Leveraging Detailed Payment History Data for Better Credit Scoring Methods to Improve Women's Access to Credit

Joshua Blumenstock[*], Laura Chioda[*], Paul Gertler[*], Sean Higgins[†], Paolina Medina[‡] and Enrique Seira[§]

[*]UC Berkeley, [†]Northwestern, [‡]Texas A&M, [§]ITAM

Lab for Inclusive FinTech

Motivation

- Although women default less than men (Karlan and Zinman, 2009; D'Espalier et.al., 2011) gender gaps in access to credit have persisted (Demirguç-Künt and Klapper, 2013, Agier and Szafarz, 2013, Alesina et.al. 2013)

Motivation

- Although women default less than men (Karlan and Zinman, 2009; D'Espalier et.al., 2011) gender gaps in access to credit have persisted (Demirguç-Künt and Klapper, 2013, Agier and Szafarz, 2013, Alesina et.al. 2013)
- We argue that this is partly due to the way in which traditional credit scoring models are calibrated and deployed

Motivation

- Although women default less than men (Karlan and Zinman, 2009; D'Espalier et.al., 2011) gender gaps in access to credit have persisted (Demirguç-Künt and Klapper, 2013, Agier and Szafarz, 2013, Alesina et.al. 2013)
- We argue that this is partly due to the way in which traditional credit scoring models are calibrated and deployed
  - Main input are credit bureau records in which women are underrepresented to begin with $\rightarrow$ cycle of exclusion

Motivation

- Although women default less than men (Karlan and Zinman, 2009; D'Espalier et.al., 2011) gender gaps in access to credit have persisted (Demirguç-Künt and Klapper, 2013, Agier and Szafarz, 2013, Alesina et.al. 2013)
- We argue that this is partly due to the way in which traditional credit scoring models are calibrated and deployed
  - Main input are credit bureau records in which women are underrepresented to begin with $\rightarrow$ cycle of exclusion
  - Gender is excluded (Mester,1997) or included without sufficient flexibility (Johnston and Morduch, 2008)

This project: Gender differentiated credit-scoring algorithms trained with a new type of alternative data

- We partner with RappiCard Mexico, the fintech arm of Rappi, the leading delivery app in Latin America

This project: Gender differentiated credit-scoring algorithms trained with a new type of alternative data

- We partner with RappiCard Mexico, the fintech arm of Rappi, the leading delivery app in Latin America
  - Can these algorithms increase access to credit for underserved women?
  - What is the impact on profitability for lenders?
  - What is their impact on economic outcomes and psychological well-being?

This project: Gender differentiated credit-scoring algorithms trained with a new type of alternative data

- We partner with RappiCard Mexico, the fintech arm of Rappi, the leading delivery app in Latin America
  - Can these algorithms increase access to credit for underserved women?
  - What is the impact on profitability for lenders?
  - What is their impact on economic outcomes and psychological well-being?
- The project has two main stages:

This project: Gender differentiated credit-scoring algorithms trained with a new type of alternative data

- We partner with RappiCard Mexico, the fintech arm of Rappi, the leading delivery app in Latin America
  - Can these algorithms increase access to credit for underserved women?
  - What is the impact on profitability for lenders?
  - What is their impact on economic outcomes and psychological well-being?
- The project has two main stages:
  - Train gender differentiated credit-scoring algorithms with data from Rappi

This project: Gender differentiated credit-scoring algorithms trained with a new type of alternative data

- We partner with RappiCard Mexico, the fintech arm of Rappi, the leading delivery app in Latin America
  - Can these algorithms increase access to credit for underserved women?
  - What is the impact on profitability for lenders?
  - What is their impact on economic outcomes and psychological well-being?
- The project has two main stages:
  - Train gender differentiated credit-scoring algorithms with data from Rappi
  - Randomize access to a credit card among women rejected by standard models but approved by the new model (randomly assign credit scoring models)

Why new types of alternative data? Breaking the cycle of exclusion

Why new types of alternative data? Breaking the cycle of exclusion

- Prior work has shown that flexible machine learning models using data from call detail records (CDR) can successfully predict wealth (Blumenstock et al., 2015) and loan repayment (Björkegren and Grissen, 2015).

Why new types of alternative data? Breaking the cycle of exclusion

- Prior work has shown that flexible machine learning models using data from call detail records (CDR) can successfully predict wealth (Blumenstock et al., 2015) and loan repayment (Björkegren and Grissen, 2015).
- However, recent regulations regarding data privacy have limited access to sensitive data, including call logs

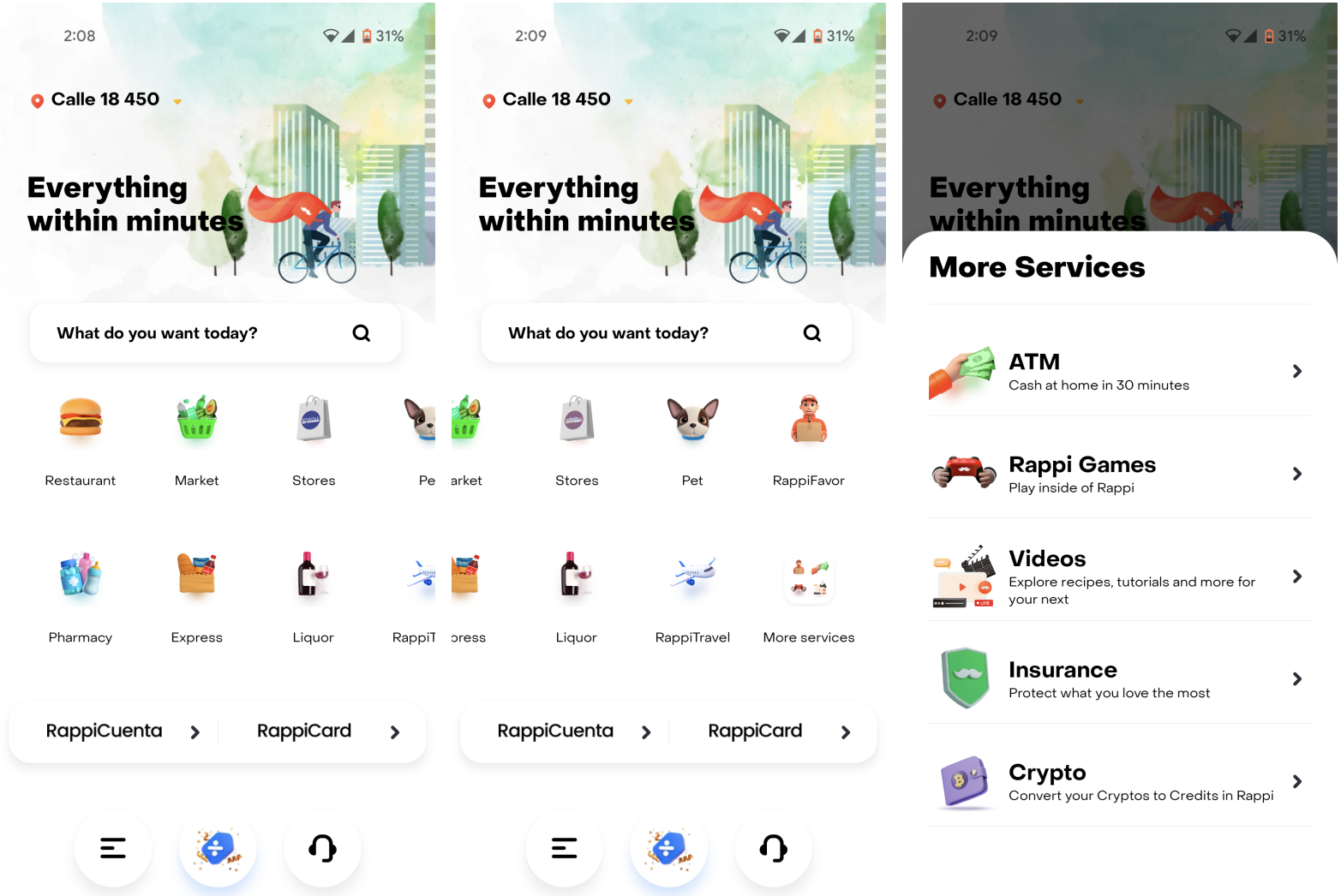
Why new types of alternative data? Breaking the cycle of exclusion

- Prior work has shown that flexible machine learning models using data from call detail records (CDR) can successfully predict wealth (Blumenstock et al., 2015) and loan repayment (Björkegren and Grissen, 2015).
- However, recent regulations regarding data privacy have limited access to sensitive data, including call logs
- Digital footprints have already proven successful in predicting default and improving access to credit among the unbanked (Berg et.al. 2019, Di Maggio et.al. 2022).

Why new types of alternative data? Breaking the cycle of exclusion

- Prior work has shown that flexible machine learning models using data from call detail records (CDR) can successfully predict wealth (Blumenstock et al., 2015) and loan repayment (Björkegren and Grissen, 2015).
- However, recent regulations regarding data privacy have limited access to sensitive data, including call logs
- Digital footprints have already proven successful in predicting default and improving access to credit among the unbanked (Berg et.al. 2019, Di Maggio et.al. 2022).
- This project: new type of digital footprint — Rappi

Rappi data: A new type of digital footprint



Data includes: groceries, pharmacies and medical lab services, cash, air-travel, special deliveries, crypto-payments (coming-up), and other retail shops

Why gender-differentiated algorithms? Intuition

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

Why gender-differentiated algorithms?

- Conditional on observable, differences in default rates could arise if women behave intrinsically differently than men: more reluctant to compete (Nieerle and Vesterlung 2011), more risk averse and less self-confident (Charness and Gneezy 2012 Eckel and Grossman 2008), and less responsive to pay-for-performance incentives (Bandiera et.al. 2021)

Why gender-differentiated algorithms?

- Conditional on observable, differences in default rates could arise if women behave intrinsically differently than men: more reluctant to compete (Nieerle and Vesterlung 2011), more risk averse and less self-confident (Charness and Gneezy 2012 Eckel and Grossman 2008), and less responsive to pay-for-performance incentives (Bandiera et.al. 2021)
- A principled framework for jointly optimizing objectives on two stratified populations, i.e., women and men (Calders and Verwer, 2009; Kleinberg et al., 2015) can capture complex ways in which gender interacts with other characteristics available in the data

Why gender-differentiated algorithms?

- Conditional on observable, differences in default rates could arise if women behave intrinsically differently than men: more reluctant to compete (Nieerle and Vesterlung 2011), more risk averse and less self-confident (Charness and Gneezy 2012 Eckel and Grossman 2008), and less responsive to pay-for-performance incentives (Bandiera et.al. 2021)
- A principled framework for jointly optimizing objectives on two stratified populations, i.e., women and men (Calders and Verwer, 2009; Kleinberg et al., 2015) can capture complex ways in which gender interacts with other characteristics available in the data
- Preliminary analysis with RappiCard Mexico shows promising results

Main outcomes of interest

- Delinquency, default, credit card use and profitability of women approved under the pooled or gender-differentiated model
- Ownership of assets and self-reported control over assets within the household
- Labor supply and access to other sources of credit
- Risk coping and consumption smoothing
- Intra-household bargaining power
- Subjective well-being, self-reported measures of stress and mental-health

Policy Implications and Next Steps

- Access to credit can give woman an equal footing in family decisions, more control over family assets, the ability to cope with risk and to start their own business, and ultimately improve their psychological well-being

Policy Implications and Next Steps

- Access to credit can give woman an equal footing in family decisions, more control over family assets, the ability to cope with risk and to start their own business, and ultimately improve their psychological well-being
- Specially true among sub-populations that have been historically left out for which the marginal loan is all the more valuable

Policy Implications and Next Steps

- Access to credit can give woman an equal footing in family decisions, more control over family assets, the ability to cope with risk and to start their own business, and ultimately improve their psychological well-being
- Specially true among sub-populations that have been historically left out for which the marginal loan is all the more valuable
- Our results can influence origination practices for a wide range of digital credit products in countries that allow the use of gender in credit allocation, and influence policy in those that do not currently allow it ...

Policy Implications and Next Steps

- Access to credit can give woman an equal footing in family decisions, more control over family assets, the ability to cope with risk and to start their own business, and ultimately improve their psychological well-being
- Specially true among sub-populations that have been historically left out for which the marginal loan is all the more valuable
- Our results can influence origination practices for a wide range of digital credit products in countries that allow the use of gender in credit allocation, and influence policy in those that do not currently allow it ...
- And inform the debate about gender-blind algorithms and how to prevent discrimination

Policy Implications and Next Steps

- Access to credit can give woman an equal footing in family decisions, more control over family assets, the ability to cope with risk and to start their own business, and ultimately improve their psychological well-being
- Specially true among sub-populations that have been historically left out for which the marginal loan is all the more valuable
- Our results can influence origination practices for a wide range of digital credit products in countries that allow the use of gender in credit allocation, and influence policy in those that do not currently allow it ...
- And inform the debate about gender-blind algorithms and how to prevent discrimination
- Next steps: ML model + Pilot + RCT