Predictably Unequal
The Effects of Machine Learning on Credit Markets
Authors: Andreas Fuster, Paul Goldsmith-Pinkham, Tarun Ramadorai, Ansgar Walther

Discussant: Tobias Berg, Frankfurt School of Finance and Management

Frankfurt, September 2020
Summary

Machine learning
  = better prediction of creditworthiness
  = different prediction of creditworthiness
→ Who benefits/who loses?

Using U.S. mortgage data from 2009-2013
  • Black and Hispanic borrowers less likely to gain
  • Attributable to flexibility, not triangulation
Dispersion PD ↑, Black-PD ↑

• General: Dispersion in PD increases
• Black borrowers: PD: 1.6% → 1.9% (mean PD increases)
Equilibrium model: Black acceptance rate ↑, black rates ↑

Results are not black and white: Rejected blacks benefit from larger variance of prediction

<table>
<thead>
<tr>
<th></th>
<th>Accept (%)</th>
<th>Mean SATO (%)</th>
<th>SD SATO (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)  (2)</td>
<td>(3)  (4)</td>
<td>(5)  (6)</td>
</tr>
<tr>
<td></td>
<td>NL  RF</td>
<td>NL  RF</td>
<td>NL  RF</td>
</tr>
<tr>
<td>Asian</td>
<td>92.4  93.3</td>
<td>-0.108 -0.123</td>
<td>0.274 0.322</td>
</tr>
<tr>
<td>White Non-Hispanic</td>
<td>90.3  91.1</td>
<td>-0.083 -0.090</td>
<td>0.296 0.356</td>
</tr>
<tr>
<td>White Hispanic</td>
<td>85.6  86.4</td>
<td>-0.031 0.008</td>
<td>0.333 0.414</td>
</tr>
<tr>
<td>Black</td>
<td>77.7  79.3</td>
<td>0.022 0.060</td>
<td>1.61</td>
</tr>
<tr>
<td>Other</td>
<td>88.9  89.5</td>
<td>-0.083 0.088</td>
<td>0.296 0.360</td>
</tr>
<tr>
<td>Population</td>
<td>89.8  90.7</td>
<td>-0.081 0.086</td>
<td>3.60</td>
</tr>
<tr>
<td>Cross-group SD</td>
<td>2.165 2.098</td>
<td>0.020 0.029</td>
<td></td>
</tr>
</tbody>
</table>
What I like about the paper

• FinTech-Lending
  – More/better data
  – Better methodology

• Host of papers on effects of more/better data
• Little known about effects of better methodology

• New topic, important question, large market
• Fundamental insight beyond specific setting:
  – Conceptual framework
  – Illustration of equilibrium effects
#1: Provocative Interpretation

• “Black and Hispanic borrowers are disproportionately less likely to gain from the introduction of machine learning.”

• “The majority of the predictive accuracy gains [...] can be attributed to the increased flexibility” (and not to triangulation)

• Provocative interpretation: Current use of coarse logistic models discriminates against Whites and subsidizes Black and Hispanic Borrowers
#2: Should Homer Simpson receive a loan?

- "No Loan Again, Naturally“ (Simpsons 2009)
  - Homer throws a party using home equity loan
  - Defaults on his loan, loses house
  - Gets rescued by Ned Flanders

- Deeper question: Worse rating = loser of new rating method?
  - PD=1% for everyone
  - PD=0.9%/1.1% $\rightarrow$ 1.1%-type = losers of new rating method
  - PD=0%/100% $\rightarrow$ Are the 100%-PDs really losers? Or saved by new rating method from private bankruptcy?

- Rating method with maximum variance (PD=0/1)
  - Seems hard to argue that there are losers
  - Seems hard to argue that this is bad for risk averse applicants
#3: Better model = more lending?

Rough signal (e.g. FICO)

- Receive cheap loan
- Receive expensive loan
- Rejected

Better signal (e.g. ML)

- Willing to lend at pooling price: Better information → less lending
- Not willing to lend at pooling price: Better information → more lending
#3: Better model = more lending?

- Pooling price: willing to lend or not?
  - If yes, then better model leads to less lending
  - If no, then better model leads to more lending
  - See Proposition 4 in Pagano and Japelli (1993)
  - See Section 3.2. in Berg et al. (2020), On The Rise in FinTechs. We find more lending for applicants with scarce data after introduction of digital footprint

- Equilibrium price and quantities highly depend on whether pooling price leads to unraveling

- Conceptual discussion in paper is great, with one exception:
  Should discuss implications of Pagano and Japelli (1993), Proposition 4
  - Data set only includes accepted loans = loans where pooling price does not lead to unraveling
  - Statements on quantity and price could be different if you look at full set of applications
  - Currently rejected borrowers should benefit most from better prediction
Summary

• Important topic, important contribution

• I personally very much enjoyed the conceptual discussion

• It will surely become a very impactful paper

• Suggestion to the audience: Read it!