

Measuring Financial Risk in China using Textual Analysis

The 12th annual international conference on the Chinese economy

Hong Kong Monetary Authority



Martino Ricci European Central Bank Joint with A. Al-Haschimi, A. Apostolou, A. Azqueta

23/11/2022

Overview of the presentation

News-based financial risk measure in China:

- Motivation
- Literature

Analytical approach

- Data collection
- Standard textual analysis (count of articles)
- Topic Modelling unsupervised machine learning for textual data (*Latent Dirichlet Allocation*: LDA)
- Construction of the index and decomposition or risk

Empirical analysis

• Impact on macro-financial variables of increase in China's financial risk



Objectives and Motivation

Motivation: Financial risk in China on the rise



Real Estate Investment: Ratio of Liabilities to Assets (%)



Source: BIS and China Statistics Yearbook, National Bureau of Statistics (NBS) of China via Haver Analytics. Notes: Credit to nonfinancial corporation is from all sectors and at market value, namely the price at which the asset would change hands if sold on the open market. It includes both credit to households and to nonfinancial corporations.

Motivation: Monitoring China using hard data is becoming increasingly difficult

China is becoming much less transparent about its economic performance, quietly discontinuing thousands of statistical series, especially in industrial production

Annual number of different economic indicators made available by China's National Bureau of Statistics, by category



Source: FT analysis of CEIC; Chinese National Bureau of Statistics. Series have been checked against historical reports to ensure that the demonstrated drop is due to discontinued series, and not a lag in publication FT graphic: John Burn-Murdoch / @jburnmurdoch © FT

Literature of financial risk in China and text based approaches

Text-based approaches in Economics:

- Baker et al. (2016): "Economic policy uncertainty".
- Caldara et al. (2019) "The economic effects of trade policy uncertainty".
- Caldara and Iacoviello (2019): "Measuring geopolitical risk".
- Angelico et al. (2020) "Can We Measure Inflation Expectations Using Twitter?".

Financial risk/text based work on China:

- Song et al. (2018) "Risks in China's financial system".
- Ma et al. (2019) "Monitoring of china's systemic financial stress".
- Wright and Rosen (2018) "Credit and Credibility: Risks to China's Economic Resilience".
- Davis et al. (2019) "Economic policy uncertainty in China since 1949: The view from mainland newspapers."
- Huang and Luk (2020) "Measuring economic policy uncertainty in China."
- Chen and Tillmann (2021) "Monetary policy uncertainty in China".



2. The construction of the Financial Risk Index

Using textual analysis and topic modelling to gauge China's financial risk and sources

Overview of the Analytical Strategy

- Step 1: We collect data (newspaper articles) from two authoritative sources on China and financial markets
- Step 2: We perform standard textual analysis -> filter articles relevant to increasing *financial risk*; data and pre-processing
- Step 3: Topic modelling -> unsupervised machine learning: assign a probabilistic distribution to identified articles of talking about specific topics.
- **Step 4:** Construct time-series based on topics; select most relevant topics aggregate in a single indicator and validate the measure.

Step 1: Data and pre-processing THE WALL STREET JOURNAL.

South China Morning Post

Data, sources and timespan

- We collect articles from the **SCMP** and **the Wall Street Journal**.
- English version of the newspaper.
- Printed editions only (more constant reporting).
- January 2005 to September 2022. Total of around 36.000 articles.

Step 2: Filtering articles discussing increasing financial risk in China

Filtering articles discussing **financial risk** in **China** over time using **keywords**:

- I. Initial terms: start search with all articles containing words *China/Chinese, risk* and *finance*.
- II. Expand initial search using *word2vec* algorithm.
- III. Select most relevant words associated with negative sentiment (i.e. danger/opaque/defaults).
- IV. Repeat the search.

Final search:

 {(china or chinese) and (financ\$ or bank\$ or lend\$ or commerc\$ or shadow bank\$) and (risk\$ or lever\$ or problem\$ or opaqu\$ or pain\$ or unstable or danger\$ or default\$ or unwind\$ or deterioat\$ or vulnerab\$ or volatil\$ or negative\$ or contagio\$ or imbalance\$ or hazard\$ or trigger\$ or delever\$ or bad or uncertain\$ or jeopard\$ or system\$ or serious\$ or consequ\$ or shock\$ or bubbl\$ or stress\$ or threat\$ or burden\$)}

Outcome: We obtain around 2000 articles per year related to financial risk in China.

Step 3: Topic Modelling and the LDA algorithm

Topic modelling is a type of statistical modelling for discovering the abstract **"topics**" that occur in a **collection of documents**. Latent Dirichlet Allocation (**LDA**) is an example of topic model and is used to classify text in a document to a particular topic.

- It uses unsupervised machine learning to unveil the distinctive topics of overall financial risk.
- This method does not require selecting additional specific keywords (unsupervised).
- It is not language specific, and can be easily adapted to multi-language setups.
- We can construct a wide range of sub-indicators related to financial risk: Banking, financial markets, trade and commodities, exchange rates....-> filter topics / decompose sources of risk
- We find an **optimal** number of topics equal to **30**

Step 3: Topic Modelling, graphical representation

Ξ

THE WALL STREET JOURNAL MARKETS | HEARD ON THE STREET China 2020: Trade Risks Become Debt Risks With trade risks receding, keeping China's rickety financial system on the rails will dominate 2020 By Nathaniel Taplin Topic proportions Dec. 27, 2019 11:25 am ET A SHARE AA TEXT China's terrible, no good, very bad year is finally drawing to a close. Next year is set to start off better, but there are some nasty icebergs lurking beneath apparently calmer waters. Investors could find opportunities in trade-exposed equities, like Foxconn. But the risk of serious financial turbulence, which they barely avoided in 2019, is rising. Though most investors didn't notice, key sections of China's economy were already on the mend by late 2019, particularly the auto sector-which was mired in a deep downturn for most of 2018 and 2019-and the electronics industry. Auto output rose on the year in November for the first time since June 2018, And electronics makers' profits have also rebounded strongly in recent months as **Topics** global smartphone and semiconductor sales have begun to rise again. In addition, the trade detente with the U.S. reached this month should help ensure some the healing of the labor market, already under way thanks to the 0.8 Trade 0.4 electronics rebound. Economy Housing 0.5 0.1 U.S. In two other key parts of the economy, things look far less rosy. Real estate profit Auto 0.3 Real estate 0.4 growth is slowing rapidly-along with the housing market itself. That, in turn, is 0.2 Electronics Infrastructure 0.1 putting pressure on state-owned industrial companies, which are concentrated 0.1 Industry in construction-dependent sectors like steel. Real estate and government-owned Debt 0.5 industry, along with infrastructure, also happen to be the sectors where China's Financial 0.3

Step 3: Topics need to be labelled. We do this looking at most frequent words and reading most representative articles

Banking







Real Estate

development broject buiding square sell sale land shanghaiprice land buy property residential event square sell sale land shanghaiprice land buy property buiding solution housing office commetal

Corporate Profitability



Corporate Investment



Step 4: Financial risk in China – aggregated index



Notes: ECB staff calculations. Time coverage 2005/01-2022/09. Y-axis in standard deviation with mean zero over the whole horizon.

Step 4: Financial risk in China - decomposition of risk



Notes: ECB staff calculations. Time coverage 2005/01-2020/11. Y-axis in standard deviation with mean zero over the whole horizon.



3.

Empirical analysis

How heightened financial risk affects China and the rest of the world

Empirical analysis: The Model

Table 2

VAR setup	Source
ln(global industrial production ex	Netherlands Bureau for Economic Policy
China)	Analysis/Haver Analytics
ln(oil prices)	Financial Times/Haver Analytics
EMBI Global spread	JP Morgan/Haver Analytics
ln(CN equity price index)	Dow Jones/Haver Analytics
Financial risk measure	Authors' calculations
ln(Chinese industrial production)	China National Bureau of Statistics/Haver Analytics
ln(CN CPI)	China National Bureau of Statistics/Federal
	Reserve Bank of Atlanta
Chinese 7-day repo rate	People's Bank of China/Federal Reserve
	Bank of Atlanta

Notes: Oil prices refer to brent crude oil (\$/bbl), Brent crude oil prices starting in 2007 are Brent blended prices from the Financial Times. Prices prior to 2007 are European Free Market Prices from the Wall Street Journal. The Chinese equity price index is the Dow Jones China 88. Industrial production measures and the consumer price index are seasonally adjusted by sources. The Federal Reserve Bank of Atlanta provides seasonally adjusted monthly data series for the Chinese economy (for additional details see also Higgins et al., 2016).

Empirical analysis: Impact of financial risk shocks to macro-financial variables





Conclusions

- We construct a text-based index of financial risk in China.
- The index seems to correctly accounting for past instances of increasing risk in China.
- The index has the advantage of being timely and can be decomposed in further sub-indices.
- Although being low compared to previous events the index shows some recent spikes due to problems in real estate.
- An empirical exercise shows that an increase in risk tightens financial conditions in China and has negative repercussion on (global) macro financial variables.

Thank you

Background slides

Financial risks in China: Difficult to extract signals from data

"In China, not all triple-A bonds are created equal" (WSJ)

Outstanding corporate bonds

Rated AAA
 With other ratings
 Unrated





1 trillion yuan=\$149.77 billion. 2020 data is as of Oct. 12. Sources: Wind, Invesco Asset Management Alternative estimates of non-performing loans

WSJ (Oct. 2020): "In China, not all triple-A rated bonds are created equal"

More than half of corporate debt in China rated AAA

In the US, bonds of only two companies are rated AAA by S&P

PBoC identified systemic risks from hidden links between non-financial holding companies and banks, how to measure this?

News database used for the analysis: Dow Jones Factiva

- More than **35,000 sources** from more than 200 countries in 26 languages.
- Nearly 400 continuously updating newswires.
- **Newspapers**: Same-day and archival coverage of several newspapers (e.g. The Wall Street Journal, The New York Times, Washington Post..).
- Factiva Search: Gives full control to build precise searches using a combination of keywords and Dow Jones Intelligent Indexing.



Freedom of the press: China ranked 177 out of 180 nations



Step 2: Filtering articles discussing increasing financial risk in China



Count the **occurrence** of **news articles** describing **financial risk** in **China** over time using **keywords**:

- Initial terms: start search with all articles containing {China or Chinese AND risk\$ AND finance\$}.
- II. Expand initial search using *word2vec* algorithm.
- III. Select most relevant words associated with negative sentiment (i.e. danger/opaque/defaults).
- IV. Repeat the search.

Step 1: Final search

{(china or chinese) and (financ\$ or bank\$ or lend\$ or • **commerc\$ or shadow bank\$)** and (risk\$ or lever\$ or problem\$ or opaqu\$ or pain\$ or unstable or danger\$ or default\$ or unwind\$ or deterioat\$ or vulnerab\$ or volatil\$ or negative\$ or contagio\$ or imbalance\$ or hazard\$ or trigger\$ or delever\$ or bad or uncertain\$ or jeopard\$ or system\$ or serious\$ or consequ\$ or shock\$ or bubbl\$ or stress\$ or threat\$ or burden\$)}

Outcome: We obtain around 2000 articles per year related to financial risk in China.

Step 4: Topics probabilities are then aggregated to form time-series



Topic Modelling: other topics

unk1



Domestic Policy-Party Issues



Investment



Trade Wars



Global Issues



Growth Outlook



Local Policy



Infrastructure



unk2



Law and Regulation



Foreign Affairs





Topic Modelling: other topics



Fintech



Industry and Manufacturing



Entertainment











Trade and Commodities



Firms and Technology

person of the second of the se

Environment



Geopolitics Asia



Energy



Art and Culture



Topic Modelling: identified topics

































Topic Modellina: identified topics











2006

2010

2014

2018



Pandemic and Health

0.04

0.02

0.03

0.02

0.01

0.015

0.005



0.005

2022

Topic Modelling: Proximity analysis



Topic Modelling: Comparison with other indicators





2006 2008 2010 2012 2014 2016 2018 2020 2022

China financial risk index



www.ecb.europa.eu ©

Alternative risk measure

Topic Modelling: Comparison with other indicators



Topic Modelling: unveiling China's sources of financial risk



Topic Modelling: unveiling China's sources of financial risk



leuropa.eu ©

VARIABLES	(1) DowJones	(2) DowJones	(3) DowJones	(4) DowJones	(5) DowJones	(6) DowJones
EinancialMarkets	-40.92**	-32.05**	-34.45**	-34.83**	-36.80**	-33.35**
Banking	25.54 (19.74)	(15.5/)	(15.03)	(15:49)	(15.03)	(15.14)
FX	6.898 (14.18)					
RealEstate	-41.02 (37.23)					
CorpProf	-62.07** (29.60)	-27.53 (31.31)	-38.46 (25.51)	-39.03 (26.11)	-46.84* (25.90)	
CorpInx	57.27** (26.81)	72.40** (30.68)	82.40*** (25.79)	82.19*** (25.92)	112.1*** (31.28)	86.47*** (27.97)
GrowthOutlook		-9.293 (15.39)				-31.31** (13.99)
LeisANDHosp		-162.5*** (51.32)	-163.9*** (51.18)	-165.0*** (52.29)	-205.0*** (56.65)	-222.1*** (56.93)
Energy		-92.48** (42.13)	-95.62** (41.74)	-96.18** (42.17)	-113.3*** (42.90)	-112.7*** (42.72)
IndANdManu				4.600 (42.69)		
ForeignAffairs					48.96* (29.47)	71.78** (32.38)
Constant	0.544* (0.290)	1.507*** (0.344)	1.452*** (0.331)	1.445*** (0.337)	1.183*** (0.367)	1.185*** (0.363)
Observations Adjusted R-squared	213 0.056	213 0.126	213 0.129	213 0.124	213 0.136	213 0.143

Table B.1.: Regression analysis results with China Dow Jones index as dependent variable.

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

.ecb.europa.eu ©

variable.	reegi cooroni ana	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		b mach ab acp	ondone	
VARIABLES	(1) CISS	(2) CISS	(3) CISS	(4) CISS	(5) CISS	(6) CISS
FinancialMarkets	-11.22	-13.57	-12.07	-9.304	-10.82	
	(16.64)	(15.90)	(15.78)	(15.69)	(15.50)	
Banking	32.50*	29.32*	32.17*	29.46*	38.05**	35.47**
	(18.20)	(17.40)	(17.02)	(16.92)	(17.07)	(16.79)
FX	-2.065	-14.92	-13.88	-10.32	-3.117	
_	(13.18)	(12.93)	(12.85)	(12.85)	(13.03)	
RealEstate	-44.55	-61.15*	-58.36*	-78.32**	-72.54**	-75.74**
	(35.79)	(34.38)	(34.18)	(35.18)	(34.81)	(34.10)
CorpProt	89.78***	24.31				
	(27.17)	(29.99)				
CorpInx	96.02***	148.3***	159.4***	185.0***	138.1***	132.3***
	(26.10)	(27.66)	(23.99)	(26.74)	(32.74)	(32.05)
GrowthOutlook		65.46***	71.58***	85.63***	80.98***	77.79***
		(15.04)	(13.00)	(14.52)	(14.46)	(13.73)
IndANdManu				-88.13**	-87.14**	-94.28**
				(42.03)	(41.49)	(40.51)
ArtandCulture					134.8**	141.3***
_					(55.76)	(53.97)
Constant	-1.986***	-2.419***	-2.401***	-2.236***	-2.234***	-2.206***
	(0.263)	(0.270)	(0.269)	(0.278)	(0.274)	(0.268)
Observations	192	192	192	192	192	192
Adjusted R-squared	0.249	0.316	0.317	0.329	0.346	0.350

Table B.2.: Regression analysis results with China CISS index as dependent

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1