Discussion of:

Making Sense of Soft Information: Interpretation Bias and Ex-post Lending Outcomes

Michael Minnis
University of Chicago Booth School of Business
Overview

• Two research questions:
  • How is soft information related to loan outcomes?
  • Do cognitive constraints mediate the relation between soft information and loan outcomes?

• Approach:
  • Gain access to textual notes written by bank employees to measure soft information.
  • Measure loan outcomes as write-offs, delinquencies, etc.

• Findings:
  • Measures of soft information negatively related to bad loan outcomes.
  • Cognitive constraints mitigate this relation.
My view

• Very cool setting – and probably the closest we can come to actually “seeing” soft information in an empirical way.
• Much improved relative to a version I read last winter.
• Suggestions to consider for this version—or maybe another paper.
  • Jumps right into cognitive constraints – less emphasis on soft info
  • Measures of soft information: How to measure this?
• I’m going to focus on soft information angle.
The strengths

• Contributes by testing theory: soft information is important, but it has been difficult to measure
  • Agarwal and Hauswald (2010) derives an implicit measure
  • Papers such as Berger et al. (2017) and Drexler and Schoar (2016) infer that something else is going on (e.g., “expertise” or “relationship”) but typically infer this.
  • Others have also collected some internal data, but far from conclusive.

• Also contributes to “cognitive constraint” literature by looking at loan officers.

• Great data: even traced loan officers to LinkedIn

• And important broader question: What is the role of soft information and cognitive constraints with “fintech”?
What is soft information? (Petersen 2004)

• Hard information: The firm generated $10 million in sales.
  • Not just a number.
  • An agreed upon approach to calculating.
  • Can be sent and receiver knows everything sender knows.

• Soft information: The firm owner is honest.
  • Can be put on a scale from 1 to 10.
  • But we might disagree what number it is.
  • Open to interpretation.

• Banks are special, in part because the cumulate soft information (e.g., Diamond 1984, Diamond 1991).
“Main” result: Table 2

<table>
<thead>
<tr>
<th></th>
<th>(I) Charge off</th>
<th>(II) Delinquency</th>
<th>(III) Bad customer</th>
<th>(IV) Credit score decline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soft information</td>
<td>-0.066***</td>
<td>-0.125***</td>
<td>-0.160**</td>
<td>-0.056</td>
</tr>
<tr>
<td></td>
<td>(-4.585)</td>
<td>(-3.033)</td>
<td>(-2.188)</td>
<td>(-0.920)</td>
</tr>
<tr>
<td>Credit score</td>
<td>-0.018***</td>
<td>-0.139***</td>
<td>-0.284***</td>
<td>-0.051***</td>
</tr>
<tr>
<td></td>
<td>(-5.278)</td>
<td>(-12.402)</td>
<td>(-11.907)</td>
<td>(-3.076)</td>
</tr>
<tr>
<td>Debt-to-income ratio</td>
<td>0.010**</td>
<td>0.115***</td>
<td>0.147***</td>
<td>0.157***</td>
</tr>
<tr>
<td></td>
<td>(2.358)</td>
<td>(11.871)</td>
<td>(8.361)</td>
<td>(11.147)</td>
</tr>
<tr>
<td>Loan interest rate</td>
<td>0.004***</td>
<td>0.037***</td>
<td>0.033***</td>
<td>0.017***</td>
</tr>
<tr>
<td></td>
<td>(9.353)</td>
<td>(38.027)</td>
<td>(19.096)</td>
<td>(12.619)</td>
</tr>
<tr>
<td>Loan exception</td>
<td>0.001</td>
<td>0.019***</td>
<td>0.039***</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.492)</td>
<td>(3.074)</td>
<td>(3.424)</td>
<td>(-0.465)</td>
</tr>
<tr>
<td>Secured loan</td>
<td>0.004</td>
<td>0.005</td>
<td>-0.054***</td>
<td>0.015*</td>
</tr>
<tr>
<td></td>
<td>(1.451)</td>
<td>(0.781)</td>
<td>(-4.450)</td>
<td>(1.674)</td>
</tr>
<tr>
<td>Loan amount</td>
<td>0.000</td>
<td>-0.009***</td>
<td>-0.011**</td>
<td>-0.018***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(-3.766)</td>
<td>(-2.499)</td>
<td>(-5.610)</td>
</tr>
<tr>
<td>Loan maturity</td>
<td>-0.001</td>
<td>-0.001</td>
<td>0.003</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td>(-0.971)</td>
<td>(-0.804)</td>
<td>(0.943)</td>
<td>(4.182)</td>
</tr>
<tr>
<td>Borrower tenure</td>
<td>-0.000</td>
<td>-0.004*</td>
<td>-0.007*</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(-0.289)</td>
<td>(-1.664)</td>
<td>(-1.817)</td>
<td>(-1.428)</td>
</tr>
<tr>
<td>Total number of accounts</td>
<td>-0.004***</td>
<td>0.004</td>
<td>0.007</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(-3.408)</td>
<td>(1.232)</td>
<td>(1.339)</td>
<td>(-1.147)</td>
</tr>
</tbody>
</table>
Suggestion 1: Big picture

• Does the use of soft information help?

• This is a big unanswered question that gets side-stepped in this paper a bit.

• Banks moving to less relationship banking and more codified/hard information, but does this result in better info?

• Idea: Take advantage of the 2005 organizational change of the credit union in which soft information was collected?
Suggestion 2: Model of lending decision

- What is objective function? Maximize profits or minimize charge-offs?
- What model of bank decision-making does the paper have in mind?
- Soft info helps bank make better lending decisions:
  - Lots of soft info $\Rightarrow$ could lead to NOT making a bad loan
  - This paper: All observations conditional on making a loan, then infers that more soft info $\Rightarrow$ less likely bad outcome. Why is this necessarily the case?
- Idea: Does soft information help the bank better price the loan?
  - Interact interest rate with soft information $\Rightarrow$ Interest rate should be a better predictor of future default with soft info?
Suggestion 3: Soft info vs. constraints

<table>
<thead>
<tr>
<th>Cognitive constraint?</th>
<th>Info type</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Hard</td>
<td>Soft</td>
</tr>
<tr>
<td>Yes</td>
<td>Paper very focused here</td>
<td></td>
</tr>
</tbody>
</table>
Suggestion 3: Soft info vs. constraints

• Paper immediately jumps into cognitive constraints
• Yet, little research investigates soft info in the first place.
• Suggestion:
  • What aspects of soft info seem to be most relevant?
  • How does soft info interact with hard info?
    • Complements? Substitutes?
• Why are main effects on cognitive constraints insignificant?
Suggestion 4: Soft measures

• Soft info construct: Information that loan officer has that is not codified in hard information source such as credit score.

• Soft info measures:
  • proportion of total text which is “soft” info;
  • absolute value of log(total text) orthogonalized to hard info sources.

• Very difficult to measure (because it is “soft” in the first place!), but I have concerns about both—but have suggestions!
Suggestion 4: Soft measures

• Proportion of “soft” text: Use a dictionary capturing potential soft info words.

• Dictionaries are nice because they have less discretion.

• But the words used imply both “good” and “bad” directions:
  • “degree” “education” “bonus” “happy” “good”
  • “overwhelmed” “frustrated” “ditch while he was drinking”

• Paper does not take advantage of the “direction” of these words but presumably different implications

• Idea: the discussion on pages 16-17 of paper including quotes of loan officers revealing (e.g., healthcare problems which may seem like a shock)
Suggestion 4: Soft measures

• Absolute value of orthogonalized log(total text)
• Based on Agarwal and Hauswald (2010) approach, but seems less applicable here.
• Why absolute value? One word seems like the least amount of soft information, but would be coded as the most?
• Suggestion: Just use log(total text)
  • Natural measure of “how much” they know
  • Still no discretion involved
  • Avoids issue of proportion or absolute value
  • Keep observations with no text ➔ No soft info?
Suggestion 4: Soft measures

• Consider the honesty example:
  • “Owner is honest”
  • “Owner is dishonest”
  • “I give the owner an honesty rating of 3 out of 10”

• I think the real question is whether the officer’s assessment predicts performance in expectation.

• Credible null: “I met J. today...what a guy! He slapped me on the back about eight times.” ⇒ Sign that a loan officer going to misuse soft info?
Suggestion 5: Call center

• Cool supplementary test with the call center loans.
  • When loan officer busy, borrower randomly assigned to call center personnel.
  • Find similar results → Mitigate endogenous matching.

• But what else could be happening?
  • Recall definition of hard information: Receiver knows everything sender knows.

• Drexler and Schoar (2016) find that when loan officer away:
  • Credit to borrowers declines and lending decisions get worse.
  • They infer the presence of non-transferable soft information.

• Did this credit union find a way to “transfer” soft information?
Other thoughts

• Generalizability
  • What is the objective of a credit union? Profitability? Helping members?
  • Top productivity credit union. Implications?

• Interest rate
  • In theory, shouldn’t this capture both hard and soft info?
  • If so, should this be included as a control variable?

• Cross sectional tests where soft info improves outcome?
  • Situations in which soft info particularly valuable.
Concluding comments

• Important topic:
  • Timely given fintech developments → hard information
  • Timeless related to theoretical constructs

• Great setting with potential for novel insights.

• Possibilities for even more follow up papers?
“Yeah, but good luck getting it peer-reviewed.”