Discussion of "On the Rise of FinTechs-Credit Scoring using Digital Footprints" by Berg, Burg, Gombović, and Puri

Xin Wang

Nanyang Technological University

ABFER 7th Annual Conference

May 30, 2019

Summary

- The paper analyzes the information content of the digital footprint for predicting consumer default
  - Digital footprints match the information content of credit bureau scores
  - Complements rather than substitutes for credit bureau information
  - Broad implications for financial intermediaries and financial inclusion

Summary

- The paper analyzes the information content of the digital footprint for predicting consumer default
  - Digital footprints match the information content of credit bureau scores
  - Complements rather than substitutes for credit bureau information
  - Broad implications for financial intermediaries and financial inclusion
- Overview
  - Very interesting and well-written paper
  - Convincing evidence
  - Minor issues on sample selection and the implications for other long-term loan markets

What are Digital Footprints?

- A digital footprint is a trail of data you create while using the Internet. It includes the websites you visit, emails you send, and information you submit to online services.

- We are living in a digital world
  - Mobile payments (Alipay, Google Pay, etc.)
  - E-commence (Amazon, Taobao, etc.)
  - Social networks (Facebook, Twitter, WeChat, etc.)
  - Sharing economy (Uber, Airbnb, Filecoin, etc)
  - Peer-to-peer lending and insurance

Use Cases of Digital Footprints

- **Alternative credit Scoring**
  - For the unbanked
  - Enables instant Point of Sale (PoS) financing
  - Peer-to-peer lending platform
  - CredoLab (Singapore)—developed a credit scoring mobile app, CredoApp, which evaluates over 50,000 data points from a client's phone and produces a credit score in under two minutes

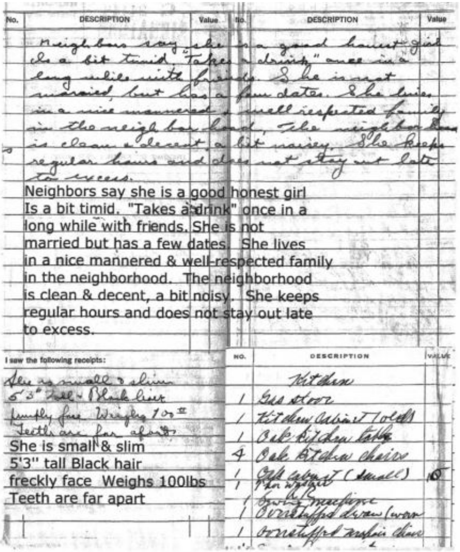
Use Cases of Digital Footprints

- **Alternative credit Scoring**
  - For the unbanked
  - Enables instant Point of Sale (PoS) financing
  - Peer-to-peer lending platform
  - CredoLab (Singapore)—developed a credit scoring mobile app, CredoApp, which evaluates over 50,000 data points from a client's phone and produces a credit score in under two minutes
- **Insurance pricing**
  - i.e. Are you often on your phone between 12 midnight - 6:00 a.m.?
  - Could increase your car and health insurance premium

Use Cases of Digital Footprints

- **Alternative credit Scoring**
  - For the unbanked
  - Enables instant Point of Sale (PoS) financing
  - Peer-to-peer lending platform
  - CredoLab (Singapore)—developed a credit scoring mobile app, CredoApp, which evaluates over 50,000 data points from a client's phone and produces a credit score in under two minutes
- **Insurance pricing**
  - i.e. Are you often on your phone between 12 midnight - 6:00 a.m.?
  - Could increase your car and health insurance premium
- **Dynamic pricing**
  - Anecdotal evidence: Orbitz shows higher prices to Mac users
  - Ride-hailing surcharge

Credit Evaluations from the 1930s



No.	DESCRIPTION	Value	No.	DESCRIPTION	Value

Neighbors say she is a good honest girl
Is a bit timid. "Takes a drink" once in a
long while with friends. She is not
married but has a few dates. She lives
in a nice mannered & well-respected family
in the neighborhood. The neighborhood
is clean & decent, a bit noisy. She keeps
regular hours and does not stay out late
to excess.

I saw the following receipts:

She is small & slim
5'3" tall Black hair
freckly face Weighs 100lbs
Teeth are far apart

NO.	DESCRIPTION	VALUE
	Kitchen	
1	Gas stove	
1	Kitchen cabinet (old)	
1	Oak kitchen table	
4	Oak kitchen chairs	
1	Oak cabinet (small)	10

"teeth are far apart"
"takes a drink" once a while
Not married, but has a few dates
Neighbors say she is a good-hearted girl

source: Eric Falkenstein, Finding Alpha: The Search for Alpha When Risk and Return Break Down

The Paper in a Nutshell

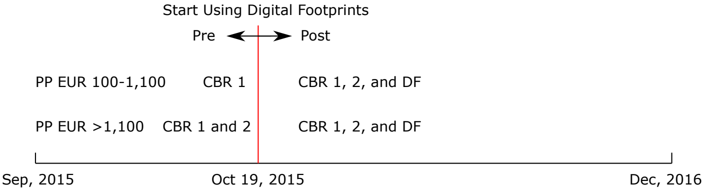
- Analyze the default prediction using approximately 250,000 purchases from an E-Commerce company selling furniture in Germany
- Customers with good creditworthiness have deferred payment option—pay after shipment
- The company started to use ten digital footprints (DF) variables for predicting default in Oct. 2015
- Main findings:
  - After using DF, the company's default rates decreased
  - DF Complements for credit bureau information
  - DF matters for other loan products such as consumer or mortgage loans

Which DF variables matter?

Variable	Standalone AUC	Marginal AUC
Computer & Operating system	59.03%	+1.71PP***
Email Host	59.78%	+2.44PP***
Email Host: paid versus non-paid dummy	53.80%	+0.98PP***
Email Host: Variation within non-paid email hosts	57.82%	+1.79PP***
Channel	54.95%	+0.70PP***
Check-Out Time	53.56%	+0.63PP***
Do not track setting	50.40%	+0.00PP
Name In Email	54.61%	+0.30PP**
Number In Email	54.15%	+0.19PP**
Is Lower Case	54.91%	+1.15PP***
Email Error	53.08%	+1.79PP***

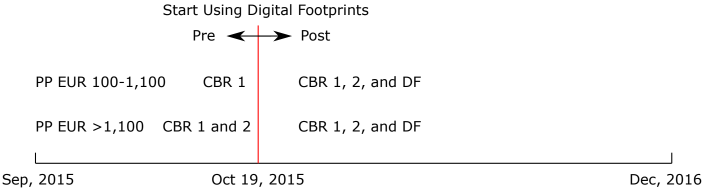
- Only “do not track” not significant
- Non-income proxies more important than income proxies

Evidence of Decreased Default Rates



Purchases Amount	Sample Period	Creditworthiness Judging Source	Invoice Offered Rate	Default Rates
EUR 100-1,100	Pre	CBR 1	96.65%	2.54%
	Post	CBR 1, 2 and DF	90.05%	1.19%
$>$ EUR 1,100	Pre	CBR 1, 2	39.00%	3.62%
	Post	CBR 1, 2 and DF	40.11%	2.33%

Evidence of Decreased Default Rates

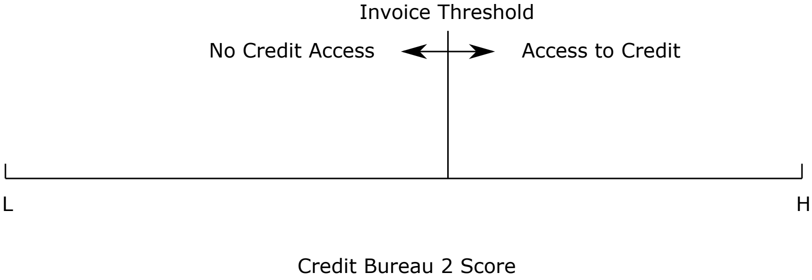


Purchases Amount	Sample Period	Creditworthiness Judging Source	Invoice Offered Rate	Default Rates
EUR 100-1,100	Pre	CBR 1	96.65%	2.54%
	Post	CBR 1, 2 and DF	90.05%	1.19%
$>$ EUR 1,100	Pre	CBR 1, 2	39.00%	3.62%
	Post	CBR 1, 2 and DF	40.11%	2.33%

Results are solely driving by digital footprints.

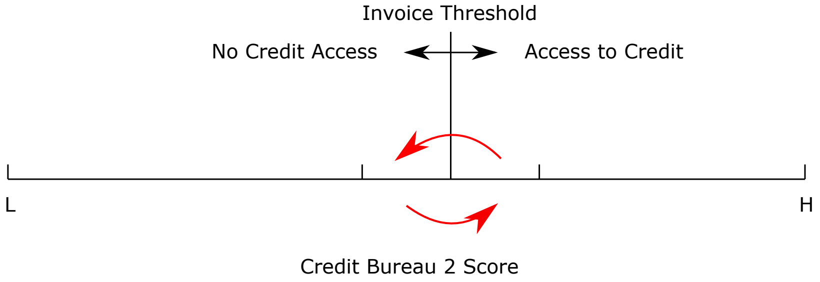
How do Digital Footprints Improve Default Prediction?

For the case of purchases amount $>$ EUR 1,100 and before Oct 19, 2015:



Using Digital Footprints (DF)

- Some above the threshold but with poor DF score get rejected
- Some below the threshold but with good DF score get the credit



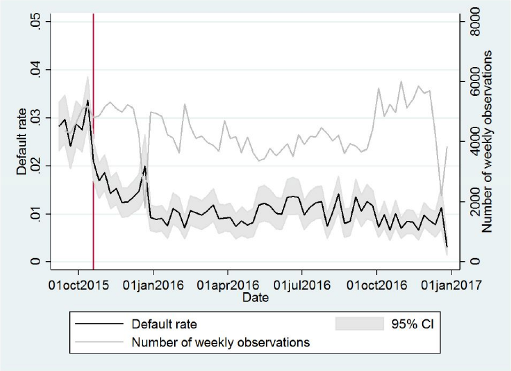
My Comments

The paper is forthcoming at the ***Review of Financial Studies***!

Only minor comments on:

- The correlation between invoice offered and default rates
- Sample selection
- Representativeness of other loans

Comment 1: Correlation between the Number of Invoice Offered and Default Rates



- Expect a positive correlation
- Positive pre-DF period, and seems to have a clear negative correlation post DF period. Why? Show more results in the pre-DF period?

Comment 2: Minor Issues on Sample Selection

- The main sample includes all purchases with access to credit after the company using digital footprints from Oct 19, 2015 to Dec 31, 2016
- Estimate default probability in a linear logistic regression
- How did the company use digital footprints to judge a customer's creditworthiness? Non-linear functional form?
- Can predictions be different for those customers rejected for credit access?

Comment 3: Are the Results Representative of long-term loan?

My prior is not. Because:

- It's a one-time short term loan with an average amount of USD 350
- Hard to think customers default because of financially constraint
- Moreover, default probability is negatively correlated with the loan amount (footnote 23)

Comment 3: Are the Results Representative of long-term loan?

My prior is not. Because:

- It's a one-time short term loan with an average amount of USD 350
- Hard to think customers default because of financially constraint
- Moreover, default probability is negatively correlated with the loan amount (footnote 23)

But the authors show that digital footprints today can forecast future changes in the credit bureau score.

Comment 3: Are the Results Representative of long-term loan?

My prior is not. Because:

- It's a one-time short term loan with an average amount of USD 350
- Hard to think customers default because of financially constraint
- Moreover, default probability is negatively correlated with the loan amount (footnote 23)

But the authors show that digital footprints today can forecast future changes in the credit bureau score. I think there is still room for future research:

- How do digital footprints proxy for soft information in lending?
- For different loans, borrowers have various reasons and cost of default
- Still interesting to see how digital footprints work in a long-term loan like mortgage loans

Conclusions

- The first paper on analyzing the information content of digital footprints
- Interesting and intuitive results
- Providing evidence that digital footprints have important implications for the unbanked

Possible future research:

- Digital footprints vs. soft information
- The role of screening vs. monitoring of digital footprints
- The impacts of digital footprints on insurance and dynamic pricing