Does Finance Benefit Society? A Language Embedding Approach

Manish Jha, Hongyi Liu, Asaf Manela (2024, *RFS*)

解读人：李子明

武汉大学金融系

2025.12.28

Overview

- Book sentiment-reflecting assumption
  - Choice of words used by books **reflect sentiment of that language** during the time.
- Measure a language's sentiment toward finance across time
  - Select **finance-mentioning sentences** published in the language and year.
  - Measure degree to which each sentence places finance in **positive context**.
  - **Aggregate sentiment** reflect attitude toward finance of books written in the language.
- Sentiment toward finance evolve over time and differ across countries
  - Time span: 1500 to 2009
  - Language coverage: 8 major languages

Motivation

- How financial industry perceived by society should be deeply cared about.
  - Public perception often diverge from economists (Goetzmann et al., 2016).
  - Positive sentiment toward finance spread its benefit (Zingales, 2012), suspicion toward finance restrict credit, risk sharing, and market efficiency (Gennaioli et al., 2022).
  - Survey reveal that trust in finance fall after 2008 (Sapienza and Zingales, 2012).
- Recency of survey limit learn how finance change over time and differ across country.
  - **Books** allow to travel through time and across borders and to study public perception about benefits of finance to society.
  - Computational linguistics approach can analyze text of millions of books.

Research Question

- How does society view finance?
  - Books written in language of capitalist countries discuss finance in more positive context.
- How does this view evolve over time and differ across countries?
  - Finance sentiment differences across countries mostly persist except China.
  - China exhibit greater volatility and positive sentiment as Italy and France.
  - Sentiment decline 1 year before rather than after financial crises.
- How does finance sentiment relate to macroeconomic growth?
  - Countries where finance sentiment increase to be more productive and to utilize more financial services.

Contribution

- Contribute to literature on public attitudes toward financial sector.
  - Prior: country's language and religion predict its creditor rights (Stulz, 2003); general lack of trust reduce stock market participation (Guiso et al., 2008).
  - This paper provide a measure of sentiment toward finance that span over century.
- Contribute to literature on culture and its effects on economic outcomes.
  - Prior: Change in culture and ideas tied to enrichment in 20th century (Mokyr, 2016); positive role finance play in society depend on public sentiment (Zingales, 2015)
  - We support the idea that change in sentiment toward productive parts of society have important effect on economic growth.
- Contribute to literature on textual analysis.
  - Increase in availability of text prompte interest in analysis of culture (Michel, 2011).
  - Prior: yet to study sentiment toward any particular sector.
  - This paper study public sentiment toward finance.

Data

- Textual data: Google Books Ngram Corpus
  - Google scan of over 8 million books (6% of all books published)
  - Consist of words and phrases and annual usage frequency from 1500 to 2009.
  - Cover books in 8 languages: American English, British English, Simplified Chinese, French, German, Italian, Russian, Spanish
  - Lower complexity n-gram counts also provided, but choose 5-gram because word's context essential for BERT. (N-gram: sequence of n consecutive words)
  - Strip case, symbols, double spaces, part-of-speech tags, and positional tags.
  - Extract all sentences mentioning stem of word for finance.
- Macroeconomic data: macro-history data set (Taylor, 2017)
  - GDP growth: annual percentage increase in real per capita GDP.
  - Credit growth: annual change in real total loans to non-financial private sector.

Step1: Sentence Embedding

- BERT embed each sentence in corpus into 768 dimensional vector space
  - **Work well out of box**: pre-trained on BookCorpus and Wikipedia
  - **Offer contextualized embedding**: distinguish connotation difference of sentences
  - **Reduce unique words**: break word into smaller subwords or tokens

Step2: Positive Minus Negative Embedding

- Language features treated as vector spaces with intuitive mathematical properties.
  - Vector arithmetic capture semantic relationships between words (Mikolov et al., 2013)
  - $vec(\text{"}king\text{"}) - vec(\text{"}man\text{"}) + vec(\text{"}woman\text{"}) \approx vec(\text{"}queen\text{"})$
  - $vec(\text{"}dictator\text{"}) + vec(\text{"}positive\text{"}) - vec(\text{"}negative\text{"}) \approx vec(\text{"}king\text{"})$
  - Positive minus negative represents a displacement in positive direction.
  - Word pairs correspond to positive step: benefit & damage, good & bad, help & hurt.
- Standard for positive and negative embedding
  - Average embedding difference across sentence containing finance with above word pairs.

Positive sentences	Negative sentences
financial services benefit society	financial services damage society
finance is good for society	finance is bad for society
finance professionals are mostly good people	finance professionals are mostly corrupt people
finance positively impacts our world	finance negatively impacts our world
financial system helps the economy	financial system hurts the economy

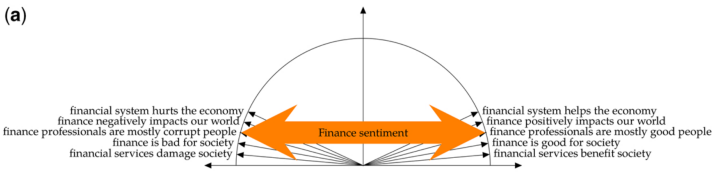
Step3: Cosine Similarity

- Define positivity dimension
  - Similar sentences keep similar company: five positive and negative sentences bunch together in embedding space
  - Language-specific positivity dimension p_i : average of positive minus negative sentence embedding for language i.
- Orthogonal projection onto positivity dimension
  - For each finance-mentioning sentence j in language i with embedding s_{ji} , calculate cosine similarity with language-specific positivity dimension p_i

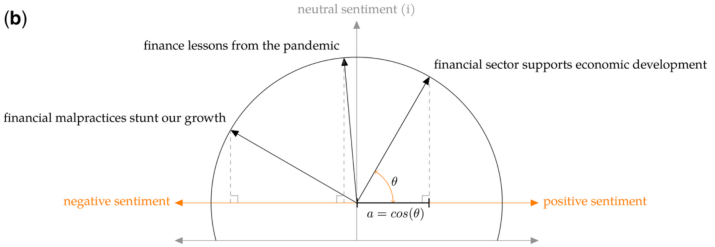
$$a_{ji} = \frac{S_{ji} \cdot P_i}{\|S_{ji}\| \, \|P_i\|} = \frac{\sum_d S_{jid} P_{id}}{\sqrt{\sum_d S_{jid}^2} \sqrt{\sum_d P_{id}^2}}$$

  - Cosine similarity between -1 and $+1$, zero indicating neutral sentence.

Conceptual Diagram of Finance Sentiment Measurement



Defining the positive minus negative finance sentiment dimension



Projection of sentences onto positive minus negative sentiment dimension

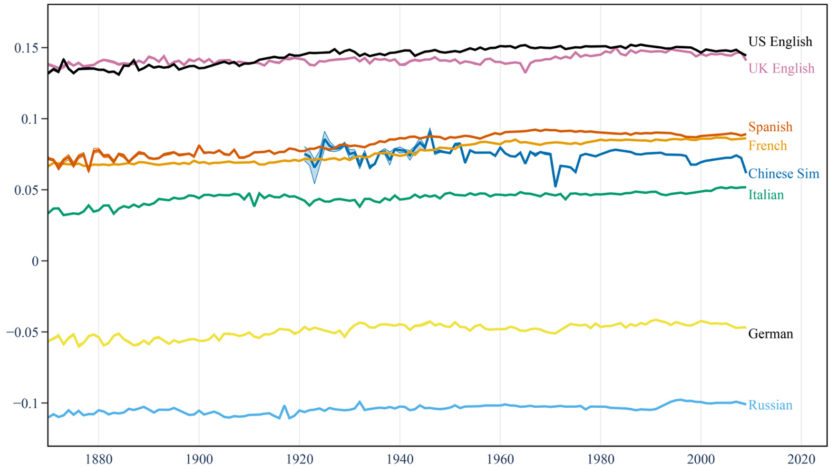
Step4: Aggregate to Annual Finance Sentiment

- Average cosine similarity of all finance-mentioning sentences occur in that language in year t, weighted by times sentence occurred that year

$$f_{it} = \sum_j a_{ji} \cdot \frac{c_{jit}}{\sum_k c_{kit}}$$

- Annual finance sentiment varies over time due to change in sentence occurrence c_{jit}
- Sentiment of particular sentences a_{ji} in each language i stay constant
- Languages evolve
  - Implicitly assume that language model constant, only frequency of language use varies
  - Change to American English do not greatly affect ability to predict with text (Manela and Moreira, 2017)

Finance Sentiment over Time and Across Countries



Capitalism and Finance Sentiment

- Regress finance sentiment on attitude toward capitalism
- Three measures: right-leaning government ; preference for private ownership; difficulty of registering business

	Finance sentiment								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Right-leaning government	0.310***	0.324***	0.00338						
	(0.100)	(0.116)	(0.00447)						
Preference for private ownership of business				0.489***	0.515***	0.00533			
				(0.0327)	(0.0333)	(0.0252)			
Difficulty of registering a business							−0.179***	−0.187***	−0.00261
							(0.00636)	(0.0109)	(0.00218)
GDP growth	0.0231	0.0246	−0.000342	0.116***	0.126***	0.00394*	0.00898	−0.0360	−0.00368
	(0.0243)	(0.0311)	(0.00109)	(0.0147)	(0.0194)	(0.00187)	(0.0302)	(0.248)	(0.00252)
Country FE			Yes			Yes			Yes
Year FE		Yes	Yes		Yes		Yes		Yes
Observations	263	263	263	67	67	67	35	35	35
R^2	0.0293	0.0518	0.999	0.522	0.551	1.000	0.222	0.231	1.000

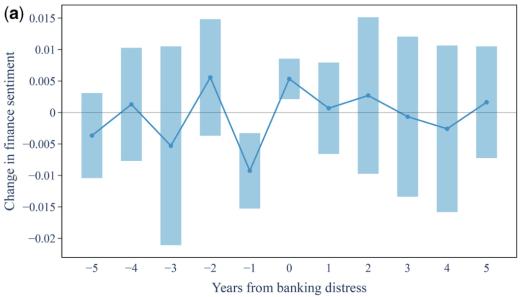
Association with Demand for Financial Services

- Direct measure: financial market participation
  - Credit market participation: total loan to household and to non-financial private sector
  - Stock market participation: equity as percent of total assets for household
  - Market participation positively correlated with sentiment toward finance
- Indirect measure: income inequality
  - Annual data on income inequality from World Income Inequality Database (WIID)
  - Income and wealth inequality negatively correlated with sentiment toward finance

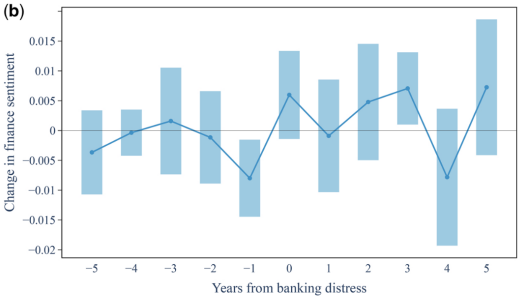
	(1) Equity/Total for households (%)	(2) Households loans / GDP	(3) Non-financial private sector loans / GDP	(4) Income inequality
Finance sentiment	57.77*	90.70***	102.6***	−45.42***
	(27.65)	(31.00)	(29.66)	(11.31)
GDP growth	−0.377	0.00873	−0.311	−0.0175
	(0.264)	(0.255)	(0.210)	(0.0944)
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	75	431	733	216
R^2	0.952	0.811	0.718	0.880

Association with Financial Crises

- Lead-lag relationship
  - Finance sentiment to decline following financial crisis
  - Decline in finance sentiment itself transform mild recession into fullblown financial crisis



Baseline



Controlling for credit growth and GDP growth

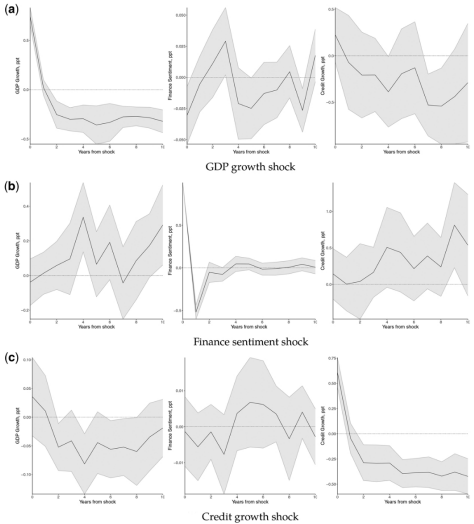
Association with Economic Growth

- Cumulative impulse response function separate shocks from other macroeconomic shocks to output and credit (Jordà, 2005)

$$\Delta_h y_{i,t+h} = \alpha_i^h + \sum_{k=1}^{3} \beta_k^h \, \Delta f_{i,t-k} + \sum_{k=0}^{3} \gamma_k^h \, X_{i,t-k} + \varepsilon_{i,t+h}, \qquad h = 0, \ldots, H$$

- $\Delta f_{i,t}$: shock to finance sentiment
- $\Delta_h y_{i,t+h} = y_{i,t+h} - y_{i,t}$: h-year cumulative growth of GDP or credit
- $X_{i,t}$: vector of control variables (three lags of credit growth and finance sentiment)

Impulse Responses of Output, Sentiment, and Credit



Discussion

- Caveat to finance sentiment from books
  - Can languages be traced to a major geographical area or a distinct country?
  - Can books represent general population or just reflect literary elite?
- Limitation of cross-language and country comparisons
  - Many characteristics differ across borders and cultures.
- Causal effect of finance sentiment on particular outcomes
  - Exogenous variation in finance sentiment
- Origin of finance sentiment fluctuation
  - How finance sentiment responds to natural disasters?