

Textual Analysis of Short-seller Research Reports

by Jules H. van Binsbergen, Xiao Han, and Alejandro Lopez-Lira

Discussant: Asaf Manela

Reichman University and Washington University in St. Louis

Jan 2023

Questions

1. What type of information is contained in short-seller research reports?
2. Do reports affect target firms' prices via cash flow or discount rate channel?



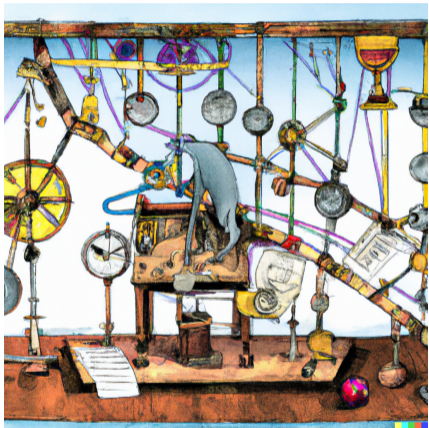
What this paper does

- ▶ Analyze the text of short-seller reports
- ▶ Use analyst CF projections to tease out the channel
- ▶ Main findings
 1. Target firm price drops -4.7% upon publication, and another -13% over subsequent year
 2. Reports more related to fraud cause larger long-term price drops
 3. Changes in CF expectations account for 70% of the price drop
 4. Target firms reduce future real investments and issue less stock

Contribution

- ▶ First to use ML to study topics in short-seller reports
- ▶ First to estimate the cash flow news share
- ▶ Closest relatives
 - ▶ Ljungqvist and Qian (2016 RFS): target firm prices respond strongly to short-sale campaigns
 - ▶ Paugam, Stolowy, and Gendron (2021 CAR): study stylized narratives related to credibility-based, emotions-based, and logic-based rhetorical strategies and assess their impact

Suggestion 1: Explain better why ML is needed



- ▶ ML is sometimes used to automate tasks on big data
- ▶ But only 528 reports here
- ▶ Perhaps the methods are more general and this is just a demo?
- ▶ Perhaps you are worried about overfit?

Suggestion 2b: More careful with the tools

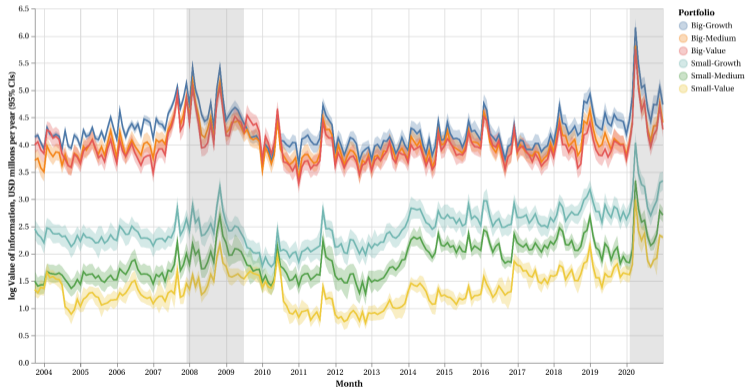
If ML textual analysis is key contribution beef it up

- b. Word embedding models like word2vec embed words in a relatively low dimensional vector (300d)
 - ▶ Paper measures similarity of research reports to the vector for 'fraud'
 - ▶ word2vec works at the word level
 - ▶ Unclear how you aggregate to the entire report
 - ▶ Newer language embedding models are much better at embedding context and longer sequences of text
 - ▶ Jha-Liu-Manela (2020) shows advantages of newer models like BERT and how to measure similarity in a dimension of interest

Suggestion 3: Emphasize firm types most targeted

- ▶ Paper finds that large growth firms are more likely to be targeted
- ▶ Kadan-Manela (2020) studies the value of information to a strategic investor

$$\text{Value of information} = \frac{\text{Return variance}}{\text{Price impact}}$$



- ▶ Do short sellers target firms with higher value of information?

Suggestion 4: Positioning

Pick one: Identification vs. prediction

1. Identification paper on the causal effect of short-seller reports on stock prices and nonprice outcomes
 - ▶ Important question and challenging problem
 - ▶ Tie to theories and their parameters
 - ▶ Clarify how the analyst CF projections help here
 - ▶ Ideally need an exogenous shifter to the report release decision
2. Predicting stock prices after the reports
 - ▶ Rationale for automating the textual analysis to beat other traders
 - ▶ Try to be as realistic as possible with trade timing
 - ▶ Market efficiency / gradual diffusion of information literature
 - ▶ Great if you can point to specific frictions

My Take

- ▶ Short-seller reports are a fascinating phenomena
- ▶ Good candidate for textual analysis
- ▶ Current draft is still a bit raw, but shows lots of promise
- ▶ Interesting connection to value of information lit
- ▶ Pick one: identification or return prediction



Appendix / Minor Comments

- ▶ Are buy-side analysts less sluggish in revising their CF projections?
- ▶ Are there short-sale constraints that explain the gradual decline?