

Big Data for Policy-Making in Central Bank: Measuring Economic Policy Uncertainty in Indonesia

Yati Kurniati

Head of Statistics Department Bank Indonesia

MyStats 2018 "Embracing Data Science and Analytics to Strengthen Evidence-Based Decision Making" Kuala Lumpur, 19 November 2018

OUTLINE

Big Data in Bank Indonesia



Measuring Indonesian Economic Policy Uncertainties Index







Technology

- Internet & social media
- Electronic devices
- Digitalization/electronification (e.g. e-commerce, FinTech, online portals)



Paradigm

- Analysis of aggregated data \rightarrow granular
- Descriptive analysis \rightarrow predictive \rightarrow prescriptive



Methods

- Data & text mining
- Image processing
- Network analysis
- Data visualization, etc

made possible by distributed & GPU computing

Digital services have touched on all aspects of our dayto-day activities, influencing how we make decisions, how we interact with each other, and giving life to innovative and more efficient business models...

(IoT)

Disruptive Technologies



Smart Phone Internet of Things



Big Data Analytics

The number of Internet and social media users in Indonesia grew extremely fast, supported by easier access to smart phones and Internet services.





BIG DATA is one source of data for STATISTICS

There are needs for data that cannot be timely fulfilled through conventional data sources, e.g. regulatory reporting, surveys, and secondary data

PRIMARY DATA

e.g.:

- 1. Bank & Nonbank Reporting
- 2. Transactional Data – Payment System, Money Market Settlement
- 3. Survey

STATISTICS

Big Data : Structured + Unstructured

SECONDARY DATA

External Sources

- 1. Administrative Data (from other government institutions)
- 2. Other Sources, e.g.:
 - News & social media
 - FinTech & e-commerce
 - Online portals
 - Internet search data
 - Satellite images
 - Sensor data



OBJECTIVE

Empowering decision-making in monetary, financial system stability, and payment system sectors

BIG DATA STRATEGIC ADVANTAGES			
Data Lag / Data Gap	Addressing the gap in data lag by providing new data/indicators in a more timely manner		
Market and Investor Behavior	Analyzing and predicting behavior in financial markets		
Network Analysis	Analyzing interconnectedness between economic agents, including for monitoring systemic risk		
Public Perception	Measuring public perception and expectation on Bank Indonesia's policies		
1.2			

5





Market/Investor Behavior E-commerce data analytics

Early warning indicators of instability in financial system from news data

Network Analysis Bank interconnectedness in payment system

Corporation interconnectedness in payment system

Public Perception/Expectation

Measuring stakeholder's perception and expectation on Bank Indonesia's policies



2. Measuring Indonesian Economic Policy Uncertainty (EPU) Index

2. Economic Policy Uncertainty (EPU) Index

Indonesian EPU Index has been developed to provide policy maker with indicator that measures economic policy uncertainties. This project is based on the research *Measuring Economic Policy Uncertainty* (Baker, Bloom, Davis, 2016). From newspaper articles we can extract proxy indicator of uncertainty on economic policy.





FRAMEWORK Indonesian EPU Keyword Selection



- Human annotation was conducted on 2,220
 newspaper articles
- From there, we obtained 113 Economic, 169 Policy, and 142 Uncertainty keywords
- The best combination of Economic, Policy, and Uncertainty keywords were selected using Genetic Algorithm*)
- Result: 10 Economic, 39 Policy, and 47
 Uncertainty keywords

*) Genetic Algorithm is a heuristic for mathematical optimization, loosely inspired by evolution theory and genetics



FRAMEWORK Examples of Indonesian EPU keywords

Есопоту		Policy		Uncertainty	
Keyword	Translation	Keyword	Translation	Keyword	Translation
Ekonomi Inflasi Keuangan Ekspor Nilai tukar rupiah	Economy Inflation Finance Export Exchange rate	Pemerintah Kebijakan pemerintah APBN Kebijakan moneter DPR Kenaikan BI Rate Penurunan BI Rate	Government Policy Government budget Monetary policy House of Representatives BI rate hike BI rate cut	Ketidakpastian Dikhawatirkan Mengkhawatirkan Kepanikan Kesulitan Terus tertekan Inkonsistensi	Uncertainty Concern Concern Panic Hardship Downward pressure Inconsistency



FRAMEWORK Filtering articles that contain EPU

Developing the model for filtering article

2

- Model for filtering article is developed based on annotated data. Each article is classified as containing EPU if it contains <u>at least one keyword from each of the list of economic, policy, and</u> <u>uncertainty keywords</u>
- b. To improve accuracy, we added constraints s.t.:
 - Uncertainty keyword cannot appear in the same sentence as a negation keyword, e.g.: "the condition is not of particular concern"
 - Distance between economic, policy, and uncertainty keywords must not exceed some number of sentences

The model is applied on all articles from 2001 to classify each article into EPU/non EPU



FRAMEWORK Constructing EPU Index

The Indonesian EPU index is constructed based on newspaper coverage frequency:

- Dividing the EPU count for each newspaper and month by the count of all articles in the same newspaper and month.
- Normalizing each newspaper's scaled count to unit standard deviation, then sum over all the papers (there are 4 papers in our case) by month to get the EPU monthly index

1
$$X_{it} = \frac{\#epu_article_{it}}{\#article_{it}}$$

2 $Y_{it} = \frac{X_{it}}{\sigma_i}$
3 $Z_t = \frac{\sum Y_{it}}{\#_i}$
4 $EPU index = \frac{z*100}{\mu_z}$



RESULT

Indonesian EPU index fluctuates more since 2015 compared to previous periods. In general, the fluctuations are driven by external factors, e.g. FFR hike, and Brexit.





RESULT & ANALYSIS

Comparison with other uncertainty indicators

- Indonesian and US EPU index move fairly hand-in-hand before 2008 and after 2012
- During 2008 to 2012, US EPU index went to very high levels while Indonesian EPU index was in a decline since 2008 GFC → indicates that GFC did not significantly impact Indonesian economy.



EVALUATING EPU INDEX Comparison with other uncertainty indicators

Beginning 2018, Indonesia 10Y yield, CDS, and NDF-spot gap rose, indicating a general increase in uncertainty. The EPU index also increased during the same period.



15

FUTURE WORKS





CONCLUDING REMARKS

- Today's digital technologies have resulted in data being produced in massive amounts, in real-time, in a variety of formats, by various institutions and individuals. Extracting relevant information from these sources is not straightforward and will require a distinct set of skills. Data science and Analytics are essential in uncovering the hidden insights in Big Data.
- Bank Indonesia have realized important benefits from data science and analytics in these 4 areas: addressing the data gap/data lag, measuring market and investor behavior, network analysis, and measuring public perception.
- An example of data science and analytics in Bank Indonesia is the Indonesian EPU index. The index, which is constructed from newspaper articles, is able to capture periods with high uncertainty, e.g. following FFR hike. The index is also shown to move in conjunction with other uncertainty indicators.

<u>Thank You</u> Terima Kasih

FRAMEWORK





Sumber Data:

CyberLibrary

- Period:
 - Jan 2001 Sept 2018
- Source:
 - Kompas
 - Bisnis Indonesia
 - Media Indonesia
 - Republika
- # articles: 150.000

Cleansing

1. <u>HTML parsing</u> Example HTML tag: <P class=BeritaBaca>PARIS (AFP/Rtr): Kelompok kreditor Paris Club sedang mempelajari penangguhan pembayaran utang negara-negara yang dilanda bencana tsunami.</P><!--AD--> <P

class=BeritaBaca>"

2. Cleansing source media name

Using string similarity algorithm

Source Media:

Bisins Indonesia, Bisnis Indonersia, Bisnia Indonesiaa → Bisnis Indonesia

Compas, Kompass -> Kompas



1. Filter EPU articles

An article is classified as EPU if it contains at least one keyword from each of the list of economic, policy, and uncertainty keywords. **Keyword:**

- 1. Economic: ekonomi, keuangan, industri
- 2. Policy: kebijakan, pemerintah, Bank Indonesia
- Uncertainty: Tidak pasti, ketidakpastian, tidak jelas, resiko, risiko, mengkhawatirkan, tidak menentu

Processing

2. Calculating EPU index

Indexing: $X_{it} = \frac{\#epu_article_{it}}{\#article_{it}}^{*}$ $Y_{it} = \frac{X_{it}}{\sigma_i}$ $Z_t = \frac{\sum Y_{it}}{\#_i}$ EPU index= $\frac{z*100}{\mu_Z}$

*) i = media (Kompas, Bisnis Indonesia, dst) t = period



FRAMEWORK



cion	Actual (human judgment)				
edict		EPU	non EPU		
el pre	EPU	ТР	FP (Type I error)		
Mod	non EPU	FN (Type error II)	TN		

Metric

Precision

Recall

F0.5

udgment)		TP	True Positive	Number of EPU articles which are correctly predicted as EPU		
	non EPU		TN	True Negative	Number of non-EPU articles which are correctly predicted as non-EPU	
	FP (Type I error)	(A)	FP	False Positive	Number of non-EPU articles which are falsely predicted as EPU	
I)	TN		FN	False Negative	Number of EPU articles which are falsely predicted as non-EPU	
	Formula	1		Intor	prototion	
Formula			interpretation			
TP / (TP + FP)% actual EPU articles from all art predicted as EPU by the model		es from all articles / the model				
TP / (TP + FN)		8 E	% predicted EPU articles from all actual EPU articles in the data			
$(1+(0.5^2))$ PR / $((0.5^2P) + R)$		H 2:	Harmonic mean of precision and recall, with 2x more weight given to precision			

20