FinTech Lending to Borrowers with No Credit History

Laura Chioda
UC Berkeley

Paul Gertler
UC Berkeley

Sean Higgins
Northwestern

Paolina Medina
University of Houston

The promise of fintech for financial inclusion

- Online FinTech lenders are an increasingly important source of credit for households and small businesses (Berg Fuster Puri 2022, Buchak et al. 2018, Gopal Schnabl 2022)

The promise of fintech for financial inclusion

- Online FinTech lenders are an increasingly important source of credit for households and small businesses (Berg Fuster Puri 2022, Buchak et al. 2018, Gopal Schnabl 2022)

- **The promise of FinTech**: alternative data sources + lower barriers (travel costs, loan processing time)
  - $\Rightarrow$ expand access to credit to the financially excluded

The promise of fintech for financial inclusion

- Online FinTech lenders are an increasingly important source of credit for households and small businesses (Berg Fuster Puri 2022, Buchak et al. 2018, Gopal Schnabl 2022)
- **The promise of FinTech**: alternative data sources + lower barriers (travel costs, loan processing time)
  - $\Rightarrow$ expand access to credit to the financially excluded
- **FinTech in practice**:
  - Fintech lending algorithms still rely on conventional credit bureau data (Johnson et al. 2023)
  - No substantial expansion of credit to those traditionally excluded! (Fuster et al. 2019, Berg Fuster Puri 2022, Cramer et al. 2025)

This project

- Develop credit scores for borrowers *with no credit history*
  - Using digital footprints and transactions data from delivery platform
  - Traditional credit scores based on credit history (FICO) not available by definition

This project

- Develop credit scores for borrowers *with no credit history*
  - Using digital footprints and transactions data from delivery platform
  - Traditional credit scores based on credit history (FICO) not available by definition
- Partnership with RappiCard Mexico, fintech arm of delivery app

This project

- Develop credit scores for borrowers *with no credit history*
  - Using digital footprints and transactions data from delivery platform
  - Traditional credit scores based on credit history (FICO) not available by definition
- Partnership with RappiCard Mexico, fintech arm of delivery app

1. Can machine learning models with alternative data predict creditworthiness for borrowers with no credit history?

This project

- Develop credit scores for borrowers *with no credit history*
  - Using digital footprints and transactions data from delivery platform
  - Traditional credit scores based on credit history (FICO) not available by definition
- Partnership with RappiCard Mexico, fintech arm of delivery app

1. Can machine learning models with alternative data predict creditworthiness for borrowers with no credit history?
2. Would lender do better trying to predict *profitability* instead of default?

Context and Data

Context: fintech in Mexico

- Fostering a dynamic fintech environment has been part of regulators' strategy to promote financial inclusion in Mexico

Context: fintech in Mexico

- Fostering a dynamic fintech environment has been part of regulators' strategy to promote financial inclusion in Mexico
- Fintech Law passed by Mexican Congress in 2018

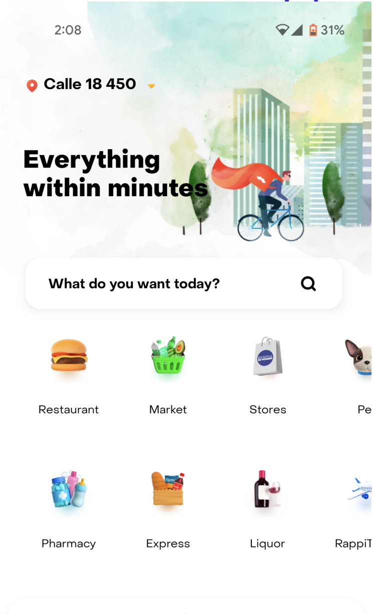
Context: fintech in Mexico

- Fostering a dynamic fintech environment has been part of regulators' strategy to promote financial inclusion in Mexico
- Fintech Law passed by Mexican Congress in 2018
- Largest fintech segment in Mexico is lending (146 fintech lenders)

Context: fintech in Mexico

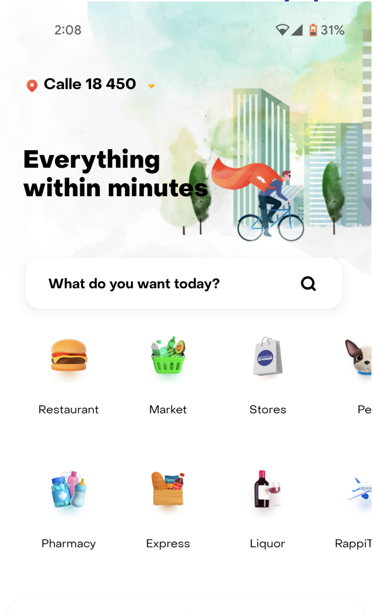
- Fostering a dynamic fintech environment has been part of regulators' strategy to promote financial inclusion in Mexico
- Fintech Law passed by Mexican Congress in 2018
- Largest fintech segment in Mexico is lending (146 fintech lenders)
- One of fintech lenders' main products is credit cards
  - And credit cards are the first formal loan type for 74% of all formal sector borrowers in Mexico (Castellanos et al. 2023)

Context: delivery platforms in Mexico



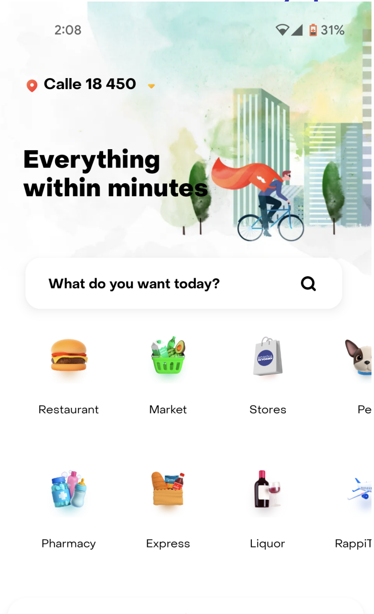
- Rappi is Latin America and Mexico's leading on-demand delivery platform
  - RappiCard has access to transactions data from Rappi

Context: delivery platforms in Mexico



- Rappi is Latin America and Mexico’s leading on-demand delivery platform
  - RappiCard has access to transactions data from Rappi
- Variety of services: purchase of groceries, household items, restaurant food, alcoholic beverages, pharmaceutical products, cash withdrawals, miscellaneous errands

Context: delivery platforms in Mexico



- Rappi is Latin America and Mexico’s leading on-demand delivery platform
  - RappiCard has access to transactions data from Rappi
- Variety of services: purchase of groceries, household items, restaurant food, alcoholic beverages, pharmaceutical products, cash withdrawals, miscellaneous errands
- In 2023Q1, 24% of mobile phone users had at least one delivery app (growing rapidly)

Sample

- RappiCard lends to individuals with and without credit histories
  - With a credit history: combine credit bureau data with transaction-level data from delivery app, use machine learning (ML) algorithms
  - No credit history: at the time of our collaboration, did not use ML algorithm for those with no credit histories; instead used ad hoc rules

Sample

- RappiCard lends to individuals with and without credit histories
  - With a credit history: combine credit bureau data with transaction-level data from delivery app, use machine learning (ML) algorithms
  - No credit history: at the time of our collaboration, did not use ML algorithm for those with no credit histories; instead used ad hoc rules
- Our sample consists of those with no credit history or a credit history so limited that the bureau deems insufficient to generate a credit score
  - No one in our sample had a credit card before
  - $N \approx 155{,}000$ cards originated between Nov 2020 and May 2023, with data on repayments through May 2024

Sample

- RappiCard lends to individuals with and without credit histories
  - With a credit history: combine credit bureau data with transaction-level data from delivery app, use machine learning (ML) algorithms
  - No credit history: at the time of our collaboration, did not use ML algorithm for those with no credit histories; instead used ad hoc rules
- Our sample consists of those with no credit history or a credit history so limited that the bureau deems insufficient to generate a credit score
  - No one in our sample had a credit card before
  - $N \approx$ 155,000 cards originated between Nov 2020 and May 2023, with data on repayments through May 2024
- Measure default as at least 60 days late over the first 12 months with the card Trade-offs

Alternative data sources

- Transaction-level data from the delivery platform
  - Timestamp, quantity of each item purchased, unit price, fees, discounts, tips, order cost, payment method, store name, store geocoordinates

Alternative data sources

- Transaction-level data from the delivery platform
  - Timestamp, quantity of each item purchased, unit price, fees, discounts, tips, order cost, payment method, store name, store geocoordinates
- “No-hit” scores, shared by credit bureau for those with no credit history
  - Based only on geographic indicators from publicly-available data, merged with the location where the individual lives

Alternative data sources

- Transaction-level data from the delivery platform
  - Timestamp, quantity of each item purchased, unit price, fees, discounts, tips, order cost, payment method, store name, store geocoordinates
- "No-hit" scores, shared by credit bureau for those with no credit history
  - Based only on geographic indicators from publicly-available data, merged with the location where the individual lives
- Digital footprint user characteristics (static), such as operating system, acquisition channel, and device model and type (as in Berg et al. 2022)

Alternative data sources

- Transaction-level data from the delivery platform
  - Timestamp, quantity of each item purchased, unit price, fees, discounts, tips, order cost, payment method, store name, store geocoordinates
- "No-hit" scores, shared by credit bureau for those with no credit history
  - Based only on geographic indicators from publicly-available data, merged with the location where the individual lives
- Digital footprint user characteristics (static), such as operating system, acquisition channel, and device model and type (as in Berg et al. 2022)
- Socioeconomics at the census tract level

Summary statistics: subset of features

Panel A: Features	Mean	Std dev	25th perc.	Median	75th perc.
Woman - dummy	0.38	0.49	0.00	0.00	1.00
User age	24.91	8.32	20.00	23.00	26.00
User iOS (Apple) operating system - dummy	0.38	0.49	0.00	0.00	1.00
No-hit score	640.29	14.32	634.00	642.00	649.00
Number of orders on app	24.56	58.70	3.00	9.00	23.00
Proportion orders paid in cash	0.48	0.36	0.14	0.46	0.80
Median amount per order (MXN)	302.70	332.07	178.00	249.75	353.76
Proportion orders at supermarkets	0.05	0.14	0.00	0.00	0.03
Proportion orders at pharmacies	0.03	0.10	0.00	0.00	0.00
Proportion orders at food establishments	0.80	0.27	0.69	0.93	1.00
Marginality (SES) index of census tract	0.97	0.01	0.96	0.97	0.97
Years of schooling among age 15+ in census tract	12.72	1.53	11.76	12.82	13.88
Proportion households own a motor vehicle in census tract	0.66	0.16	0.55	0.67	0.79

Summary statistics: credit card terms and use

Panel B: Credit Card Terms and Use	Mean	Std dev	25th perc.	Median	75th perc.
Credit limit (MXN)	4,573.20	1,964.69	3,888.89	4,666.67	5,000.00
Interest rate	0.78	0.16	0.72	0.87	0.87
Minimum payment (MXN)	339.75	541.03	54.17	93.92	326.72
Statement balance (MXN)	2,940.74	2,023.36	1,433.94	2,846.80	4,188.28
Repayment (MXN)	2,093.93	2,254.45	517.78	1,490.12	2,935.38
Delinquent within 12 months - dummy	0.22	0.41	0.00	0.00	0.00

Results

Can we accurately predict creditworthiness for borrowers with no credit history?

- XGBoost with Bayesian hyperparameter optimization, 3-fold CV More details

Can we accurately predict creditworthiness for borrowers with no credit history?

- XGBoost with Bayesian hyperparameter optimization, 3-fold CV ▸ More details
- 80–20 split for training–test samples

Can we accurately predict creditworthiness for borrowers with no credit history?

- XGBoost with Bayesian hyperparameter optimization, 3-fold CV ▸ More details
- 80–20 split for training–test samples
- All results and measures of predictive accuracy from test sample

Can we accurately predict creditworthiness for borrowers with no credit history?

- XGBoost with Bayesian hyperparameter optimization, 3-fold CV ‣ More details
- 80–20 split for training–test samples
- All results and measures of predictive accuracy from test sample
- Measure model performance using AUC, recall, and F1
  - AUC = area under the receiver operating characteristics curve, which plots true positive rate vs false positive rate for all possible thresholds

Can we accurately predict creditworthiness for borrowers with no credit history?

- XGBoost with Bayesian hyperparameter optimization, 3-fold CV ▸ More details
- 80–20 split for training–test samples
- All results and measures of predictive accuracy from test sample
- Measure model performance using AUC, recall, and F1
  - AUC = area under the receiver operating characteristics curve, which plots true positive rate vs false positive rate for all possible thresholds
- **Yes.** AUC of 0.796
  - Exceeds threshold of Iyer et al. 2016 of 0.7 for data-rich environments
  - In the range of RappiCard's ML models for borrowers *with* credit history
  - Upper end of AUCs with alternative data in middle income countries (with caveat that AUCs will vary depending on default rates)

Comparison to other studies in middle-income countries

Citation	Country	% with Credit Bureau Score	AUC
This paper	**Mexico**	**0%**	**0.796**
Lee et al. 2024	Multiple countries in Asia	50%	0.679
Agarwal et al. 2023	India	63%	0.738 for sample w/ credit history, 0.674 for w/o
Björkegren Grissen 2020	Middle Income South American country	85%	0.711
Rishabh 2024	India	95%	0.68
Frost et al. 2019	Argentina	100%	0.764
Gambacorta 2024	China	100%	0.607
Huang et al. 2023	China	100%	0.841
San Pedro et al. 2015	Latin American country	100%	0.725
DeCnudde et al. 2019	Philippines	Not reported	0.825

Note: See Table 1 of paper for full table. We use AUC as recall and F1 not reported in most studies. Other papers in the US and advanced economies have AUCs of 0.67–0.95 using samples where 94–100% of individuals have a traditional credit bureau score.

Digital footprint and transaction data matter most

Marginal contribution of each data source to AUC

Feature set	AUC	Reduction in AUC
All	0.7959	0
All, but digital footprint user characteristics	0.6724	0.1235
All, but transaction-level data from delivery platform	0.7702	0.0257
All, but no-hit score and limited credit history	0.7899	0.0060
All, but census tract socioeconomic characteristics	0.7949	0.0010

“Thicker” transaction history ⇒ better model performance

AUC, by quintile of number of transactions on Rappi app

Quintile	Number of transactions	AUC
1	2 or fewer	0.7410
2	2–6	0.7739
3	6–12	0.7638
4	12–27	0.7817
5	27 or more	0.8498

Profitability

Should lenders instead be predicting profitability?

- With credit cards, both revenue (interest + fees) and cost (charge-offs) depend on consumer behavior $\rightarrow$ profitability could differ from default:
  - Are there high-debt consumers who do not default and thus generate interest fee revenue? Can we accurately predict who they are?
  - Are there high-spending consumers who generate interchange fee revenue but do not default? Can we accurately predict who they are?
  - These questions have largely not been addressed in the literature due to data limitations

Should lenders instead be predicting profitability?

- With credit cards, both revenue (interest + fees) and cost (charge-offs) depend on consumer behavior $\rightarrow$ profitability could differ from default:
  - Are there high-debt consumers who do not default and thus generate interest fee revenue? Can we accurately predict who they are?
  - Are there high-spending consumers who generate interchange fee revenue but do not default? Can we accurately predict who they are?
  - These questions have largely not been addressed in the literature due to data limitations
- Compare model predicting default (used by most lenders) to a model predicting positive profits

Should lenders instead be predicting profitability?

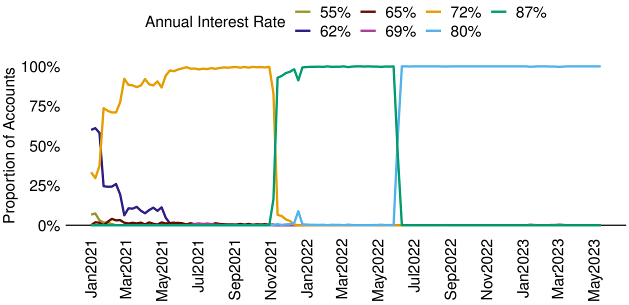
- With credit cards, both revenue (interest + fees) and cost (charge-offs) depend on consumer behavior $\rightarrow$ profitability could differ from default:
  - Are there high-debt consumers who do not default and thus generate interest fee revenue? Can we accurately predict who they are?
  - Are there high-spending consumers who generate interchange fee revenue but do not default? Can we accurately predict who they are?
  - These questions have largely not been addressed in the literature due to data limitations
- Compare model predicting default (used by most lenders) to a model predicting positive profits
- <u>Key result</u>: Model predicting default does *better* (higher profits) than a model predicting profits

Card-by-month level data on revenues and cost

	Mean	Std dev	25th perc.	Median	75th perc.
Interest revenue	63.46	80.67	0.00	30.15	99.90
Charge-offs	52.83	169.18	0.00	0.00	0.00
Interest revenue net of charge-offs	10.67	196.43	0.00	24.66	95.76
Interchange fee revenue	37.06	38.01	10.32	27.08	51.63
Cost of rewards	36.23	39.30	9.73	25.86	49.92
Interchange fee revenue net of cost of rewards	1.08	18.78	-3.66	1.22	8.02
Late payment fee revenue	12.03	30.90	0.00	0.00	0.00
Funding costs	28.01	17.26	15.15	27.84	38.75
Other fee revenue	2.28	7.36	-0.01	0.00	0.00
Other costs	11.27	12.70	2.90	7.60	14.88

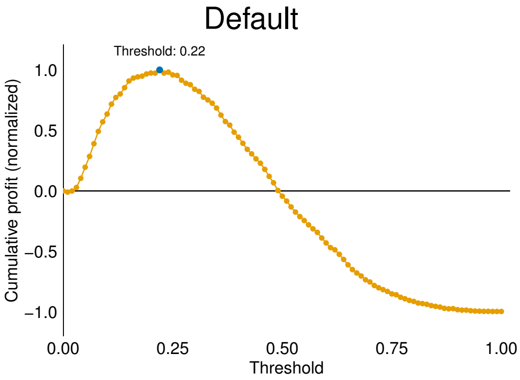
No risk-based pricing for those with no credit history

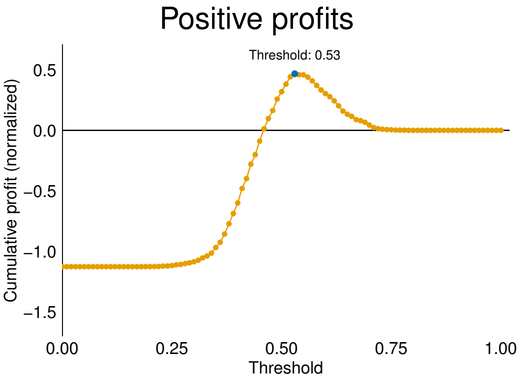
- Not uncommon for FinTech lenders to charge same rate to all borrowers (Berg et al. 2022, Yang 2025)
- To maximize profits this would require experiments to learn elasticities of demand and default with respect to interest rates
  - Where default elasticity comes through adverse selection and causal effect of interest rates on default



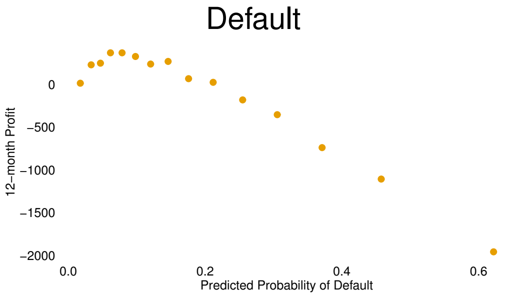
Predicting default more profitable than predicting profits

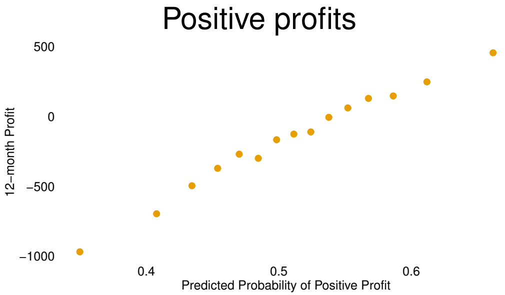
- AUC of profits model 0.598 (compared to 0.796 for defalt)
- Profitability of each model by approval threshold:





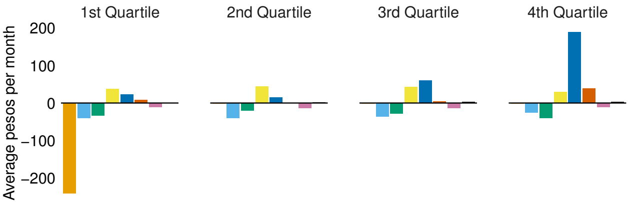
Why is predicting default profit-maximizing?



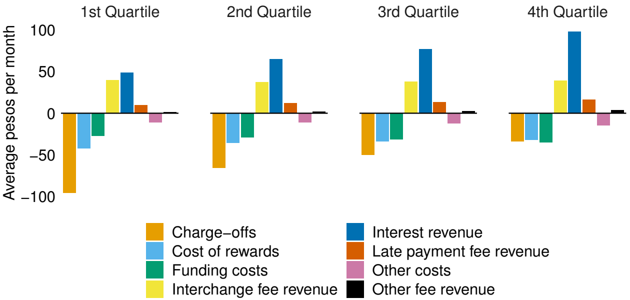


Why is predicting default profit-maximizing?

By quartile of **actual** profitability:



By quartile of **predicted** profitability:



Conclusions and policy implications

1. The promise of FinTech to expand credit access relies on alternative data to predict creditworthiness for people with no credit history
   - Our model with data from a delivery app can accurately predict creditworthiness for individuals without credit history

Conclusions and policy implications

1. The promise of FinTech to expand credit access relies on alternative data to predict creditworthiness for people with no credit history
   - Our model with data from a delivery app can accurately predict creditworthiness for individuals without credit history
2. Why do we observe FinTech lenders typically predicting *default* rather than *profitability*?
   - Negative profits largely driven by default
   - FinTechs use aggressive rewards to compete so high-purchase-volume transactors not profitable
   - Other components of profitability (high debt and interest fees but no default) are hard to predict

   $\Rightarrow$ Suggests FinTech lenders are profit-maximizing by predicting default

Appendix

Gender-segmented models

- Can gender-segmented models provide more equitable outcomes without deteriorating portfolio default rates?

Gender-segmented models

- Can gender-segmented models provide more equitable outcomes without deteriorating portfolio default rates?
- Split the sample for men and women
  - Note the pooled model still had access to gender as a feature (was not gender-blind)!

Gender-segmented models

- Can gender-segmented models provide more equitable outcomes without deteriorating portfolio default rates?
- Split the sample for men and women
  - Note the pooled model still had access to gender as a feature (was not gender-blind)!
- Train separate XGBoost models in each sample
  - Allow all features of the algorithm to change: feature selection, feature importance, hyperparameters, etc.

Gender-segmented models

- Can gender-segmented models provide more equitable outcomes without deteriorating portfolio default rates?
- Split the sample for men and women
  - Note the pooled model still had access to gender as a feature (was not gender-blind)!
- Train separate XGBoost models in each sample
  - Allow all features of the algorithm to change: feature selection, feature importance, hyperparameters, etc.
- Measure performance by calculating out-of-sample AUC on pooled sample, and separately for women (and men)

Gender-segmented models

- Can gender-segmented models provide more equitable outcomes without deteriorating portfolio default rates?
- Split the sample for men and women
  - Note the pooled model still had access to gender as a feature (was not gender-blind)!
- Train separate XGBoost models in each sample
  - Allow all features of the algorithm to change: feature selection, feature importance, hyperparameters, etc.
- Measure performance by calculating out-of-sample AUC on pooled sample, and separately for women (and men)
- Describe portfolio default of both models holding approval threshold fixed

An analogy

Harvard Health Publishing
HARVARD MEDICAL SCHOOL

Clinicians sometimes misread heart attack symptoms in women

TIME SUBSCRIBE

healthline Health Conditions Discover Plan Connect

HEALTH NEWS

Doctors Missed Heart Attack Signs in Women 50% of the Time

Women are more likely than men to report heart attack pain not in the chest.

HEALTH • RESEARCH

Here's Why Doctors Often Miss Heart Attack Symptoms in Women

ESC European Society of Cardiology — The ESC · Congresses & Events · Journals · Guidelines · Education · Research

European Society of Cardiology › The ESC › ESC Press Office › Press releases

ESC Press Office
Press releases

Heart attack diagnosis missed in women more often than in men

12 Mar 2021

Can gender-segmented models produce more equitable outcomes without affecting performance? **Yes**

- AUC of gender segmented model is 0.750 vs 0.752 for pooled model $\rightarrow$ no meaningful loss in accuracy ▸ Full comparison

Can gender-segmented models produce more equitable outcomes without affecting performance? **Yes**

- AUC of gender segmented model is 0.750 vs 0.752 for pooled model → no meaningful loss in accuracy ▸ Full comparison
- Distinct predictions and credit recommendations

Can gender-segmented models produce more equitable outcomes without affecting performance? **Yes**

Women:

		Pooled model	
		Approved	Rejected
Gendered model	Rejected	0.022	0.402
	Approved	0.520	0.057

- Under gender-segmented model (with same approval threshold):
  - 12% of women rejected by pooled model would be accepted

Can gender-segmented models produce more equitable outcomes without affecting performance? **Yes**

Women:

Gendered model \ Pooled model	Approved	Rejected
Rejected	0.022	0.402
Approved	0.520	0.057

- Under gender-segmented model (with same approval threshold):
  - 12% of women rejected by pooled model would be accepted
  - These are 2.6x as many women approved by segmented but rejected by pooled than vice versa

Can gender-segmented models produce more equitable outcomes without affecting performance? **Yes**

Women:		Pooled model: Approved	Pooled model: Rejected
Gendered model	Rejected	0.022	0.402
Gendered model	Approved	0.520	0.057

- Under gender-segmented model (with same approval threshold):
  - 12% of women rejected by pooled model would be accepted
  - These are 2.6x as many women approved by segmented but rejected by pooled than vice versa
  - Portfolio default rate not meaningfully changed, about 10% in both models

Can gender-segmented models produce more equitable outcomes without affecting performance? **Yes**

Men:

Gendered model	Pooled model: Approved	Pooled model: Rejected
Rejected	0.028	0.426
Approved	0.520	0.026

- Under gender-segmented model (with same approval threshold):
  - 12% of women rejected by pooled model would be accepted
  - These are 2.6x as many women approved by segmented but rejected by pooled than vice versa
  - Portfolio default rate not meaningfully changed, about 10% in both models
  - No significant change to credit allocated to men

Why does gender-segmented model have different predictions for women?

1. Estimating the models separately allows *all aspects* of the model to vary, *including hyperparameters*
   - e.g., gender-specific decision trees, learning paths, aggregation into the ensemble learner, and regularization

2. Features from transactions data differentially predict creditworthiness of men and women across quintiles of transaction history

Percentage of top 25 features that are Rappi app transactions features, in separate models by gender and quintile of transaction history

	Quintiles				
Model	1st	2nd	3rd	4th	5th
Pooled model	40%	56%	56%	60%	56%
Women	48%	64%	52%	76%	68%
Men	28%	48%	56%	60%	68%

Marginal contribution of each data source, by gender

Marginal contribution of each data source to AUC, by gender

	Men only		Women only	
Feature set	AUC	Reduction in AUC	AUC	Reduction in AUC
All	0.7549	0	0.7398	0
All, but digital footprint user characteristics	0.7086	0.0463	0.7061	0.0337
All, but transaction-level data from delivery platform	0.7283	0.0266	0.7159	0.0239
All, but no-hit score and limited credit history	0.7376	0.0173	0.7269	0.0129
All, but mobile phone-based proprietary score	0.7466	0.0083	0.7347	0.0051
All, but census tract socioeconomic characteristics	0.7545	0.0004	0.7432	–0.0034

Back

“Thicker” transaction history ⇒ better model performance

AUC, by quintile of number of transactions on Rappi app and gender

		AUC	
Quintile	Number of transactions	Men only	Women only
1	2 or fewer	0.7098	0.6898
2	2–6	0.7443	0.7223
3	6–12	0.7361	0.7379
4	12–27	0.7631	0.7335
5	27 or more	0.7731	0.7691

▸ Back

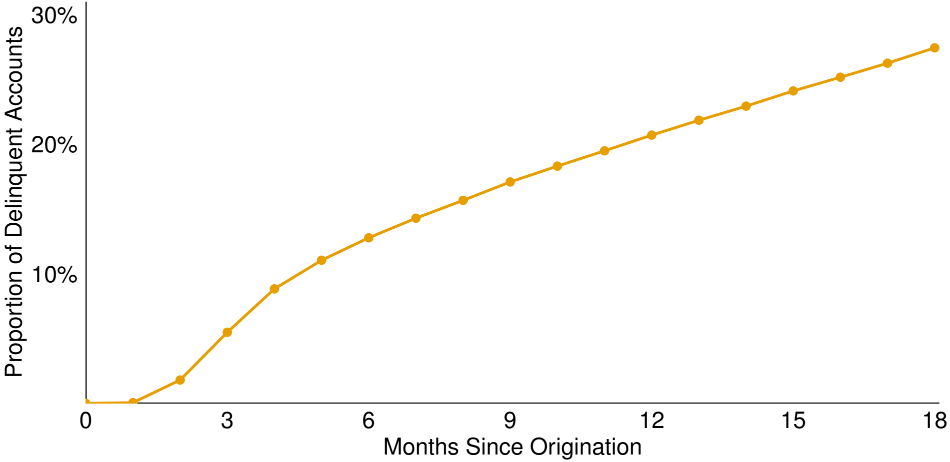
Predictive performance of gender-segmented vs. pooled models

	Full sample			Men only			Women only		
Model	AUC (1)	Recall (2)	F1 (3)	AUC (4)	Recall (5)	F1 (6)	AUC (7)	Recall (8)	F1 (9)
Pooled model	0.7522	0.4781	0.7490	0.7571	0.4894	0.7529	0.7443	0.4597	0.7425
Gender-segmented model	0.7496	0.4770	0.7338	0.7549	0.4875	0.7523	0.7398	0.4588	0.7021

Back

Trade-offs in definition of default

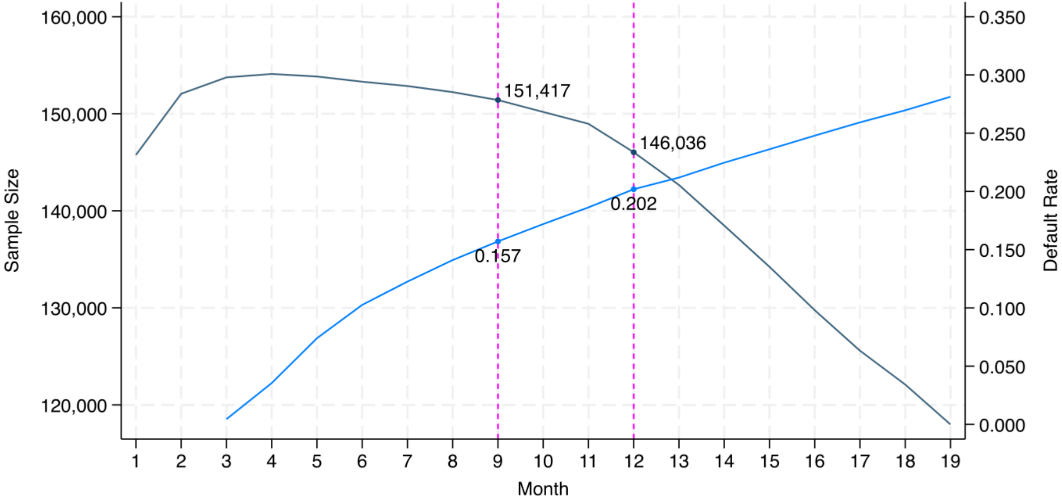
- Trade-off in using full data window for each user vs. fixed period since origination



Back

Trade-offs in definition of default

- Trade-off for fixed period since origination



Back

More details on model and hyperparameter space

Panel A: XGBoost Classifier	
Evaluation metric	logloss
Tuning	hyperopt, max eval 1250
Panel B: Hyperparameter Space	
Tree-specific hyperparameters	
max_depth	hp.quniform('max_depth', 1, 100, 1)
min_child_weight	hp.loguniform('min_child_weight', -2, 3)
subsample	hp.uniform('subsample', 0.5, 1),
colsample_bytree	hp.uniform('colsample_bytree', 0.5, 1),
n_estimator	hp.quniform('n_estimators', 100, 5000,1)
Learning task-specific hyperparameters	
eta, learning rate	hp.loguniform('learning_rate', -9, 0),
gamma	hp.loguniform('gamma', -10, 10),
alpha (L1)	hp.loguniform('reg_alpha', -10, 10),
lambda (L2)	hp.loguniform('reg_lambda', -10, 10),

Back

Threshold determination for default model

- Credit origination decision:
  - Status-quo: Threshold for target default rate.
  - Can be rationalized as profit maximizing under certain assumptions:

$$Profit_i = E[Revenue] - P(default)_i * E[LossGivenDefault]$$

- MR = MC $\rightarrow$ Select marginal borrower to have zero profit and use corresponding P(default) as threshold.

Threshold determination for profit model

- Credit origination decision:
  - Threshold for probability of positive profit.

$$
\begin{aligned}
Profit_i = {} & P(Profit > 0)_i \cdot E[Profit|Profit > 0] \\
& + (1 - P(Profit > 0)_i) \cdot E[Profit|Profit < 0]
\end{aligned}
$$

- MR = MC $\rightarrow$ Select marginal borrower to have zero profit and use corresponding P(Profit $>$ 0) as threshold.

Back