Abstract

Our team explored state-level mortgage enforcement actions and experimented with a new approach to understanding the financial crisis from a bottom-up perspective. This project was done in two phases: collecting and organizing mortgage enforcement actions in the states of Ohio and Massachusetts from 2000 - 2010, and then analyzing both trends and texts in these enforcement actions. Before we continued with analysis, we used a mixture of specialized domain knowledge and NLP to process the text and used LDA to categorize texts into different topic models. Our primary finding was that regulations that were signed or enacted proactively were effective deterrents of malpractice.

Our Data

The primary focus of our data was on Mortgage Enforcement Actions. A mortgage enforcement action is a legal document issued by a state regulatory agency that notifies the guilty party of malpractice charges and the associated punishments. Ohio stored their MEAs as PDFs on a government database, while Massachusetts stored their MEAs as webpages. We extracted text from these data sources and filtered out the relevant sections from the text due to the presence of irrelevant legal clauses that colored our analysis.

Methods

After we extracted the text from various sources, our focus turned to NLP (using SpaCy) and LDA (using gensim) to produce topic models that classified wrongdoing in the text. We tuned our methods by removing specific irrelevant terms, ("commonwealth") and adjusting the number of topics, to find topic models with high coherence (high c_v scores) and understandability.

NLP Workflow

- Extraction of Texts
- Tokenization
- Stopword Removal
- Lemmatization
- Stemming
- Document-Level Feature Set
- Ngrams

Topic Modelling Visualization

- Topic Modelling Process
- Topic Modelling Evaluation
- Topic Modelling Interpretation

Constructing Ngrams

- Extracted Texts
- Remove irrelevant terms
- Visualize topics with pLDAvis

Ohio Findings

Our primary finding for the state of Ohio was centered around the 2007 Homebuyer’s Protection Act. Prior to this bill, not only was Ohio not as strict, but its main regulatory focus was reactive. This bill represented a shift in policy priorities, as regulations became proactive, and much more stringent. This resulted in a steep decline of MEAs in the ensuing years because it was now harder to obtain a license to issue loans and the fees to which lenders were held tightened. Thus, we see a decrease in MEAs across the board, but especially those related to criminal backgrounds and conduct.

Figure 1: Ohio MEAs classified as “Criminal Conduct” and “Criminal Background” (2000 – 2010)

We also found that when Ohio pursued reactive legislation, their regulatory agencies adapted to the new change, but it did not prevent any actions. When Ohio passed legislation in 2001 that established a mandatory education requirement, as well as a mandatory number of hours to be able to maintain ownership of a license, we see the number of administrative MEAs increase steadily. This implies that the changing rules did not impact the behavior of current actors in the mortgage market but did reveal a problem in the market that had previously been unaddressed.

Figure 2: Ohio MEAs classified as “Administrative” (2000 – 2010)

Conclusions

Our work this summer has validated this new experimental approach to understanding the dynamics between state regulatory agencies and mortgage lenders, brokers and loan originators. Our summary analysis of overall trends in MEAs revealed that state legislation of mortgage regulatory statues was very much needed and effective. Our topic modeling revealed the types of activities that state regulatory agencies identified and sought to regulate and showed clearly the dynamics between regulation and enforcement. The patterns and trends we found in the MEAs can provoke further questions about enforcement priorities and sentiments on the ground that the Bass Connections oral history team can look to answer.

References and Acknowledgements

Our team would like to thank Ms Pamela Prude-Smithers and her team in the Ohio Department of Commerce for being so accommodating with data requests and clarifications regarding data classification and labels, as well as specific legal terminology. We would also like to thank Mr Lee Reiners and Mr Joseph Smith for setting up and maintaining this project for the past 3 summers, as well as the domain guidance provided. We would like to thank Malcolm for laying out the roadmap for the project, for providing much-needed technical advice and for clarifying the roles of data scientists and Kate Wright for her assistance in clarifying legislative details for Ohio and Massachusetts.

Our Sources