Discussion by
Stephen H. Shore (GSU)

November 1, 2019

Timothy Dore and Traci Mach (Fed Board)

Marketplace Lending and Consumer Credit Outcomes: Evidence from Prosper

J. Christina Wang and Charles B. Perkins (Boston Fed)

How Magic a Bullet is Machine Learning for Credit Analysis?

Georgia State University | J. MACK ROBINSON COLLEGE OF BUSINESS

Robinson

Framing: FinTech

New technology is changing the products, process, and pricing in financial markets

- Products (Dore and Mach)
- Pricing (Wang and Perkins)

Does this help us

- Move wealth over time and between states in a more targeted way?
- Reduce the cost of moving this risk?
- Reduce informational asymmetries?

Are there distributional consequences?

Contributions: Dore and Mach

P2P Platform expands credit access (or lowers rates) for borrowers

- What does this do to their financial lives?
- Compare platform applicants who do and don’t borrow on the platform
  - How are treatment and control groups different?
  - Does this represent causal effect of borrowing?
- Findings for those who use P2P lending
  - Improved FICO
  - less CC debt (substitution)
  - more total debt (imperfect substitution)
  - lower delinquency

Finally, a paper about me…

From Shore's prosper.com 2006 loan application:

Event type	details	date
Listing Expired	Your listing, "Seeing from borrower's perspective" (Listing #37842) for $1,000.00 at 1.00% expired without sufficient funding.	Sep-15-2006 1:18 PM
Listing Added	Your listing, "Seeing from borrower's perspective" (Listing #37842) for $1,000.00 at 1.00% has been created and is active on the marketplace.	Sep-08-2006 1:17 PM

Identification: Dore and Mach

Comparing Treatment and Control

- Compare those who agreed to borrow to those who applied but then declined.
- Would these groups have behaved differently absent P2P lending?
- Why didn’t my loan application succeed? (Given later rates, I would have declined the loan. How might my loan trajectory absent P2P lending differ from those who chose to take a loan?)
- Treatment group more serious about borrowing, would have had upward borrowing trajectories with or without access to P2P lending.
- Control group may have had lower cost outside options.

What would ideal causal design reveal?

- General equilibrium: P2P cream skimming
- Effect of getting P2P loan ≠ effect of P2P rollout

Implications: Dore and Mach

What could P2P lending be doing?

1. **Loaning money to people who couldn't otherwise get it (at price wtp)**
   a) This is good if you think their revealed preference to borrow means borrowing helps them

2. **Loaning money to people who would borrow otherwise but at lower rates**
   a) If due to improved risk classification, then this is just cream-skimming. P2P lenders/borrowers benefit at the expense of legacy lenders and borrowers.
   b) If due to lower costs of platform, this makes P2P borrowers better off (w minimal harm)

How can we tell which is happening?

- Both (1) and (2). Observed imperfect substitution from credit card borrowing suggests some borrowing is new and other borrowing displacing traditional borrowing.
- Hard to know if (2.a) or (2.b) but matters a lot for welfare, but I'm skeptical it is (2.a).

Distributional consequences?

Contributions: Wang and Perkins

Compare performance of machine learning v. logit for default prediction

- Ex-ante unknown covariates (market conditions) to estimate systematic risk properties of loans
- In-sample vs out-of-sample
- As number of observations grows

- Sensible tools, well implemented!

Findings: Wang and Perkins

Compare performance of machine learning v. logit for default prediction

- Find improved performance of ML methods in sample
- Benefits of ML limited in
  - larger samples
  - out-of-sample (risk of over-fitting?)

Do findings reflect specific ML implementation and not ML generally?

- Firms may use different ML methods more sensitive to over-fitting?
- This paper uses ML to tackle non-linearity/interaction of existing covariaties. Might not apply to using ML to increase and search among increased number of sparse covariates (e.g., text of tweets, consumer purchases)

Future Covariates and Out-of-Sample Results

When does prediction accuracy matter most?

- Results based on future covariates (where ML doesn’t work particularly well)
- Most important for systemic risk
- If you had the same customers and same average risk prediction (because competitors’ rates adjust with yours when everyone uses ML, and logit is right on average), it wouldn’t matter if you could spot good and bad risks.

Why do we care about use of ML?

ML as a (somewhat) better technology to predict risk. Competitive advantage.

- ML allows better cream-skimming of best risks by users of ML, at the expense of legacy logit firms.

Is added information from ML otherwise private information of customers?

- If so, incorporating ML improves risk classification and reduces adverse selection. Introducing ML increases efficiency.
- Look for evidence that ML information known to customers? Positive correlation test? Compare tendency to borrow by those with p(ML)>p(logit) vs those with p(ML)<p(logit)

Absent private info, what happens in equilibrium when all firms use ML?

- **ML as arms race.** Same customers (competitors' rates adjust too) and same average risk. (Prices slightly higher on average to pay the smart and highly compensated ML data scientists.)
- Once people know ML in widespread use, behavior changes.

Winners and losers

- ML helps white borrowers at the expense of minorities, relative to logit benchmark (see Fuster, Goldsmith-Pinkham, Ramadorai, Walther, R&R, JF)

Conclusion

Thanks for two great papers.

- **Both tell us more about how technology may be changing the efficiency and distributional consequences of financial markets.**
- **I’m excited to see more on how new statistical tools and platforms might transform financial markets.**