HOT OFF THE PRESS NEWS-IMPLIED SOVEREIGN DEFAULT RISK

Chukwuma Dim¹, Kevin Koerner², Marcin Wolski² and Sanne Zwart²

¹George Washington University and ²European Investment Bank

EEA-ESEM 2022 23 August, Milan

The views expressed in this presentation are those of the authors and do not necessarily reflect the position of the European Investment Bank.



Motivation - EIB perspective

EIBG on the ground

- Over 160 countries of operations
- 53 local offices (22 EU + 31 OEU)



Outstanding loan book as of 31.12.2021



Substantial need to monitor country-specific risk in real time (perfectly)



Assessing sovereign default risk

How to track sovereign default risks if the usual resources are not available?

- No CDS spreads
- No risk ratings
- Debt data published with a long lag, sometimes with imperfect quality

- Only about 75% of countries worldwide are rated
- 60% of LICs are experiencing debt distress or have high risks

Alternative data sources can help in assessing the current sovereign credit risk situation



Do newspaper articles have informative/predictive power over sovereign default risk?



3

Building a high-frequency proxy for sovereign default risk

Aim of our research: Apply news media text to develop a reliable high-frequency proxy for sovereign default risk that can be computed in real-time for any country without market liquidity constraints

Approach: Use a repository of international and local newspapers, employ Natural Language Processing (NLP) techniques to quantify sovereign default concerns, and assess properties of the measure

The framework does <u>not</u> rely on supervised learning which requires an outcome variable This allows us test the performance/additionality of the measure relative to existing proxies

Uses:

- Analyse trends over time in countries lacking high-frequency default risk indicators
- Analyse wider regional trends (due to straightforward aggregation)
- Conduct proactive monitoring (e.g., Early Warning Signals under IFRS9): real-time information
- Revisit theoretical predictions on the influence of default risk on e.g., equity markets



4

Main results

- 1. News text *alone* can deliver an informative News-Implied Sovereign Risk Index (NSRI)
- NSRI captures major global risks episodes: great recession, euro debt crisis, Covid-19 pandemic
 - Co-moves with market indicators of global risk – G7 sovereign CDS spreads and VIX
- 3. NSRI captures country-specific default risk
 - Comoves/predicts CDS spreads & provides default risk signals beyond CDS spreads
 - Has significant incremental information about future sovereign rating downgrades

Global News-Implied Sovereign Risk Index Weekly 3-month moving-average



Source: Authors' calculations, Refinitiv Datastream



CONSTRUCTING THE NSRI: MEASURING SOVEREIGN RISK USING MACHINE LEARNING



6

Constructing the NSRI





Constructing the NSRI: 1) Text corpus

Objective: Obtain news covering countries' economic, financial and political developments

Approach: Use the PressReader News API over the period 2006 – 2022

- PressReader provides historical news coverage for many countries, including local and global news publishers
- For each country, retrieve **English language news** that mentions the country's name and one of the key words "economy" and "government" and published either in the country, the US, UK or China
- In total, we retrieved approximately 10 million news articles published by roughly 2,100 publishers covering 184 countries



8

Constructing the NSRI: 2) Topic modeling

Objective: Isolate narratives likely to matter for sovereign default risk to reduce noise in the text data

Approach: Use the Latent Dirichlet Allocation (LDA) algorithm to classify news text into a low-level number of topics $T \in \{4, 5, ..., 10\}$ within which the text data looks similar

- LDA gives text a hierarchical structure, where documents (news articles) are composed of topics which in turn contain words
- Each document has a probability distribution over latent topics (one document has many topics) and each topic is defined by a probability distribution over words (one word present in many topics)
- **Training the model** boils down to deriving the optimal parameters and, more importantly, the number of latent topics *T* that best fit the data
- Identified nine topics:
 - Economy, Business, Governance I & II, Security, Health, Environnent, Tourism, and Miscellaneous



Constructing the NSRI: 3) Text similarity

Objective: Quantify the level of sovereign default concern in each news article about a given country

Approach: Use the **Cosine Similarity algorithm** to compute how similar an article's language is to that of typical texts that signal heightened default risk

- Text similarity allows the use of **unsupervised learning techniques that do not depend on market data**
- Text similarity algorithms are simple, have low computational cost, and suits our setting where supervised learning seems problematic
- A complement of Cosine Distance, CS(u,v)= 1 CD(u,v)



Constructing the NSRI: 3) Text similarity (ctd)

Cosine Similarity (CS) is calculated as the inner product of two documents' vectors divided by the product of the vectors' Euclidean lengths

$$CS = \cos(\boldsymbol{u}, \boldsymbol{v}) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{u}| \times |\boldsymbol{v}|} \in [0, 1]$$

where, *u* is the vector of an arbitrary news article

v is the vector of a representative sovereign default risk document

u, *v* are based on *term frequency–inverse document frequency (tf-idf)* vectorization – *mitigates spurious association between document vectors*

Specify v as in studies that construct representative text in other domains (e.g., Hassan et al., 2019, Engle et al., 2020):

Compile a training library of 448 documents centered on worsening sovereign risk and default concerns:

- Textbooks on sovereign default
- Sovereign credit rating downgrade reports by Moody's, Standard & Poor's, and Fitch
- Excerpts from the EIB's sovereign rating rationales for lowly rated sovereigns
- Country risk analysis reports by the Economist Intelligence Unit
- Debt sustainability analysis reports by the IMF/World Bank
- News articles on sovereign defaults and debt problems

11

Constructing the NSRI: 4) Aggregation

Objective: Aggregate the information obtained from individual news articles by country and period

Approach: Average article-level CS scores at a daily frequency using only news articles in sovereign risk-relevant topics:

$$NSRI_{j,t} = \frac{1}{n_{j,t}^r} \sum_{i}^{n_{j,t}^r} CS_{i,j,t}$$

where $CS_{i,j,t}$ is the CS score of article *i* published on day *t* about country *j*

 $n_{j,t}^r$ is the number of news articles in the sovereign risk-relevant topics published about country j on day t

Sovereign risk-relevant topics are "Economy", "Business", "Governance" and "Security"



NSRI: VALIDATION



Constructing the NSRI: 5) Validation

Objective: Confirm **good separation of noise from likely relevant narratives**

Approach: Inspect topics and terms

S/n	Topic 1 Environment	Topic 2 Security	Topic 3 Miscellaneous	Topic 4 Governance I	Topic 5 Health	Topic 6 Business	Topic 7 Governance II	Topic 8 Tourism	Topic 9 Economy
1	oil	police	woman	right	health	business	president	water	market
2	energy	attack	world	law	pandemic	development	minister	food	bank
3	climate	kill	school	party	virus	project	leader	flight	growth
4	climate change	rebel	family	political	medical	technology	prime minister	tourist	economy
5	fuel	force	team	people	coronavirus	service	war	passenger	rate
6	coal	investigation	sports	court	vaccine	company	peace	tourism	investor
7	mining	intelligence	game	public	covid	sector	millitary	airline	tax
8	oil gas	report	cup	election	hospital	industry	sanction	ship	financial
9	emission	security	play	vote	disease	cooperation	opposition	airport	trade
10	electricity	arrest	love	corruption	patient	programme	administration	hotel	economic
11	oil price	government	film	government	outbreak	opportunity	state	road	debt
12	carbon	terrorist	book	act	infection	digital	meeting	island	price
13	power	millitary	friend	bill	lockdown	international	deal	transport	stock
14	renewable	millitant	football	freedom	people	local	political	travel	investment
15	natural gas	violence	player	state	doctor	partnership	nation	traffic	income
16	solar	army	young	rule	death	innovation	conflict	sea	currency
17	natural	civilian	life	voter	treatment	system	diplomatic	park	export
18	petroleum	crime	home	legal	health care	economic	international	vessel	sale
19	crude	protest	old	democracy	worker	infrastructure	negotiation	boat	gdp
20	gas	soldier	child	president	test	country	ally	route	business
% of News	4%	12%	12%	15%	6%	11%	15%	7%	17%

Security + Governance I + Governance II + Business + Economy \approx 70% of the news corpus



Constructing the NSRI: 5) Validation ctd



Terms with the highest tf-idf scores in our Sovereign Default Risk Library Font size is proportional to term weight



Constructing the NSRI: 5) Validation *ctd2*

News-level Topics and CS Scores vs. News Sentiment

Nows Article Title	(1)	(2) Sontimont	(3) Topic
News Article Title	CS	Sentiment	Topic
Panel A: News signaling sovereign default concern	5		
Zambia on brink of defaulting on foreign debt	0.81	0.72	Economy
Zambia's default fuels fears of African 'debt tsunami' as Covid impact bites	0.80	0.74	Economy
Argentina defaults yet again, but hopes to get off lightly	0.63	0.94	Economy
Already in default, Argentina hits an impasse with creditors over debt restructuring	0.62	0.82	Governance II
'Painful' downgrades will raise South Africa's borrowing costs, minister says	0.55	0.90	Economy
What the latest rating downgrades mean for the average South African	0.64	0.87	Economy
World Bank says world leaders moving away from debt cancellation for Africa	0.67	0.88	Business
Panel B: News on socioeconomic problems			
Myanmar: Military Coup Kills Fragile Democracy	0.13	0.91	Security
Nigeria's Economy Faces Worst Recession in Four Decades, says New Report	0.23	0.73	Economy
Recession: Nigeria's Economic Crisis Requires a Political Solution	0.31	0.69	Economy
Turkey's Erdogan sacks central bank governor after rate hike	0.11	0.50	Economy
Venezuela crisis: How the political situation escalated	0.08	0.87	Governance I
Chadian President Idriss Deby dies on frontline, rebels vow to keep fighting	0.13	0.78	Governance I
Nigeria's inflation rises again, hit four-year high February	0.12	0.50	Economy
Panel C: Miscellaneous news			
Minister orders intelligence-led operation against violence in South-East	0.03	0.89	Security
Man who drove into News Cafe in Rosebank appears in court	0.03	0.98	Security
Uber pledges to boost safety for SA drivers as accidents rise	0.07	0.67	Health
UK imposes sanctions on Russians, Guptas in first use of anti-corruption law	0.09	0.86	Security
Naomi Osaka withdraws from French Open amid row over press conferences	0.09	0.80	Miscellaneous
Chelsea beat Man City to win Champions League	0.07	0.58	Miscellaneous

CS is markedly different from news sentiment and sufficiently differentiates news signaling elevated sovereign credit risk from those that do not



NSRI: EMPIRICAL RESULTS



NSRI and CDS spreads

	NS	RI % chang	e rel. to <i>t</i> -	- 1	NSRI % change rel. to 3-month me					
	Contemp	ooraneous	Lag	ged	Contem	ooraneous	Lag	Lagged		
Dep. var.: CDS Spread % change	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Panel A: Weekly frequency										
NSRI % change	0.399	0.258	-0.055	-0.095	0.391	0.213	-0.132	-0.158		
	(5.05)	(4.39)	(-1.15)	(-2.12)	(5.35)	(3.58)	(-1.91)	(-2.43)		
Controls	NO	YES	NO	YES	NO	YES	NO	YES		
R^{2} (%)	25.45	38.73	25.19	28.71	24.86	38.04	24.64	28.32		
Obs.	14,004	14,004	14,004	14,004	13,779	13,779	13,779	13,779		
Panel B: Monthly frequency										
NSRI % change	3.117	1.713	1.495	1.021	4.804	2.384	1.673	1.100		
	(7.11)	(9.59)	(3.53)	(3.47)	(9.08)	(8.62)	(3.87)	(3.27)		
Controls	NO	YES	NO	YES	NO	YES	NO	YES		
R ² (%)	10.81	46.33	8.66	18.99	14.67	47.64	8.97	19.27		
Obs.	3,015	3,015	3,015	3,015	3,073	3,073	3,073	3,073		

$\Delta CDS_{j,t} = \beta_0 + \beta_1 \Delta NSRI_{j,t-\tau} + \gamma \mathbf{X} + \xi_j + \lambda_t + \epsilon_{j,t}$

X includes lag of $\Delta CDS_{j,t}$, country-specific news sentiment, stock market return, and volatility; it also includes global macro-financial variables – the world stock market return, implied volatility index (VIX), US economic policy uncertainty Index (EPU), and US economic activity index (ADS)

NSRI is significantly associated with CDS spread contemporaneously and predicts future CDS spread dynamics



NSRI and sovereign rating downgrades

		(1)	(2)	(3)	(4)	(5)	(6)
NSR	1	0.114	0.115	0.092	0.092	0.080	0.088
		(6.51)	(6.63)	(4.31)	(4.31)	(3.78)	(3.76)
Sent	iment		0.051	0.048	0.047	0.037	0.092
			(3.12)	(1.80)	(1.77)	(1.38)	(3.60)
Ret				-0.017	-0.023	-0.025	-0.044
				(-1.45)	(-1.23)	(-1.44)	(-2.62)
Vol				0.155	0.158	0.115	0.119
				(9.46)	(8.95)	(5.35)	(3.98)
Wor	ld Ret				0.011	0.031	0.066
					(0.48)	(1.48)	(4.08)
EPU						0.041	0.044
						(2.08)	(1.54)
ADS						0.010	0.012
						(0.57)	(0.65)
VIX						0.056	0.051
						(1.97)	(1.20)
CDS	Spread						0.008
- 2							(0.46)
R^2 (%)	16.5	17.3	27.5	27.5	28.9	33.5
Obs.		2,605	2,605	1,716	1,716	1,716	1,074

 $Downgrade_{j,t} = \beta_0 + \beta_1 NSRI_{j,t-\tau} + \gamma \mathbf{X} + \xi_j + \epsilon_{j,t}$

One SD \uparrow in average NSRI over the 30-day window ending t - 2 predicts \uparrow in the probability of downgrade by 15–22% relative to the fraction of downgrades in the data

NSRI complements CDS spread; useful as a stand-alone high-frequency default risk proxy



NSRI and equity markets I

 $Ret_{j,t} = \beta_0 + \beta_1 \Delta NSRI_{j,t-\tau} + \gamma \mathbf{X} + \xi_j + \lambda_t + \epsilon_{j,t}$

	Panel A: Dollar index return					Panel B: Local-currency index return					
	Contemporaneous		Lag	agged Contemporaneous				Lagged			
	(1)	(2)	(3)	(4)		(5)	(6)		(7)	(8)	
ΔNSRI	-9.316	-7.120	7.560	6.443		-6.111	-3.978		6.294	5.466	
Controls	(-4.25) NO	(-3.50) YES	(2.75) NO	(2.39) YES		(-3.14) NO	(-2.31) YES		(2.70) NO	(2.37) YES	
R ² (%) Obs.	11.74 24,101	36.78 24,101	11.72 24,101	13.40 24,101		10.03 24,004	32.50 24,004		10.03 24,004	11.61 24,004	

 $\textit{Ret}_{j,t} = \beta_0 + \beta_1 \Delta \textit{NSRI}_{j,t-\tau} + \beta_2 \Delta \textit{NSRI}_{j,t-\tau} \times \textit{HighFiscalConstraint}_{j,t}$

+ HighFiscalConstraint_{j,t} + γX + ξ_j + λ_t + $\epsilon_{j,t}$

	Contemp	oraneous	Lag	ged
Dollar index return	(1)	(2)	(3)	(4)
ΔNSRI	-3.078	-1.651	12.023	11.863
	(-1.11)	(-0.68)	(2.61)	(2.61)
Δ NSRI × High Fiscal Constr.	-10.005	-9.098	-8.586	-9.958
	(-2.48)	(-2.44)	(-1.74)	(-2.04)
High Fiscal Constr.	8.221	6.912	9.242	10.431
	(1.81)	(1.55)	(2.05)	(2.13)
Controls	NO	YES	NO	YES
R^{2} (%)	11.74	36.73	11.72	13.41
Obs.	23,862	23,862	23,862	23,862

X includes lag of $Ret_{j,t}$ country-specific news sentiment, and stock market volatility. Also includes global macro-financial variables – the world stock market return, implied volatility index (VIX), US economic policy uncertainty Index (EPU), and US economic activity index

 $Ret_{j,t}$ is in basis points

Continuous right hand side variables are normalized to unit variance



NSRI and equity markets II

	Panel A: Dollar index return					Panel B: Local-currency index return					
	Contemporaneous		Lag	Lagged		Contemporaneous			Lagged		
	(1)	(2)	(3)	(4)		(5)	(6)		(7)	(8)	
ΔNSRI	-7.093	-9.908	5.287	5.138		-3.770	-5.765		4.129	4.350	
	(-3.07)	(-4.23)	(1.86)	(1.80)		(-1.90)	(-2.89)		(1.72)	(1.80)	
∆GNSRI	-18.900	-21.554	19.640	16.683		-19.914	-21.741		18.551	15.976	
	(-6.53)	(-7.93)	(6.45)	(5.71)		(-7.41)	(-8.24)		(6.21)	(5.66)	
Ret(t-1)		-41.845		-39.172			-33.777			-31.676	
		(-6.22)		(-5.96)			(-5.36)			(-5.15)	
Sentiment		-18.085		10.578			-13.709			9.520	
		(-3.94)		(2.85)			(-3.25)			(2.70)	
R ² (%)	11.88	13.27	11.87	13.06		10.23	11.45		10.20	11.26	
Obs.	24,101	24,101	24,101	24,101		24,004	24,004		24,004	24,004	

- ΔGNSRI is the percentage change of the Global NSRI. Global NSRI is computed as the cross-sectional weighted average of NSRI using last year-end's equity market capitalization as weights
- Investors tend to respond more to default concerns that are more global in nature compared to country-specific concerns global (push) factors are crucial for local asset prices



Conclusion

- We develop a **novel framework for quantifying sovereign default risk in real-time** using big data and natural language processing techniques
- We show that **news media narratives contain valuable information about countries' riskiness** useful for tracking default concerns in real-time
- The NSRI is informative about sovereign CDS spread dynamics and future credit rating downgrades
 - Useful as a high-frequency default risk proxy, especially where daily market-based data or other proxies are not available
 - The NSRI can be integrated as an important early warning indicator of any sovereign risk monitoring framework
- The NSRI allows the measurement of both global risk episodes, such as the Covid-19 pandemic, the GFC and the war in Ukraine, and country-specific risks (see Annex, if needed)



ANNEX

Details and supporting information



Case study I: NSRI and Covid-19



Note: All data 3-month moving averages, last observation 28.03.2022 Source: Authors' calculations, Refinitiv Datastream

- The NSRI quickly reacted to the Covid-19 pandemic, showing a substantial increase in China and other countries from the onset of the crisis
- In Europe, the NSRI responded more rapidly in Italy, the first country of significant transmission, than e.g. Spain – NSRI is different across countries
- The NSRI reasonably tracks or leads the evolution of sovereign CDS spreads at the height of the pandemic



24

06/07/2022



Case study I: NSRI and Covid-19 (ctd)

NSRI during Covid-19 – differences between selected advanced and developing economies



Note: *: Canada, France, Germany, Italy, Japan, UK, US **: here: Angola, Ghana, Kenya, Nigeria, South Africa 3-month MA over a four-day window Source: Authors' calculations

NSRI – Top 20 increases during Covid-19 outbreak (Nov/Dec 2019 - Mar/Apr 2020), % change



Note: Out of 99 countries, 3-month MA over a four-day window Source: Authors' calculations

Case study I: NSRI and Covid-19 (ctd)

Share of countries with an NSRI that is higher/lower at end-2021 compared to December 2019, %

The pandemic has a continuing/structural impact on news media's perception of sovereign risk On average, the NSRI remains more elevated compared to pre-pandemic levels

- Persistent concerns about sovereign solvency as a
- result of increased Covid-19 public spending
- Newspapers more keen on broadcasting sovereign distress topics (the need to recalibrate the model to the new steady state?)



NSRI more than 25% higher
NSRI up to 25% higher
NSRI unchanged or up to 25% lower
NSRI more than 25% lower

Note: Out of 106 countries, 3-month MA over a four-day window Source: Authors' calculations



Case study II: NSRI and Russia's invasion of Ukraine



NSRI - Ukraine

NSRI - Russia



Note: All data 3-month moving averages, last observation 27.03.2022 Source: Authors' calculations, Refinitiv Datastream

- NSRI sends clear warning signals weeks before Russia's invasion of Ukraine on 24 February 2022 for both countries
- The NSRI remains muted compared to CDS spreads and the index's reaction to the 2008 financial crisis
 - Possible explanation: NSRI focuses on news on sovereign risk while the news coverage of the war is centred more on political developments with occasional warnings of rating downgrades and Russian default