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
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
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

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
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

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.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teeth are far apart</td>
<td></td>
</tr>
<tr>
<td>Freckly face</td>
<td>Weighs 100 lbs</td>
</tr>
<tr>
<td>5'3&quot; tall Black hair</td>
<td></td>
</tr>
</tbody>
</table>

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

"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
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?

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

- Only “do not track” not significant
- Non-income proxies more important than income proxies
Evidence of Decreased Default Rates

Start Using Digital Footprints

Pre → Post

PP EUR 100-1,100 CBR 1 CBR 1, 2, and DF
PP EUR >1,100 CBR 1 and 2 CBR 1, 2, and DF

<table>
<thead>
<tr>
<th>Purchases Amount</th>
<th>Sample Period</th>
<th>Creditworthiness Judging Source</th>
<th>Invoice Offered Rate</th>
<th>Default Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR 100-1,100</td>
<td>Pre</td>
<td>CBR 1</td>
<td>96.65%</td>
<td>2.54%</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>CBR 1, 2 and DF</td>
<td>90.05%</td>
<td>1.19%</td>
</tr>
<tr>
<td>&gt; EUR 1,100</td>
<td>Pre</td>
<td>CBR 1, 2</td>
<td>39.00%</td>
<td>3.62%</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>CBR 1, 2 and DF</td>
<td>40.11%</td>
<td>2.33%</td>
</tr>
</tbody>
</table>
Evidence of Decreased Default Rates

Start Using Digital Footprints

Pre → Post

<table>
<thead>
<tr>
<th>Purchases Amount</th>
<th>Sample Period</th>
<th>Creditworthiness Judging Source</th>
<th>Invoice Offered Rate</th>
<th>Default Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR 100-1,100</td>
<td>Pre</td>
<td>CBR 1</td>
<td>96.65%</td>
<td>2.54%</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>CBR 1, 2 and DF</td>
<td>90.05%</td>
<td>1.19%</td>
</tr>
<tr>
<td>&gt; EUR 1,100</td>
<td>Pre</td>
<td>CBR 1, 2</td>
<td>39.00%</td>
<td>3.62%</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>CBR 1, 2 and DF</td>
<td>40.11%</td>
<td>2.33%</td>
</tr>
</tbody>
</table>

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:

Credit Bureau 2 Score

Invoice Threshold

No Credit Access  Access to Credit

Credit Bureau 2 Score

Xin Wang
On the Rise of FinTechs (Discussion)
May 30, 2019 9 / 15
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