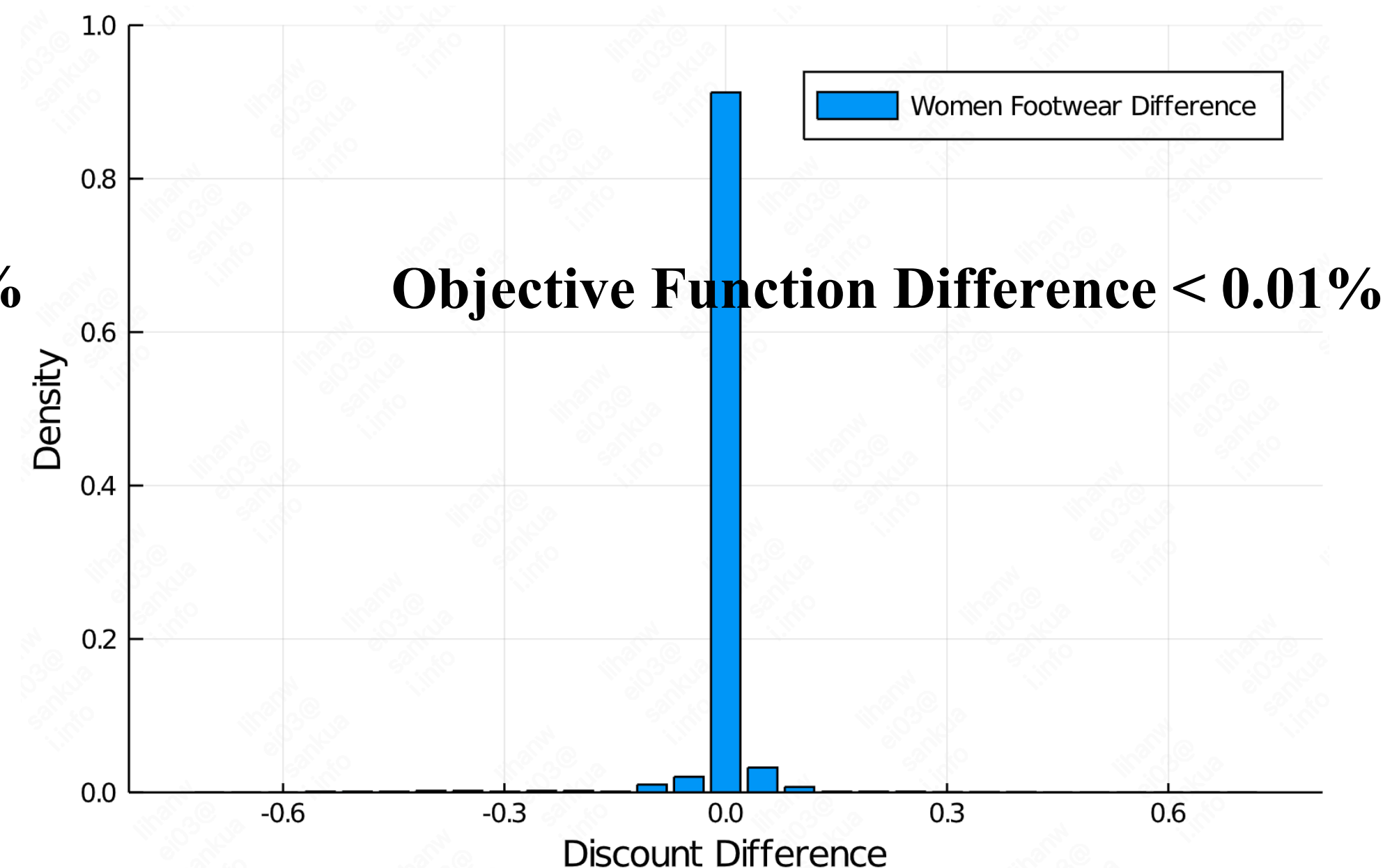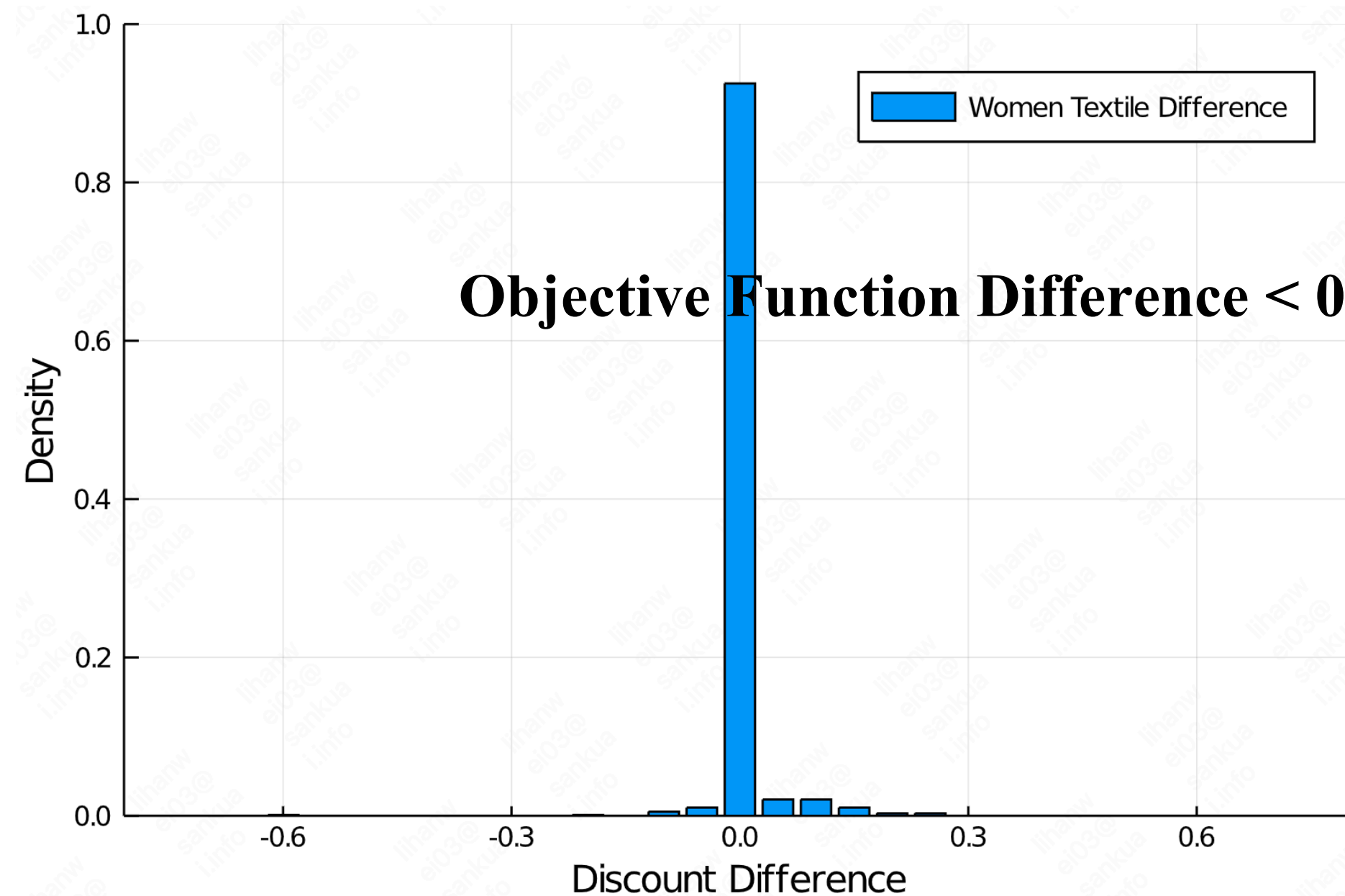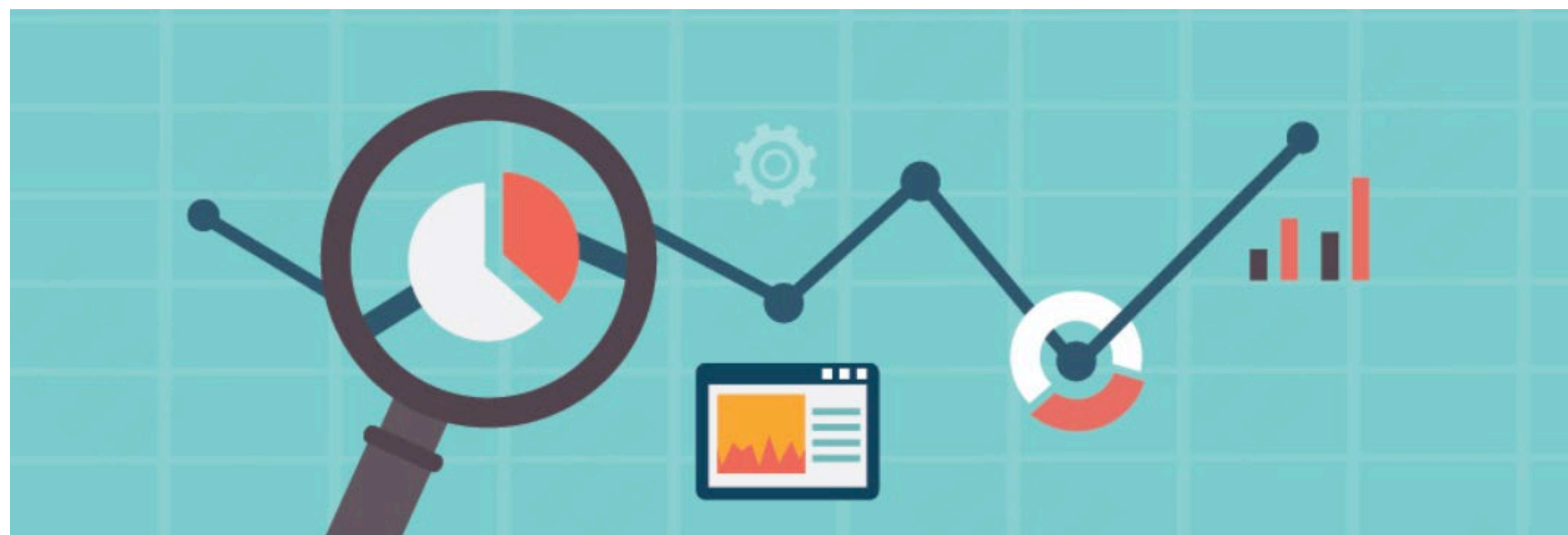

Discount Difference by SKU-Country in the First Week

From Aggregated Model to Planning Tool

- Impact of global targets on Zalando's profit

- Tradeoffs between Target, Profit and Market Share

- Impact of budget constraints on Profit

- Sensitivity to available inventory

Customer Experience: Stock Hedging

- Maximizing profits is not the only objective, we care about customer experience

- Zalando must satisfy demand whenever there's available inventory: it cannot reserve inventory for other (more profitable) countries or later time periods

- Sales equals demand if stock is sufficient

- Sales equals stock distributed proportionally by country if stock is insufficient

- Incorporated into the MIP

Customer Experience: Size Availability

- Since the model is at the SKU level, and each SKU includes multiple sizes, it is not clear that the system can match supply with demand when inventory is low

- Using historical data, we approximate stock availability by

$$\mathrm{sr}_t(y_t) = 1 - \exp\left(-\alpha \cdot (y_t/\mathcal{N})^{\beta}\right)$$

  - sr= Stock Response: a ratio between 0 and 1 to approximate size availability
  - y: stock level
  - N: number of different sizes
  - α, β: parameters tuned by using historical data

Customer Experience: Size Availability

- We adopt the piecewise linear approximation method

- There is a tradeoff in approximation goodness and running time (MIP complexity), we observe that 4 to 5-piece linear approximation is good enough

Performance Comparison between Piecewise Linear Approximation Schemes

Methods: PL_x	Relative Difference in Objective
PL_2	9.03%
PL_3	5.35%
PL_4	1.76%
PL_5	1.07%
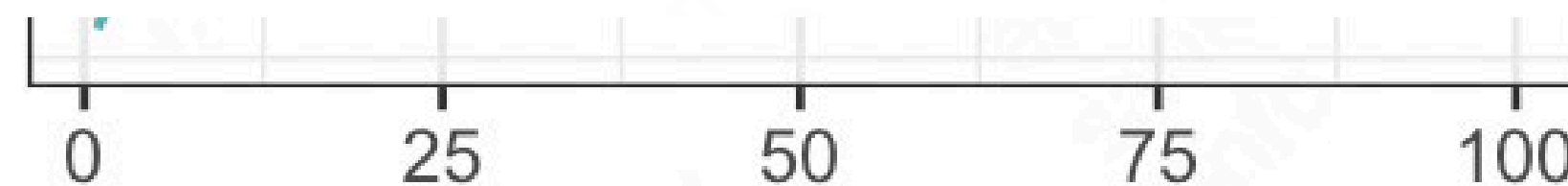
Offline and Field Experiments

- Offline:

Experiments	SKUs	Iterations	Run Time (m)	Number of Cores	Total Memory	Target Deviation
1	51745	5	105	2000	4.0	0.00
2	12798	5	90	1500	3.0	0.00

- Online (field experiment):

Experiments	SKUs	Iterations	Run Time	Actual Target deviation	
				Test	Control
1	12632	4	1h	**0.00%**	-3.27%
2	12757	1	40m	**0.00%**	0.00%
3	8961	4	1h	-2.85%	-10.43%

Neural network meets Optimization: Impact

Black Friday 2021:

- Pilot system used to discount the assortment by having daily demand forecast
- Popular products disappear in hours; late customers maybe frustrated
- By providing higher discounts to articles that can substitute popular ones, the system increased the visibility of substitute articles and thus diversify sales.

Results:

- Significant improvement in customer satisfaction regarding discount and availability throughout the entire Black Friday event.

zalando

Zalando SE, 11501 Berlin

Berlin, 02.05.2022

RE: MIT x Zalando submission to 2022 RM&P Section Practice Award

Dear Colleagues,

It is with great enthusiasm that I am writing this letter to serve as an evidence of the practical impact of the paper "Large-scale Price Optimization for an Online Fashion Retailer" by Li H., Simchi-Levi D., Sun R., and Wu M. X from MIT and Fux V., Gellert T., Greiner T., and Taverna A. from Zalando.

Let me start by introducing myself. Until recently I served as the VP of Pricing and Forecasting at Zalando based in Berlin, Germany. Zalando is a European online fashion retailer with 48 million active customers in 23 countries offering 1.4m articles (styles) from more than 5,800 brands and a revenue of 10.4bn Euros (2021). Currently I am the VP of Product Data & Experience. During our collaboration with the MIT team, my main role was leading the pricing team and forecasting team to define pricing strategy, product development, annual plans and weekly commercial actions.

Starting July 2018, a team from the MIT Data Science Lab, led by Professor Simchi-Levi, has been collaborating with my team at Zalando on utilizing the massive amount of data we have to optimize price discount decisions over a large number of products in multiple countries on a weekly basis.

The project had multiple challenges: First, this a large-scale problem with millions of SKUs across 23 countries and where the planning horizon is 40 weeks. Second, there are a variety of business requirements such as constraints on the average discount across product categories; revenue targets by country and by product category, all of which imply that the problem cannot be decoupled by product and country. Third,

The collaboration has made a considerable impact on our business. The core pricing algorithm central to this work has been launched on our platform in 2021. The pilot field experiment empirically validates that the optimization framework successfully steers the discounts towards the business targets, and the model is integrated into the company's weekly operation pipeline. The new system automates the decision-making process for the Zalando online platform and demonstrates the power of integrating statistical learning (neural-network) and optimization to make a big impact on business performance.

Lawrence Jewsbury
VP Product Data & Experience
Zalando SE, Berlin

Data Driven Business Transformation

- **Supply Chain Resiliency**
  - **Ford Case Study**

- **Online Pricing for Fashion Retailers**
  - **Zalando Case Study**

- **Conclusions**

The Future of Management Science Research

- **Emphasize data driven research and teaching**
  - Today, there is too little reliance on data in formulating models and identifying research opportunities
  - Systems involving people can be difficult to analyze unless you have data about behavior
- **Develop new engineering and scientific methods that explain, predict and change behavior**
  - Use Stat, CS, Econ and OR to address operational problems
  - Apply OR techniques to solve open CS and Stat problems
- **Emphasize research that provides better understanding of human-algorithms interactions and how to improve it**

Management Science Special Issue

Call for Papers—*Management Science* Special Issue on the Human-Algorithm Connection

Felipe Caro,[a] Jean-Edouard Colliard,[b] Elena Katok,[c] Axel Ockenfels,[d] Nicolas Stier-Moses,[e] Catherine Tucker,[f] D. J. Wu[g]

[a] Anderson School of Management, University of California, Los Angeles, Los Angeles, California 90095; [b] HEC Paris, 78351 Jouy-en-Josas, France; [c] Naveen Jindal School of Management, University of Texas at Dallas, Richardson, Texas 75080; [d] Department of Economics, University of Cologne, D-50923 Cologne, Germany; [e] Facebook, Inc., Menlo Park, California 94025; [f] MIT Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142; [g] Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30308

Contact: felipe.caro@anderson.ucla.edu, https://orcid.org/0000-0003-0947-3958 (FC); colliard@hec.fr, https://orcid.org/0000-0003-1936-8641 (J-EC); ekatok@utdallas.edu, https://orcid.org/0000-0002-7037-7896 (EK); ockenfels@uni-koeln.de, https://orcid.org/0000-0003-1456-0191 (AO); nicostier@yahoo.com (NS-M); cetucker@mit.edu, https://orcid.org/0000-0002-1847-4832 (CT); dj.wu@scheller.gatech.edu (DJW)

Submission Deadline: September 9, 2022

MIT Data Science Lab

Theoretically Elegant & Practically Relevant Research

David Simchi-Levi
Professor of Engineering Systems, MIT
Home | MIT Data Science Lab
dslevi@mit.edu