Reading China:

Predicting Policy Change with Machine Learning

Julian TszKin Chan   Weifeng Zhong

(Bates White)   (AEI)

November 17, 2018

MAPOR annual conference

Policy Change Index (PCI) for China

The first *leading* indicator of China's policy changes.

- Covers 1951 Q1 – 2018 Q3.
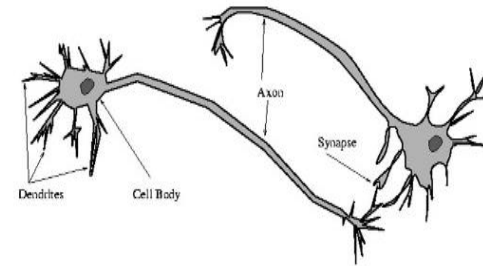
- Can be updated in the future.

How to predict policy changes?

Build a machine learning algorithm to

- "read" the *People's Daily*;

- detect changes in how it prioritizes policy issues.

Official newspaper, 1946-2018   (Artificial) neural networks

Source of predictive power

People's Daily:
nerve center of China's propaganda system

+

Propaganda often precedes policies.

↓

Detect changes in newspaper's priorities   **≈**   Predict changes in gov't policies

Source of predictive power

People's Daily:
nerve center of China's propaganda system

+

Propaganda often precedes policies.

↓

Detect changes in newspaper's priorities

≈

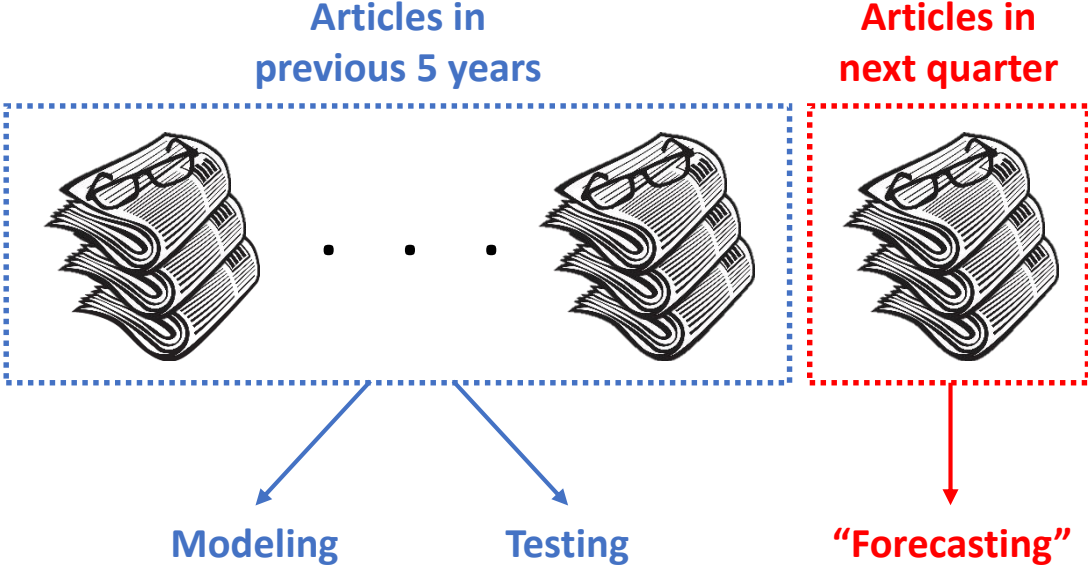Predict changes in gov't policies

Front page?

Methodology

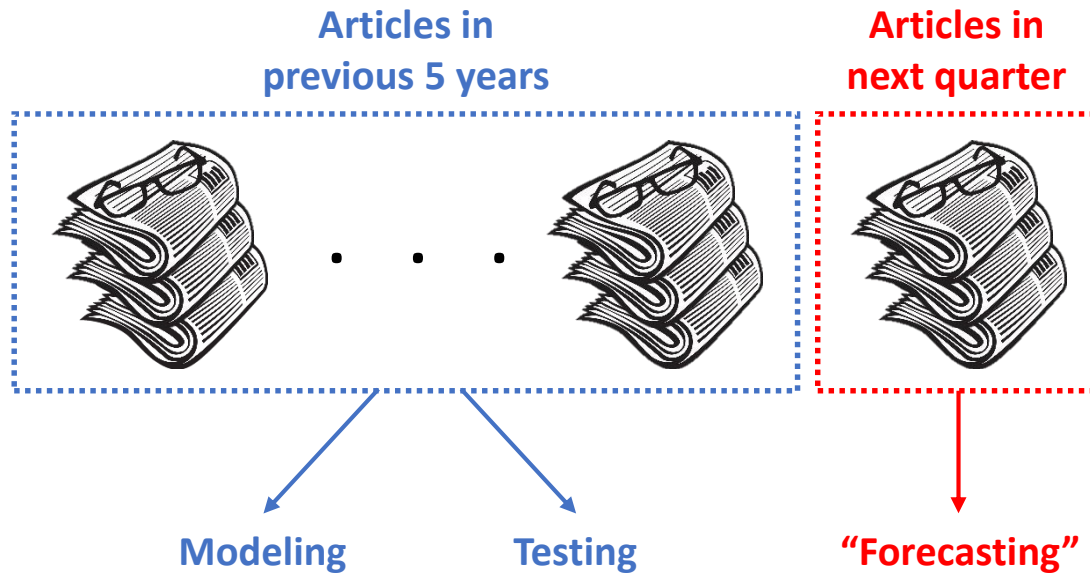Imagine an avid reader of the *People's Daily* who

1. reads recent articles;

2. forms a paradigm about front-page content;

3. tests the paradigm on new articles.

Methodology

Articles in previous 5 years   **Articles in next quarter**

Modeling   **Testing**   **"Forecasting"**

$$\text{PCI} = \left| \text{Testing performance} - \text{"Forecasting" performance} \right|$$

Methodology

Articles in previous 5 years   Articles in next quarter

Modeling   Testing   "Forecasting"
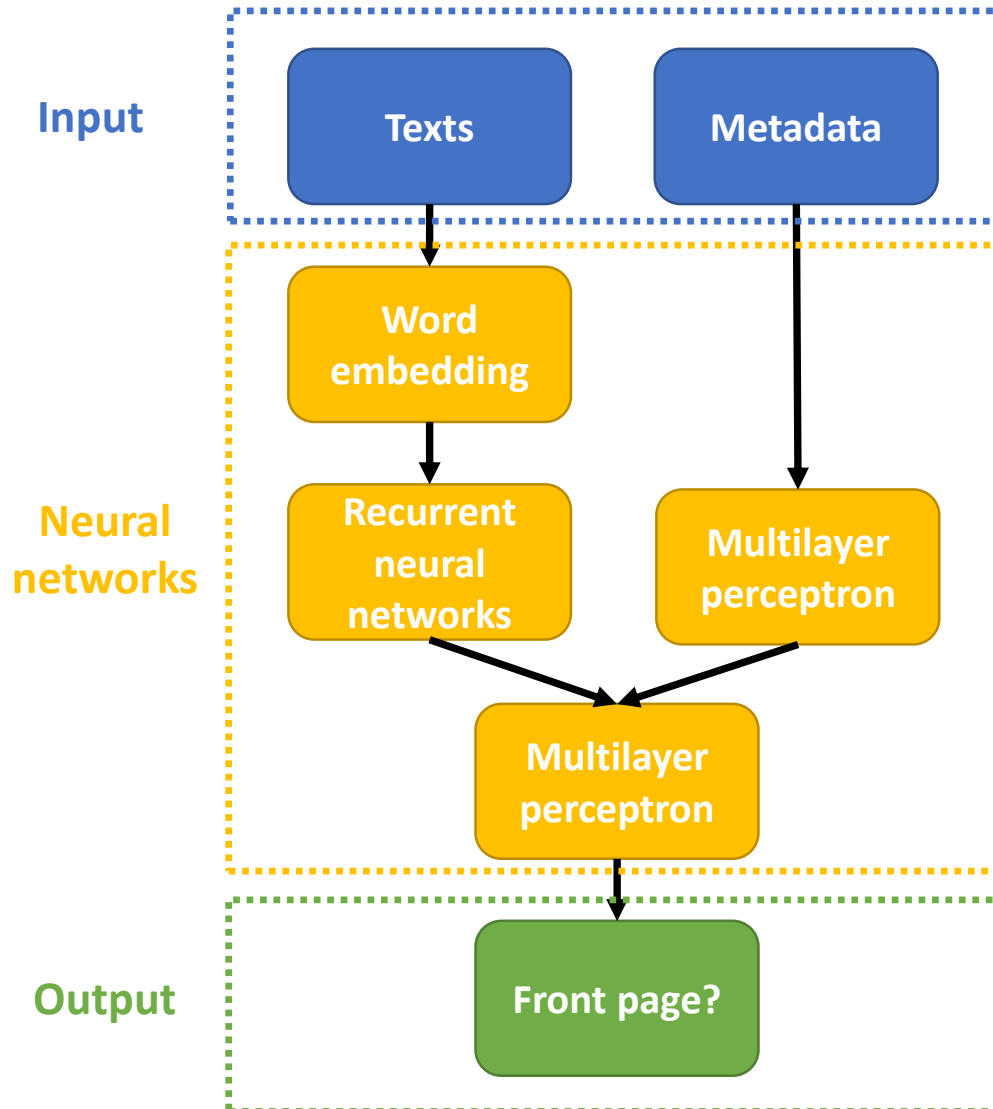
$$\text{PCI} = \left| \text{Testing performance} - \text{"Forecasting" performance} \right|$$

*"Language-free" — it does **not** require the reading of the Chinese text.*
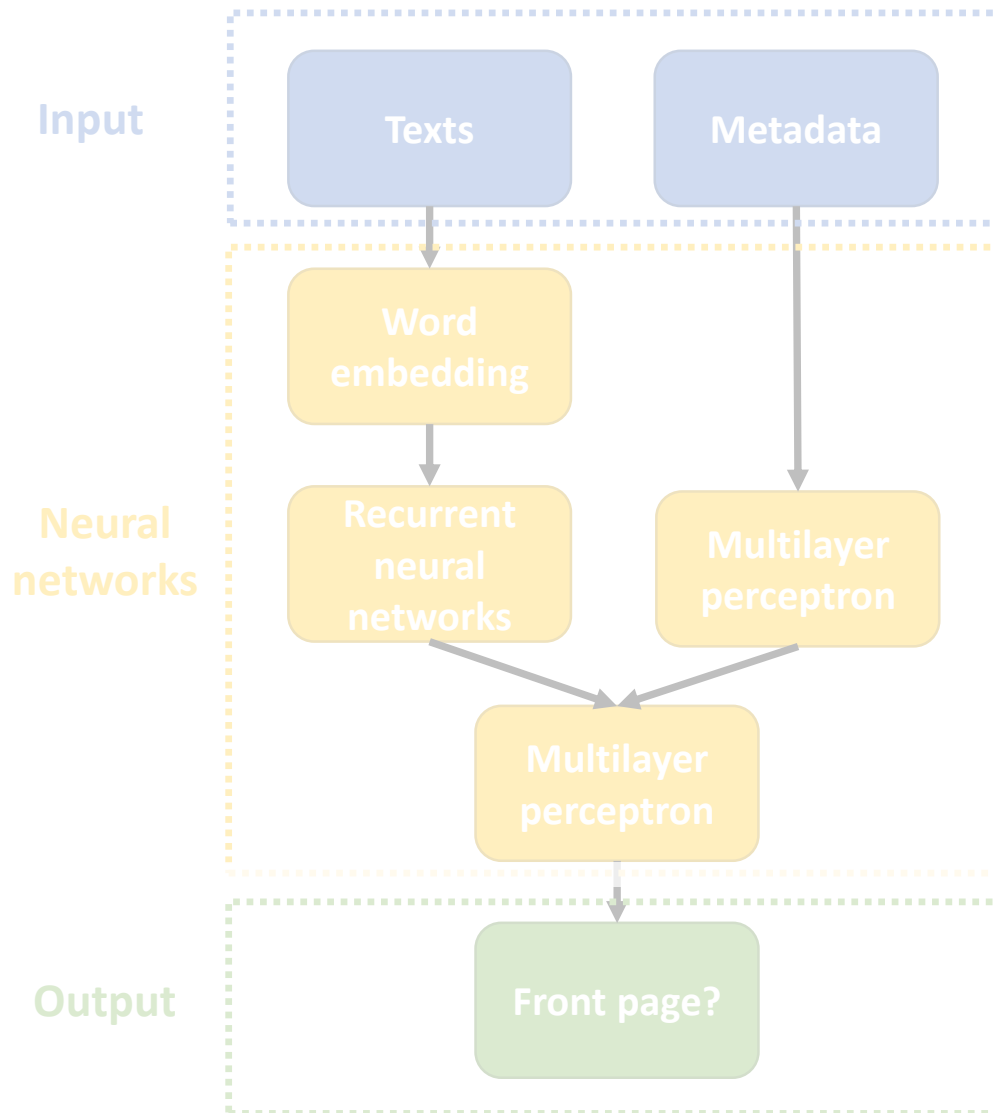
Methodology: modelling

Input

Texts   Metadata

x : each article as an observation.

Neural networks

Word embedding

Recurrent neural networks   Multilayer perceptron

Multilayer perceptron

f : a complicated function.

Output

Front page?

$y = f(x)$

Methodology: modelling

Input

Texts   Metadata

x : each article as an observation.

Neural networks

Word embedding

Recurrent neural networks   Multilayer perceptron

Multilayer perceptron

f : a complicated function.

Output

Front page?
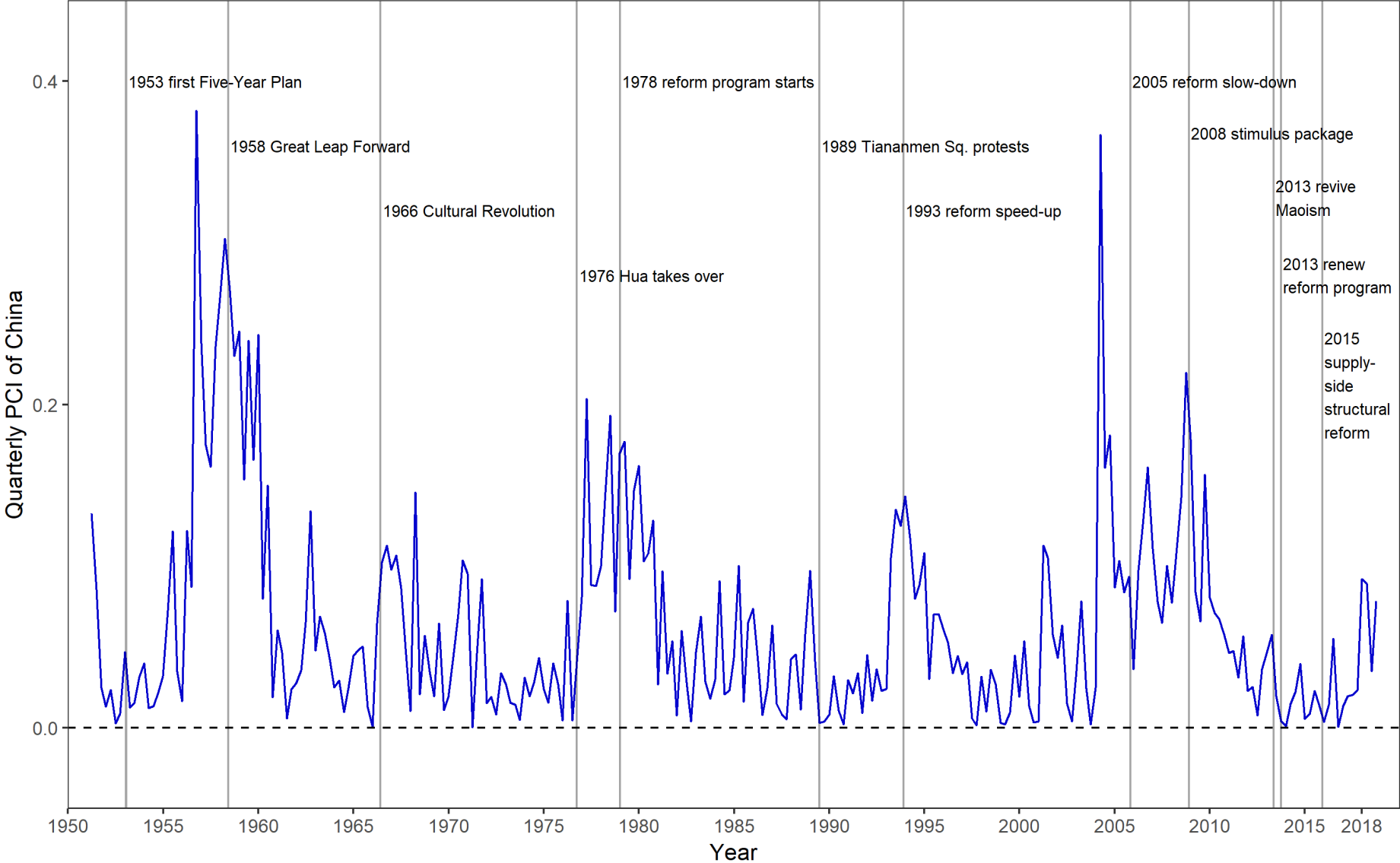
$y = f(x)$
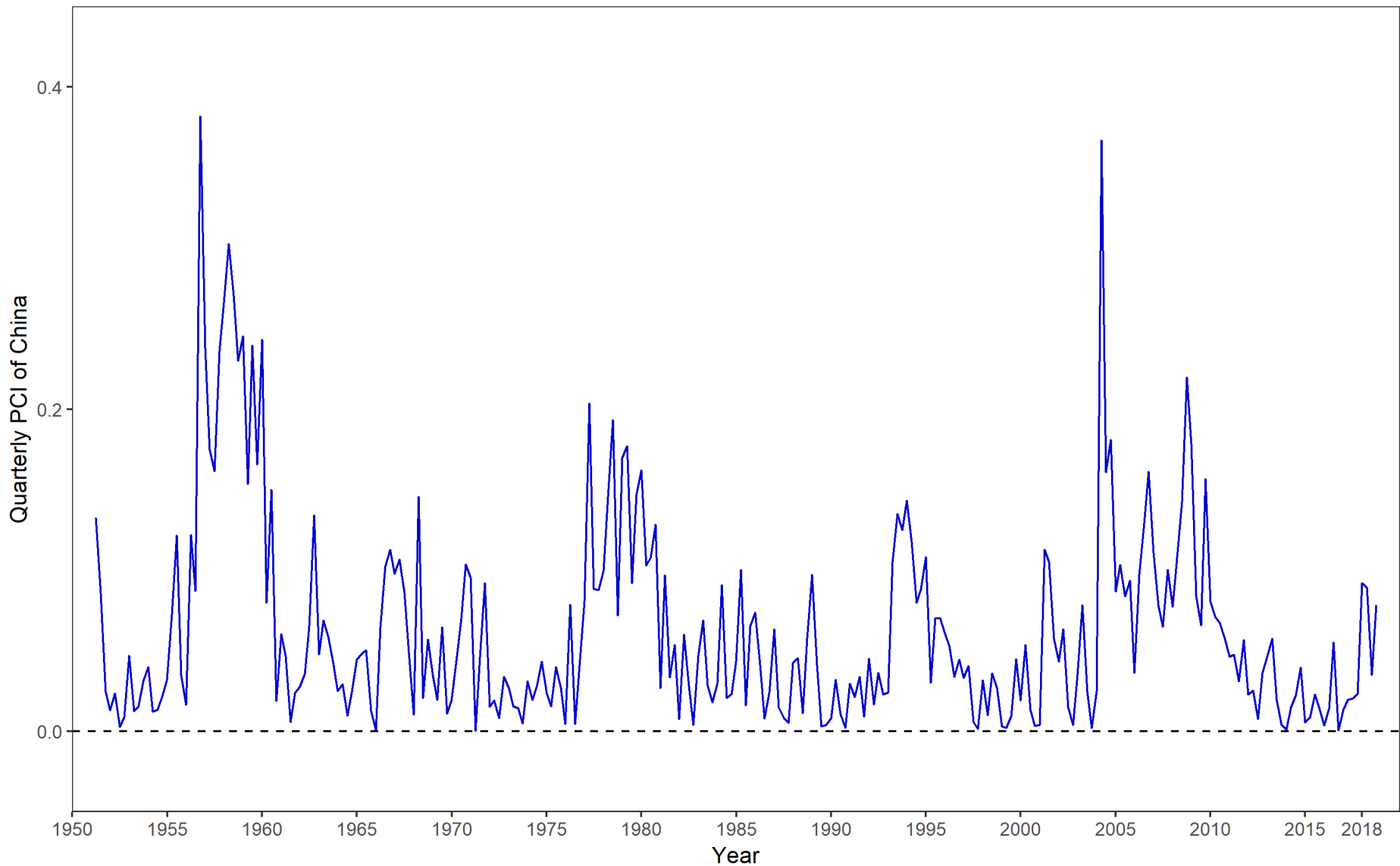
Results

Result: PCI

Result: PCI — with ground truth

Understanding substance of change

		Classified on front page?	
		No	Yes
Front page?	No	√	false positives
	Yes	false negatives	√

- Content of *mis*-classified articles has policy substance.

Potential applications

App 1: PCIs for other countries

Predicting other (ex-)Communist regimes' policies:

- USSR's *Pravda*

- GDR (East Germany)'s *Neues Deutschland*

- DPRK (North Korea)'s *Rodong Sinmun*

- Cuba's *Granma*

- Vietnam's *Nhân Dân*

App 1: PCIs for other countries

Predicting other (ex-)Communist regimes' policies:

- USSR's *Pravda*

- GDR (East Germany)'s *Neues Deutschland*

- DPRK (North Korea)'s *Rodong Sinmun*

- Cuba's *Granma*

- Vietnam's *Nhân Dân*

PCI Cuba is in the works!

App 2: measuring media bias

Measuring media bias in the US:

- Replicate the same analysis on US newspapers.

- Compare $"PCI"_{WaPo}$, $"PCI"_{NYT}$, $"PCI"_{WSJ}, \ldots$

- Divergence among $"PCI"$s $\dashrightarrow$ Polarization in media

Other apps

Omitted here. See our research paper.

Takeaways

- A "crystal ball" to predict China's next moves.

- A new way to uncover hidden patterns from trivial labels.

- A wide range of potential applications.

Interested in DIY?

- Website: policychangeindex.com (newsletter sign-up)

- Research paper: policychangeindex.com/Reading_China.pdf

- Source code: github.com/open-source-economics/PCI

Contact us:

- Julian TszKin Chan: ctszkin@gmail.com

- Weifeng Zhong: weifeng@weifengzhong.com

Questions?