August 2023

C0 - PUBLIC

Combining Open Banking & Credit Bureau data in Credit Scoring

enables a marked improvement to assess credit risk for an underserved population in the UK

Oussama Bouraoui & Marc Gaudart

algoan

Agenda

- Project Objectives
- Population profile
- Results
- Approach
- Scorecards

Project Objectives

Develop new credit scorecards for FairFinance credit applicants to improve the discriminatory power of its credit granting process, in particular:

- Take into account the changes in distribution and performance of key features post-pandemic
- Use Open Banking features that were not available for the previous generation of scorecards
- Build separate scorecards for new and existing customers, to take into account the significant differences in profile between the two

Note: this presentation focusses on the scorecards developed for **new customers** (i.e. new to FairFinance)

Who we are

FairFinance is a Social Enterprise with a vision to revolutionise financial services to make it fair and inclusive for everyone.

FairFinance provide affordable personal loans in the UK to individuals who don't have any or have limited access to mainstream finance

Algoan is a Fintech company launched in 2018 and specializing in Open Banking analytics.

Algoan designs, develops and operates technological products enabling the digitalisation and the optimisation of credit applications processes, including supporting the credit decisioning of the lender.

Based on Open Banking data, Algoan's solutions allow lenders to grant more loans while reducing credit risk.

Trent is a consulting firm providing consulting and executive coaching services to leaders in financial services and other industries

Agenda

- Project Objectives
- **Population profile**
- Results
- Approach
- Scorecards

Population profile

Fair Finance loans are short in duration (average 10 months) and amounts are low (average £670).

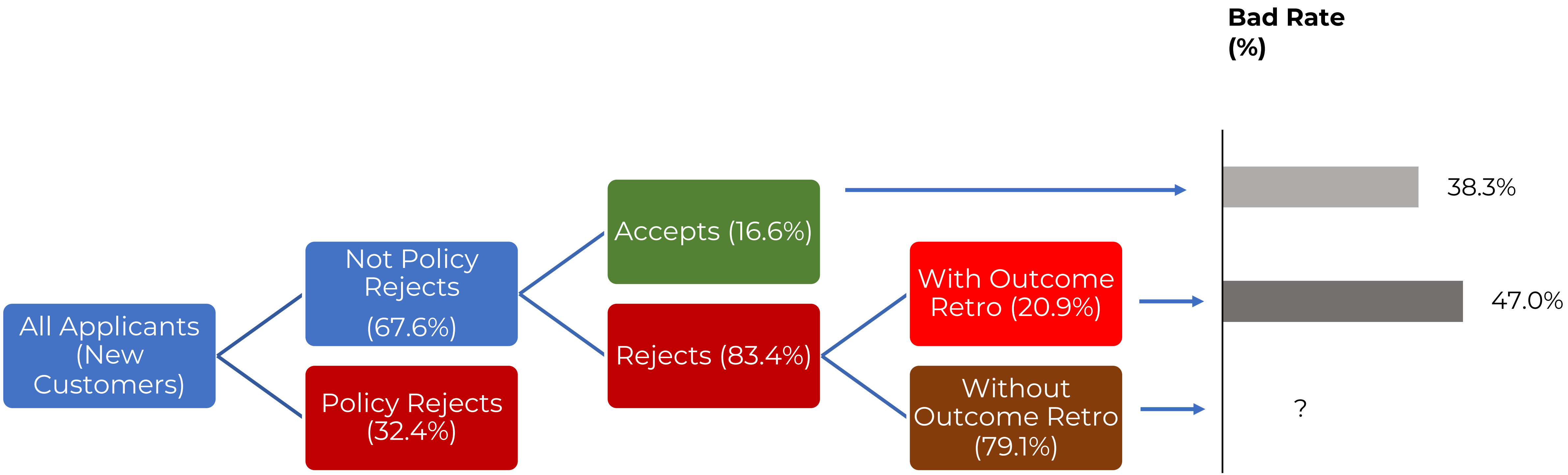
They provide an alternative to pay day loans for an underserved population with a history of financial hardship.

New Fair Finance customers are particularly tricky to serve:

Reject rate: 87%

Bad rate: 38%

Population flow



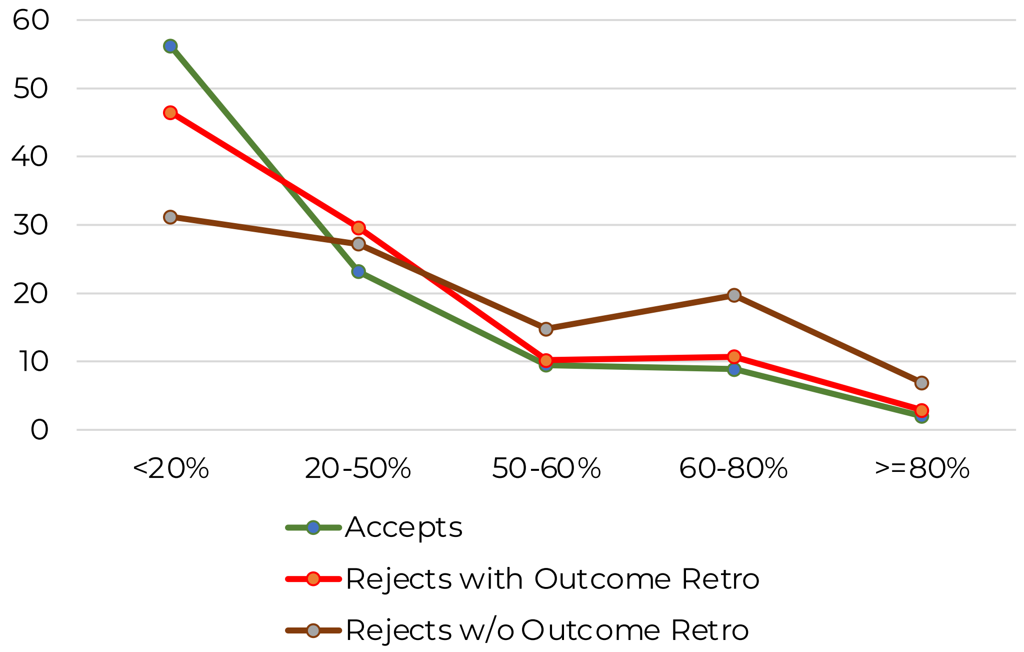
algoan

Example Variable report 1

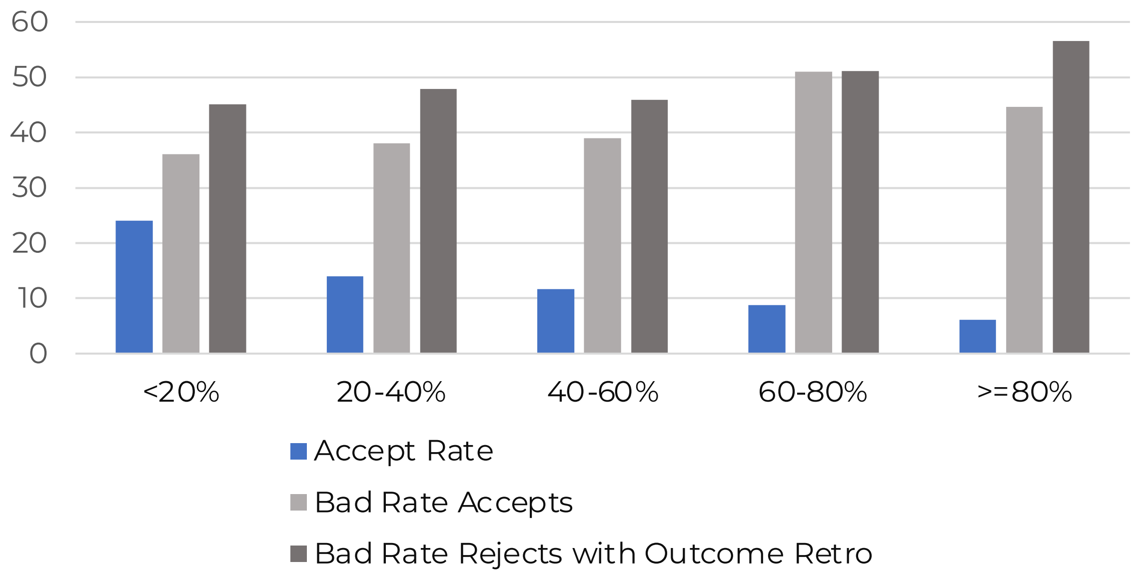
Credit Bureau variable

Default rate on active accounts (Credit Bureau): Number of active accounts where the worst payment status has been 3+ in the last 12 months divided by the number of active accounts. If there are no active accounts, the default rate is 0.

Population Distribution



Accept Rate and Bad Rates

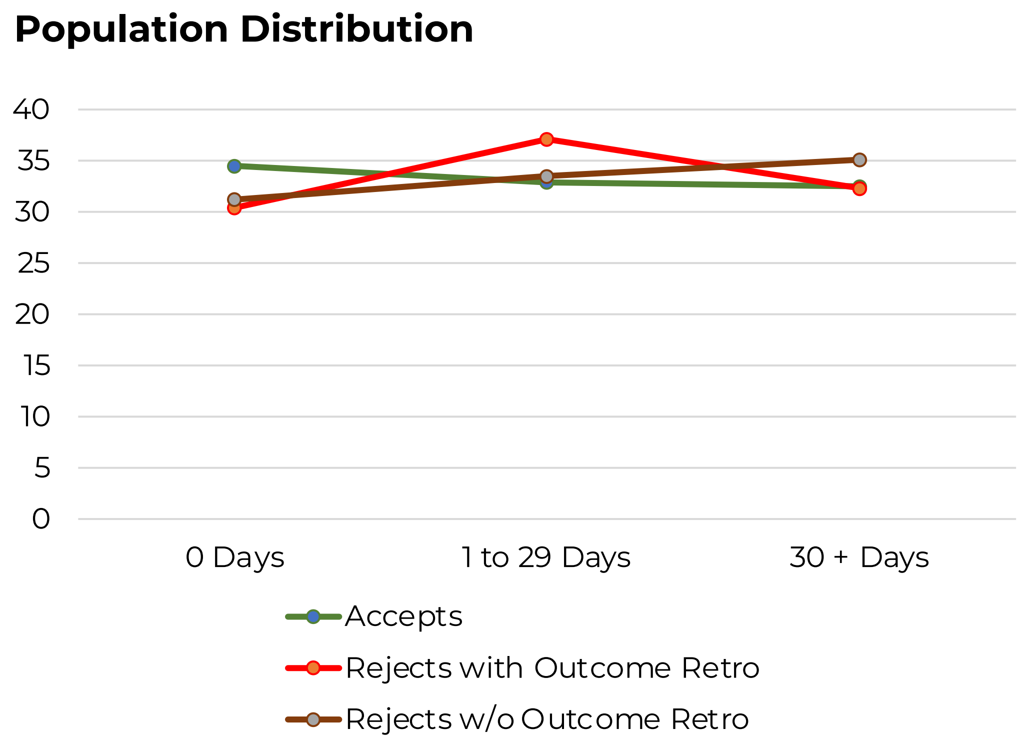


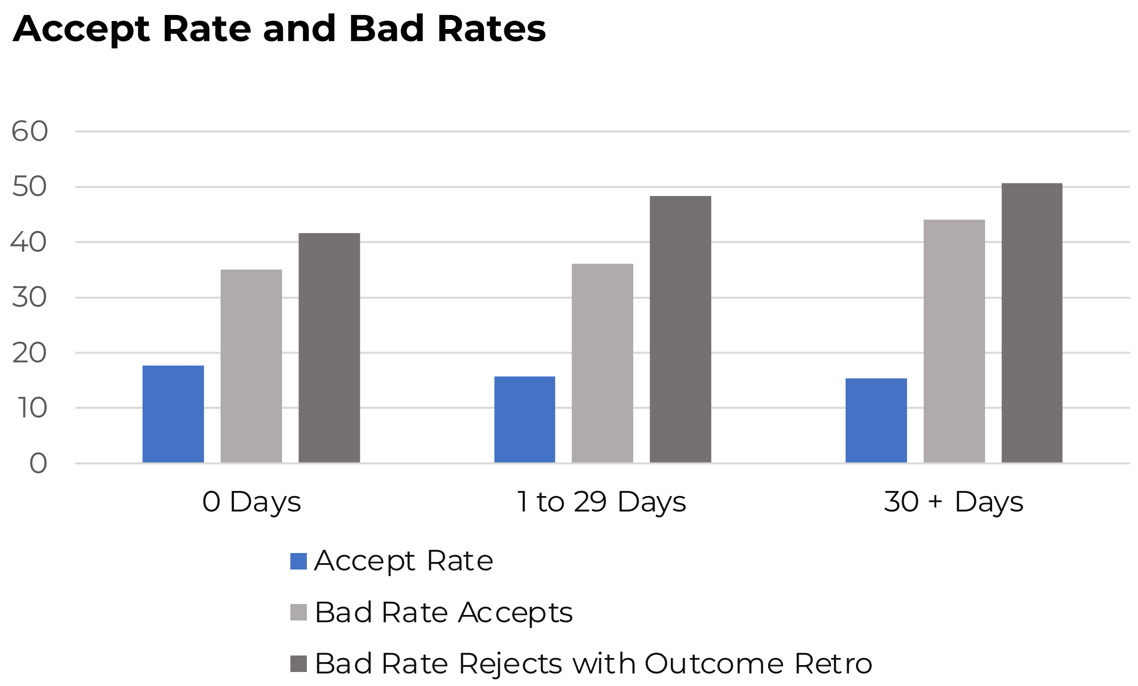
algoan

Example Variable report 2

Open Banking variable

Largest number of consecutive days where the total banking balance is negative.





Agenda

Project Objectives

Population profile

Results

Approach

Scorecards

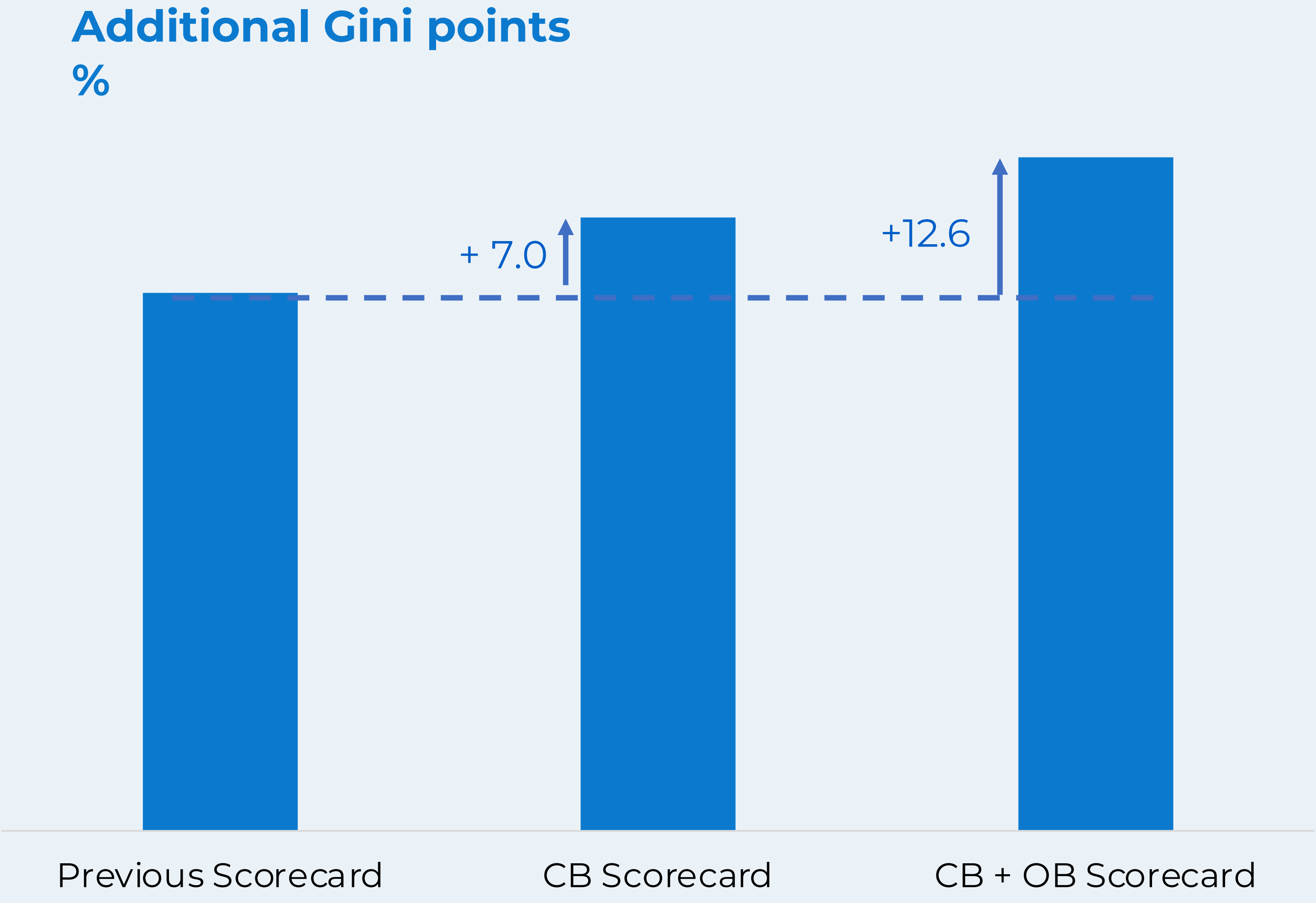
The results

2 scorecards were developed for new customers:

CB data only:	Credit Bureau data
CB + OB data:	Credit Bureau and Open Banking data

The scorecard with CB data only is to be used for brokers-introduced applications, enabling a rapid initial decision at the brokers before adding Open Banking data for the final decision

Results – Gini improvement

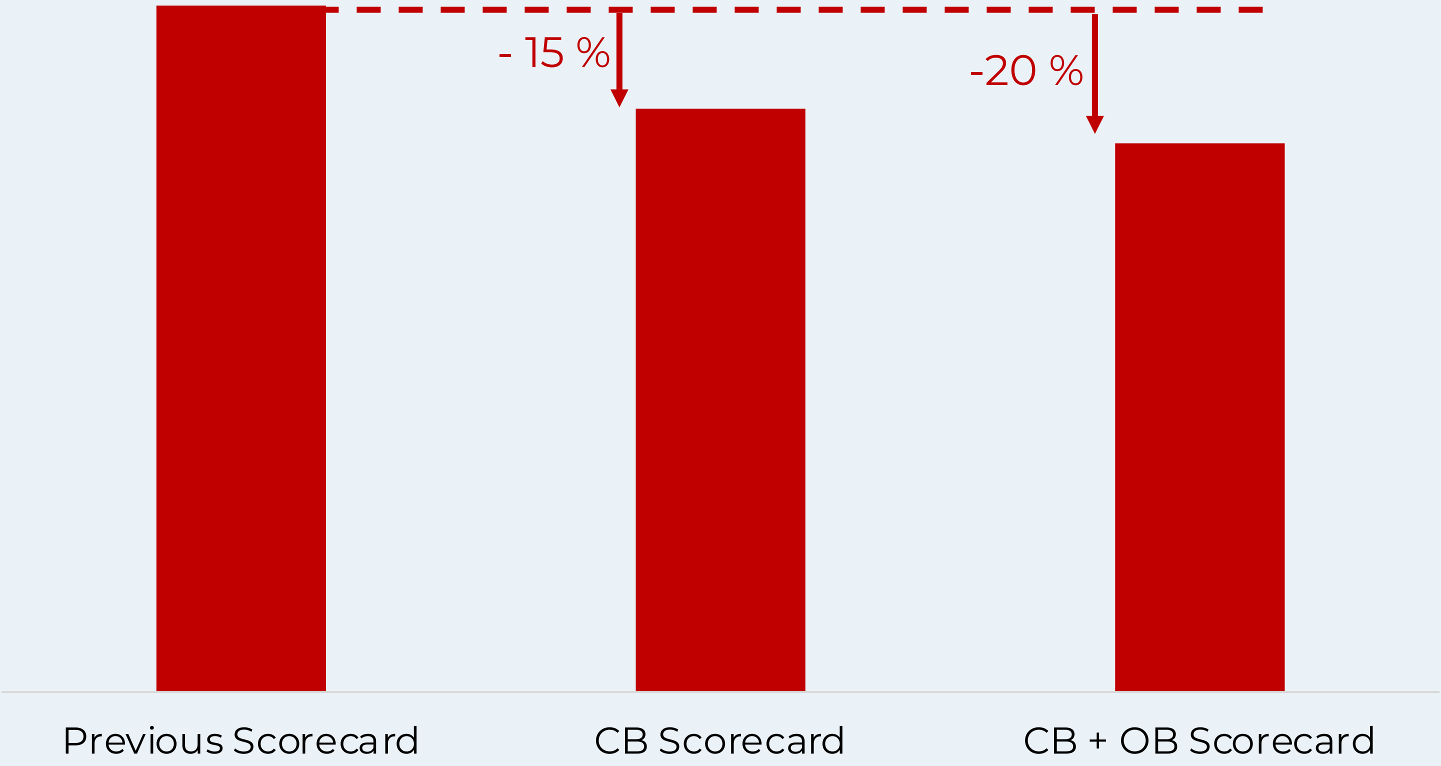


The scorecard with Credit Bureau data shows a marked improvement compared to the previous scorecard,

and the addition of Open Banking variables creates a significant lift to the performance

Results – Swap Sets

Reduction in Bad Rate at constant Accept Rate %

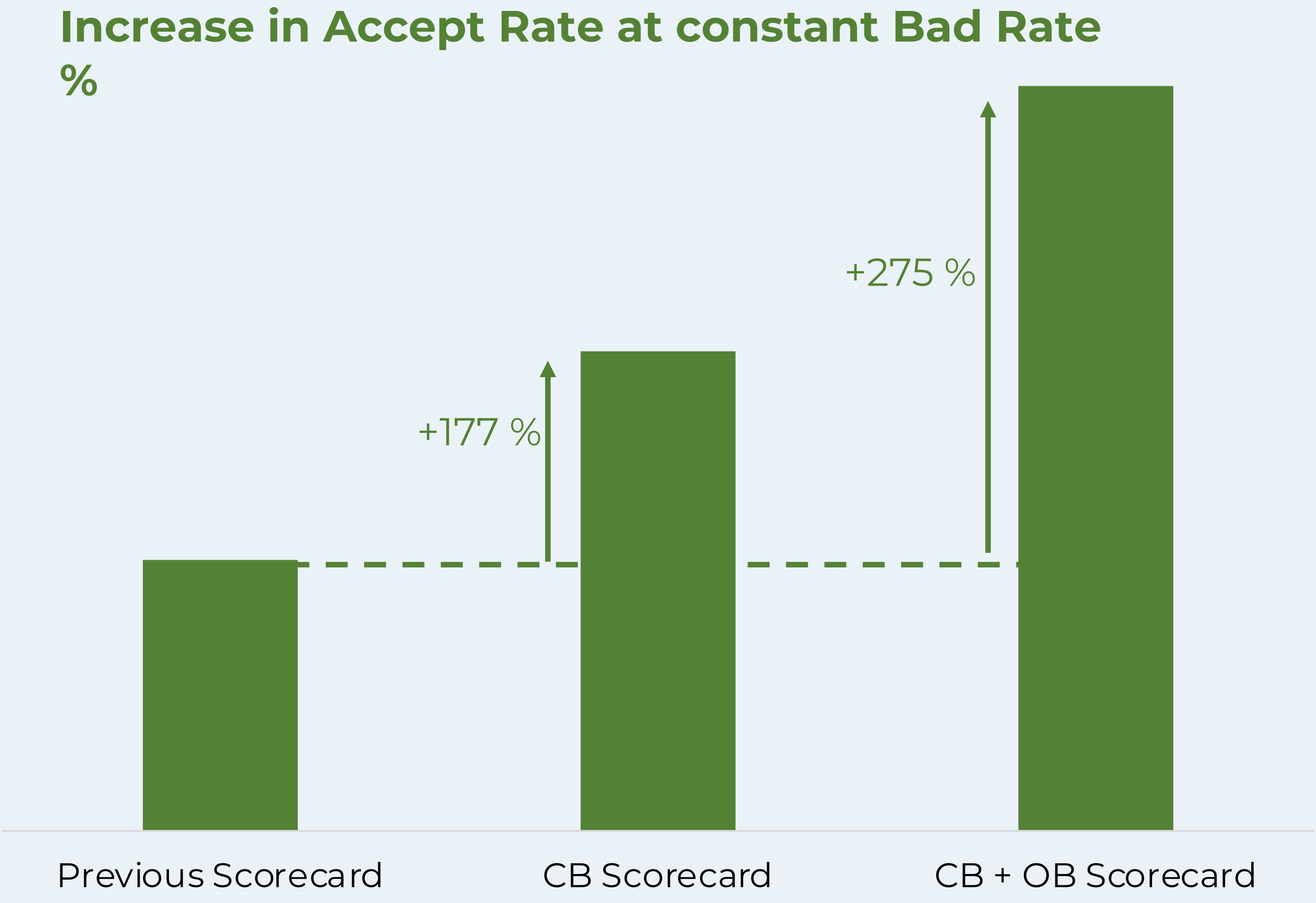


A cut-off is chosen for each new scorecard that enables the same Accept Rate as previously:

the scorecard with Credit Bureau data enables a significant decline in Bad Rate,

here again, the addition of Open Banking variables creates a significant lift to the performance

Results – Swap Sets



A cut-off is chosen for each new scorecard that enables the same Bad Rate as previously:

The scorecard with Credit Bureau data enables a significant increase in Accept Rate,

here again, the addition of Open Banking variables creates a significant lift to the performance

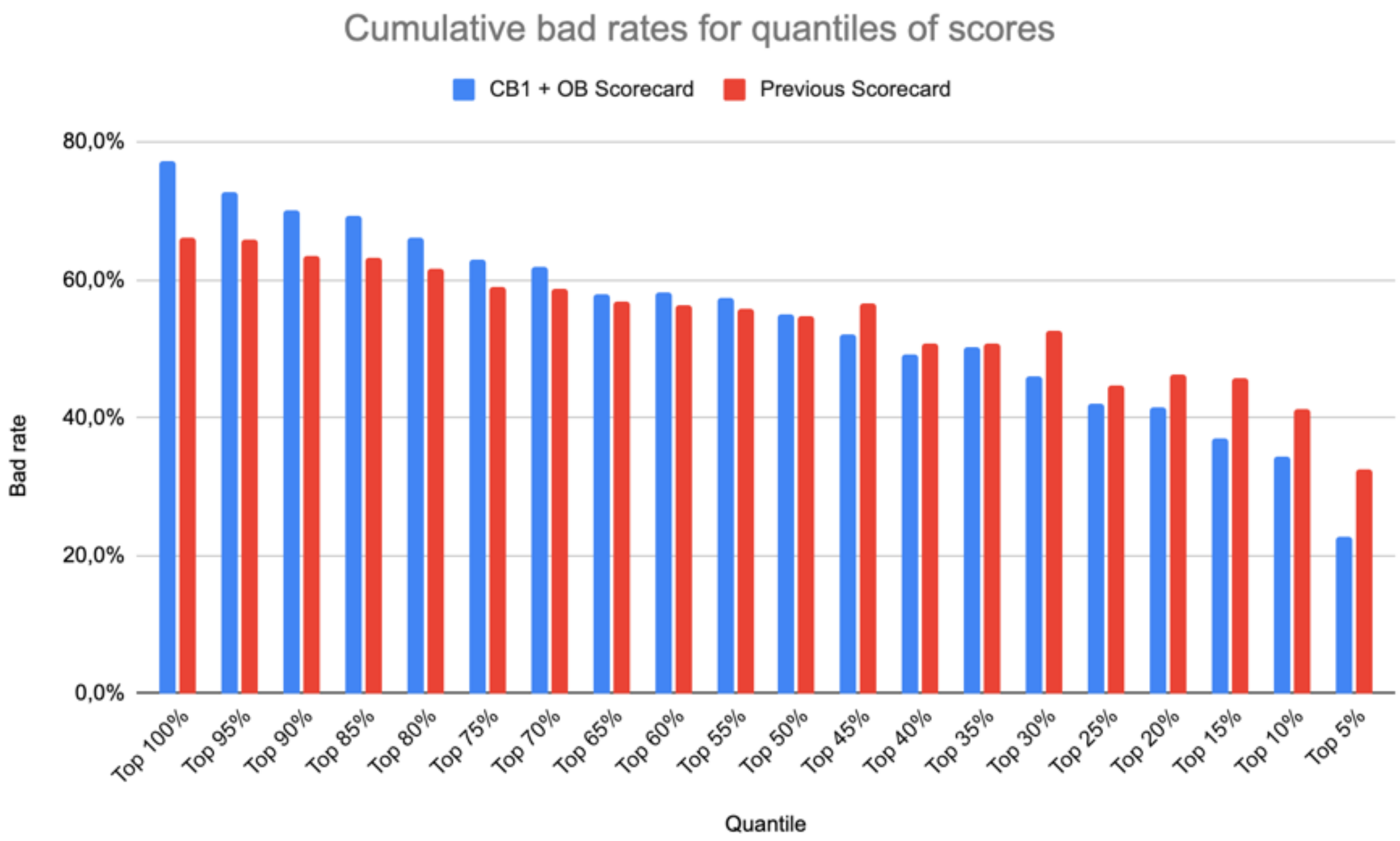
Swap set discussion

At the same bad rate Fair Finance can serve a much larger number of customers.
This reduces sample selection bias for future scorecards.
Accept rate increases from 12.4% to 34.3%.

The introduction of a large number of additional new customers is made safer with the proven good payment history in the outcome retro.
38% of additional new customers have loans in the outcome retro (compared to 21% for all of Fair Finance rejects).

Additional new customers become existing customers with lower bad rate thanks to behavioural data.

The results


Cumulative bad rates for quantiles of scores
CB1 + OB Scorecard
Previous Scorecard
80,0%
60,0%
40,0%
20,0%
0,0%
Bad rate
Top 100%
Top 95%
Top 90%
Top 85%
Top 80%
Top 75%
Top 70%
Top 65%
Top 60%
Top 55%
Top 50%
Top 45%
Top 40%
Top 35%
Top 30%
Top 25%
Top 20%
Top 15%
Top 10%
Top 5%
Quantile

The results

	Precision	Recall	F1 Score
CB + OB Scorecard	77%	76%	73%
CB Scorecard	74%	73%	70%
Current Scorecard	57%	50%	43%

For precision, recall and F1 Score, the accept rate is set as double the accept rate of the current Fair Finance Scorecard.

A weighted average between the two classes ”Good” and ”Bad” is calculated.

Agenda

- Project Objectives
- Population profile
- Results
- **Approach**
- Scorecards

The approach

The development sample consists of customers with Credit Bureau and Open Banking data.

Outcomes are available for:

- Applications accepted by Fair Finance: 3 months in arrears.
- Customers who have loans with other lenders in a short time window around the rejection by Fair Finance (“outcome retro”): “bad” is one month in arrears for short term loans or 3 months in arrears in the first 12 months for other loans.

Reject inference is then applied to assign probabilities of default to other rejected customers.

The same development sample is used to train the different scorecards with different variables.

Reject Inference

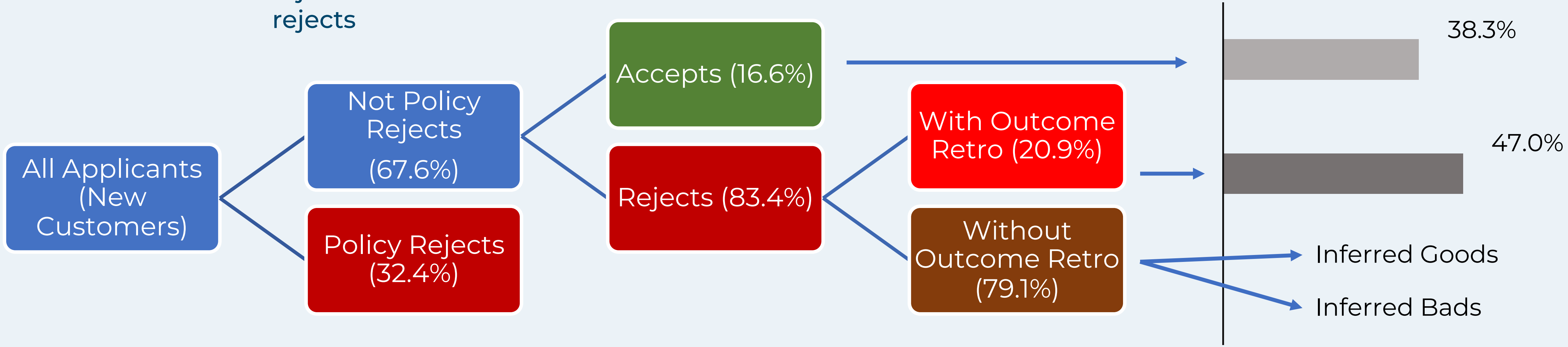
Given the high reject rate of , special attention had to be given to inferring a Good/Bad performance to the rejects

Step 3, carry out “standard” reject inference process to assign probability of default to rejects without outcome retro information

Step 2: perform “outcome retro” on the rejects: identify, at the credit bureau, applicants rejected by Fair Finance who were granted a loan by another lender, within a short time window of the rejection by Fair Finance

Step 1: identify policy rejects vs score rejects

Bad Rate (%)



algoan

The approach: Credit Bureau

Credit Bureau:

- **Standard variables.**
  Examples: number of settled accounts without any delinquency, total limits on active revolving accounts.
- **Bespoke variables**: created from raw accounts and searches data.
  Examples: number of searches in the last X days, number of missed payment by account sector.

Bespoke variables can be more relevant to the Fair Finance population and more stable over time.

The approach: Open Banking

Open Banking consists of bank account transactions over the last 90 days (date, amount, anonymized description, etc.) with some account metadata (account type, overdraft limit, etc.).

2 enrichments to the Open Banking data is performed:

- A balances time series consisting of a daily value of the total account balances is created.
- Transactions are enriched with Algoan's NLP models to add information that is relevant to credit scoring (category and type).

Examples of Open Banking variables:

- Median balance.
- Number of transactions with an amount greater than £X
- Total amount of transaction in category X.

NLP

NLP models classify transactions into 64 categories and 12 types tailored for credit decisioning.

Incomes	ALLOWANCE	WAGE	FREELANCE	RENTAL INCOME
Expenses	LOAN REPAYMENT	INSURANCE	RENT	UTILITIES
Red Flags	BANK INCIDENT FEES	REJECTED PAYMENTS		

The NLP models are built with a language model fine-tuned on data that is labelled in-house.

The same language model architecture powers other Algoan products.

Agenda

- Project Objectives
- Population profile
- Results
- Approach
- Scorecards

Credit Bureau Scorecard

This scorecard is built with 5 variables:

- A bureau score from the Credit Bureau. This is a very predictive variable.
- 4 standard variables based on account data.

Searches have proven to be strongly volatile in the previous scorecard so they were not included.

Only standard credit bureau variables (not bespoke) to make implementation easier by brokers.

Credit Bureau Scorecard – Feature contributions

Variable	Contribution
Bureau score	34.9%
Credit Bureau feature 2	34.3%
Credit Bureau feature 3	8.4%
Credit Bureau feature 4	4.3%
Credit Bureau feature 5	17.9%

Variable contribution is the average absolute number of points of each variable across the development sample.

The score is driven by 2 main variables including the credit bureau score.

Variables with lower contributions provide value by changing the score significantly for a small number of customers.

Credit Bureau + Open Banking Scorecard

Incorporate Open Banking variables and bespoke Credit Bureau variables for better discrimination.

The scorecards uses 8 variables:

- 3 standard credit bureau variables.
- 1 bespoke credit bureau variable.
- 4 Open Banking variables.

Credit Bureau + Open Banking Scorecard

	Variable	Contribution by feature	Contribution by data source
Credit Bureau	Bureau score	16.7%	44.9%
	Credit Bureau feature 2	2.1%	
	Credit Bureau feature 3	13.5%	
	Bespoke credit bureau feature	12.6%	
Open Banking	Open Banking feature 1	15.3%	55.1%
	Open Banking feature 2	11.4%	
	Open Banking feature 3	12.9%	
	Open Banking feature 4	15.4%	

The scorecard is balanced between Credit Bureau and Open Banking features.

The Bureau score is less important in this scorecard as Open Banking adds new information.

Thank you !

www.algoan.com

34, rue Laffitte
75009 Paris - France

contact@algoan.com

algoan