

# Discussion of:

## On the Rise of FinTechs - Credit Scoring using Digital Footprints

Daniel Paravisini

London School of Economics and CEPR

September 5, 2019

## Interesting questions

- Does digital footprint contain information useful to evaluate consumer credit risk?
- Is the information different/better than that in the credit footprint?

# Interesting questions

- Does digital footprint contain information useful to evaluate consumer credit risk?
- Is the information different/better than that in the credit footprint?
- The paper answers a narrower question
  - ▶ Does digital footprint contain different/better information than the credit score?

# Credit footprint

- 50-100 different variables
  - ▶ Amount borrowed by type, by institution, used and unused credit lines, past repayment behavior (of loans and utilities), credit searches, etc.
- Credit Bureau estimates a model to predict an outcome
  - ▶ Outcome is related to credit risk
  - ▶ What model and what the model predicts are proprietary
- Score: unidimensional summary statistic

# Easy benchmark to beat

- Paper compares predictive power of the score...
  - ▶ Ignores information in credit footprint that may be useful to predict on-line purchase defaults
  - ▶ Predicts something other than default in on-line purchases
- ...with predictive power of a model
  - ▶ Estimated using all the information in the digital footprint
  - ▶ Built to predict the same outcome against which performance will be measured

# Scores alone are bad default forecasters

- Lenders rarely rely on bureau score alone to lend
- Paper documents poor predictive performance

## AUC and changes in the Area Under the Curve using other variables in addition to the credit bureau score

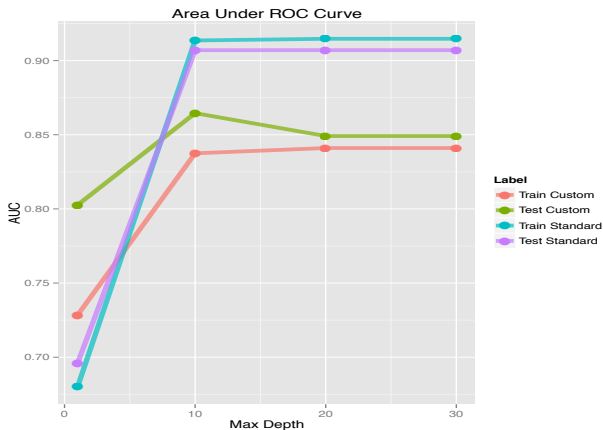
		AUC Change	Combined AUC
This study	Digital footprint versus credit bureau score only	+ 5.3PP	73.6%
Berg, Puri, and Rocholl (2017) <sup>a</sup>	Bank internal rating (which includes credit bureau score) versus credit bureau score only	+8.8PP	75.4%
Puri, Rocholl, and Steffen (2017) <sup>a</sup>	Bank internal rating (which includes credit bureau score) versus credit bureau score only	+11.9PP	78.4%
Iyer, Khwaja, Luttmer, and Shue (2016)	Interest rates versus credit bureau score only	+5.7PP	68.2%
Iyer, Khwaja, Luttmer, and Shue (2016)	All available financial and coded information (including credit bureau score) versus credit bureau score only	+8.9PP	71.4%
Lending Club (own analysis)	Lending Club loan grade (which includes credit bureau score) versus credit bureau score only	+11.9PP	71.7%

<sup>a</sup> These results are not in the original papers but were provided to us by the authors using exactly the same data set from the paper.

<sup>b</sup> Results are very similar for 60-month loans.

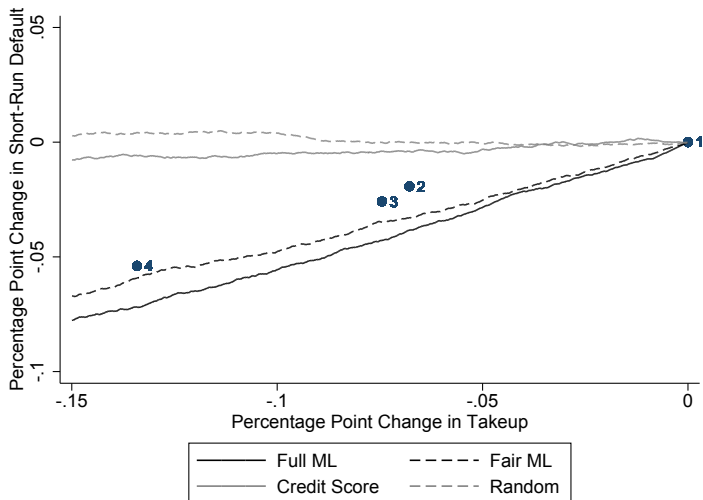
## AUC when using all traditional bureau data

- Baker (2015): U.S., 50,000 bank customers, 5,000,000 transactions, 13 months, ML (random forests)



## How bad can scores be?

Dobbie et al. (2019) compare prediction algorithms on marginal loans (U.K. lender,  $n > 20K$ , similar loan sizes)





# Technological change

- More data, better models
- Important equilibrium implications
- Two recent working papers:
  - ▶ Fuster et al. (2019): “Predictably Unequal? The Effects of Machine Learning on Credit Markets”
  - ▶ Liberman et al. (2019): “The Equilibrium Effects of Information Deletion: Evidence from Consumer Credit Markets”

## Fuster et al. (2019)

- U.S. mortgage data (HMDA, McDash),  $n > 10$  million, 2009-2014

Model	ROC AUC Score	
	No Race	Race
Logit	0.8477	0.8484
Logit Non-Linear	0.8518	0.8524
Random Forest Isotonic	0.8577	0.8577

- Estimate a structural model of the credit market:
  - ▶ Technology leads to larger fraction of Whites and Asians to be offered lower rates (opposite for Black and Hispanic)

## Liberman et al. (2019)

- In 2012, Chilean credit bureaus were forced to stop reporting defaults for 2.8 million individuals (21% of the adult population)
- Increases in the predicted cost of lending are largest for lower-income non-homeowners who resemble defaulters but have good credit records
- On net deletion caused borrowing to fall by 3.5
- Welfare effects of deletion are negative under a variety of assumptions about lenders pricing strategies

# Conclusion

- Fundamental message of the paper is sound
  - ▶ More information has important equilibrium implications in credit markets
- But empirical exercise makes a narrower point
  - ▶ Digital footprint contains information that credit scores do not