

Reading China:

Predicting Policy Change with Machine Learning

Julian TszKin Chan (Bates White Economic Consulting)

Weifeng Zhong (Mercatus Center at George Mason University)

EcoMod 2019 (Ponta Delgada, Portugal)

The views expressed here are the authors' and do not represent the views of Bates White, the Mercatus Center, or their other employees.

子亮相

人民日报

RENMIN RIBAO

（人 員 與 機 械）

习近平应约同美国总统特朗普通电话

特朗普祝贺中共十九大胜利召开，祝贺习近平再次当选中共中央总书记。



據悉，該公司負責人表示，該公司目前正積極籌集資金，並計劃在未來幾個月內，將該公司重新組建，並恢復正常生產。該公司負責人表示，該公司目前正積極籌集資金，並計劃在未來幾個月內，將該公司重新組建，並恢復正常生產。

一、中央政治局委员

丁其昌	丁其昌	丁其昌	丁其昌	丁其昌	丁其昌
李其昌	李其昌	李其昌	李其昌	李其昌	李其昌
李其昌	李其昌	李其昌	李其昌	李其昌	李其昌
李其昌	李其昌	李其昌	李其昌	李其昌	李其昌
李其昌	李其昌	李其昌	李其昌	李其昌	李其昌

二、中山政出兩堂多難期全案

三、中央委员候补书记

四、中央书记处书记

[illegible][illegible][illegible]

地址：北京市西城区西便门大街2号
邮编：100031 电话：010-63792111

[illegible]

1997年12月1日

[illegible]

100

Figure 1

Figure 1

100

10



100



会议并作重要讲话
正在中央政治局常委
中央军委主席
主持机构



证券时报

中国

State-run media

- Example: the *People's Daily*.
 - Official newspaper of the Communist Party of China;
 - Toed by other media in the country;
 - ...
 - The *Pravda* of China.
- The primary mission is not to report news.

State-run media tell us a lot — about the state!

- Example: the *People's Daily*.
 - Official newspaper of the Communist Party of China;
 - Toed by other media in the country;
 - ...
 - The *Pravda* of China.
- The primary mission is not to report news... but to propagate official viewpoints and signal policies.

Machine learning can help decipher propaganda

- Deciphering propaganda content has been complicated due to
 - Opaque system;
 - Manipulated data;
 - Difficulty in validation;...
- Machine learning — with some twists — can help!
- The method has wide applicability in other contexts.

Policy Change Index (PCI) project

- First use case: China.
- Data: *People's Daily* text, 1946-present.
- Use neural networks to detect changes in its priorities.
- Goal: predict policy change based solely on the newspaper.
- Output: the **Policy Change Index (PCI)** for China.



Source of predictive power

Constructing the PCI

Results

Source of predictive power



Source of predictive power

- The Soviet tradition:
 - “[T]he whole task of the Communists is to be able to **convince** the backward elements.” (Lenin, 1920)
 - It is a fundamental necessity “to transform the press... into a serious organ for the **economic education** of the mass of the population.” (Lenin, in early USSR years)
 - Even the film was considered “a great and invaluable force... aiding the working class and its Party to **educate** the toilers in the spirit of socialism.” (Stalin, 1935)

Source of predictive power

- The Soviet tradition, in China:
 - “Newspapers **must** be run by politicians.” (Mao, 1957)
 - The media should “uphold the **correct orientation** of public opinion in all aspects and at all stages.” (Xi, 2016)



Source of predictive power

People's Daily:
central to China's propaganda system

+

Propaganda often precedes policies.

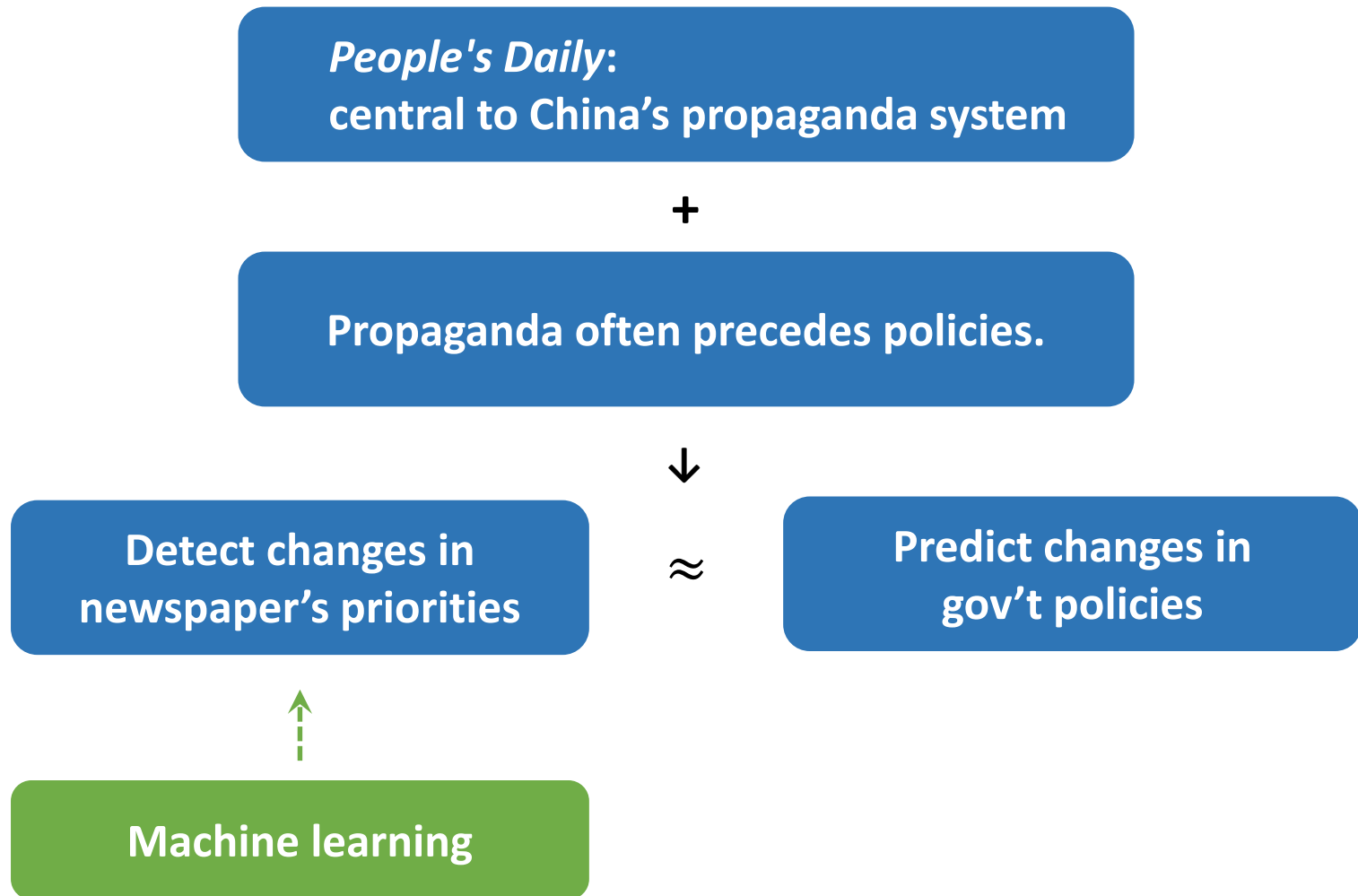
↓

Detect changes in
newspaper's priorities

≈

Predict changes in
gov't policies

Source of predictive power



Source of predictive power

Constructing the PCI

Results

Start with something *seemingly* trivial

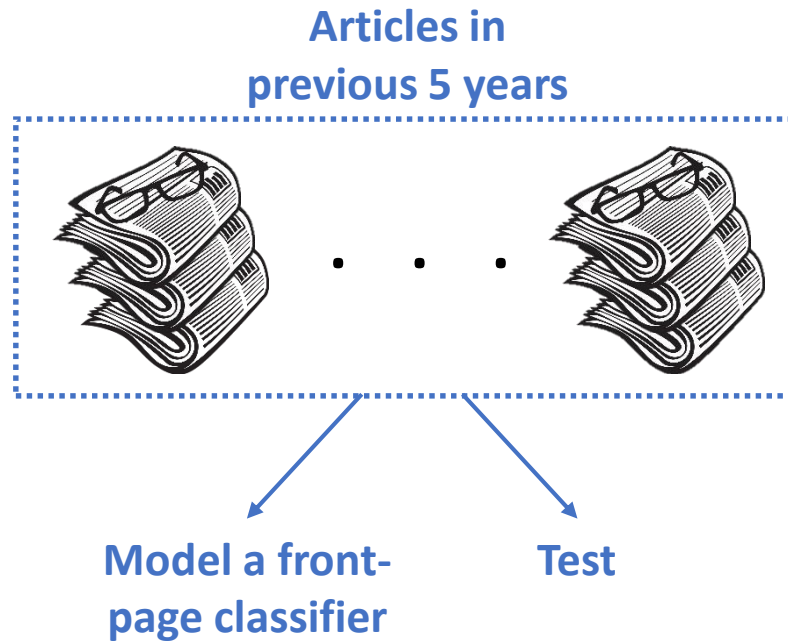
Build an algorithm that classifies **front-page** articles
("reverse-engineering" the editor's mind).

Start with something *seemingly* trivial

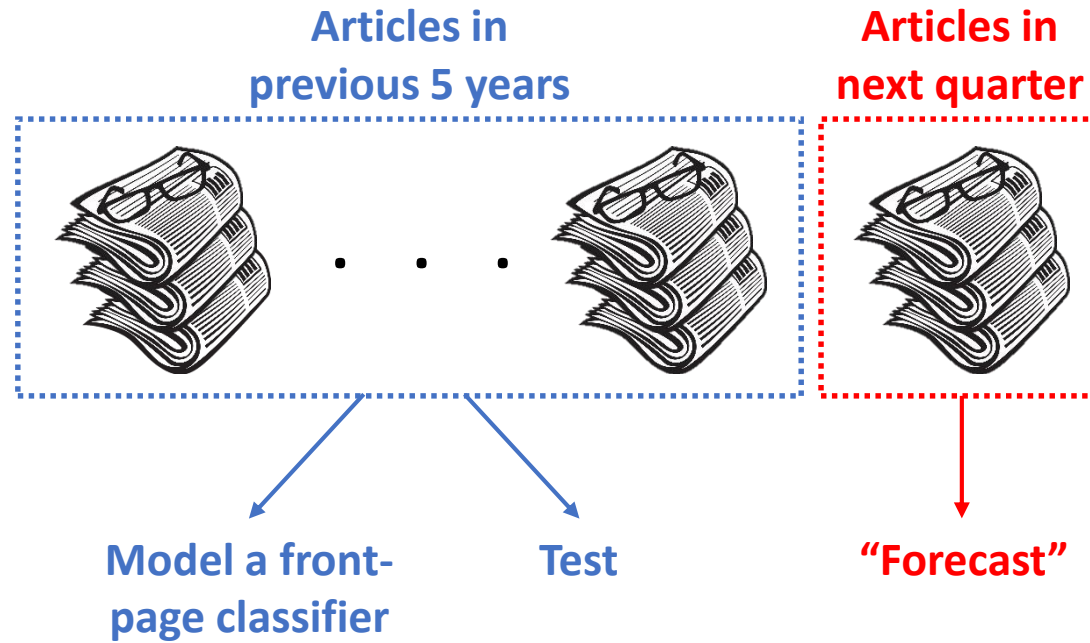
Build an algorithm that classifies **front-page** articles
("reverse-engineering" the editor's mind).

But what's the point?

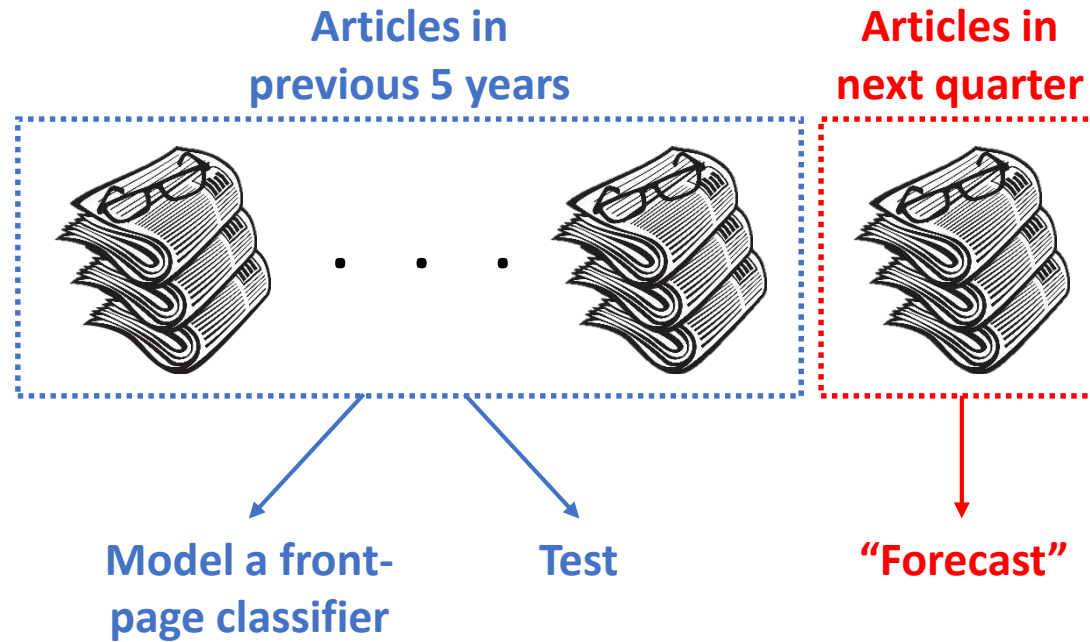
The framework



The framework



The framework



$$\text{Policy Change Index} = \left| \begin{array}{c} \text{Test} \\ \text{performance} \end{array} - \begin{array}{c} \text{“Forecast”} \\ \text{performance} \end{array} \right|$$

PCI $\gg 0 \Rightarrow$ Structural difference in the newspaper's priorities

Data structure

sample_df

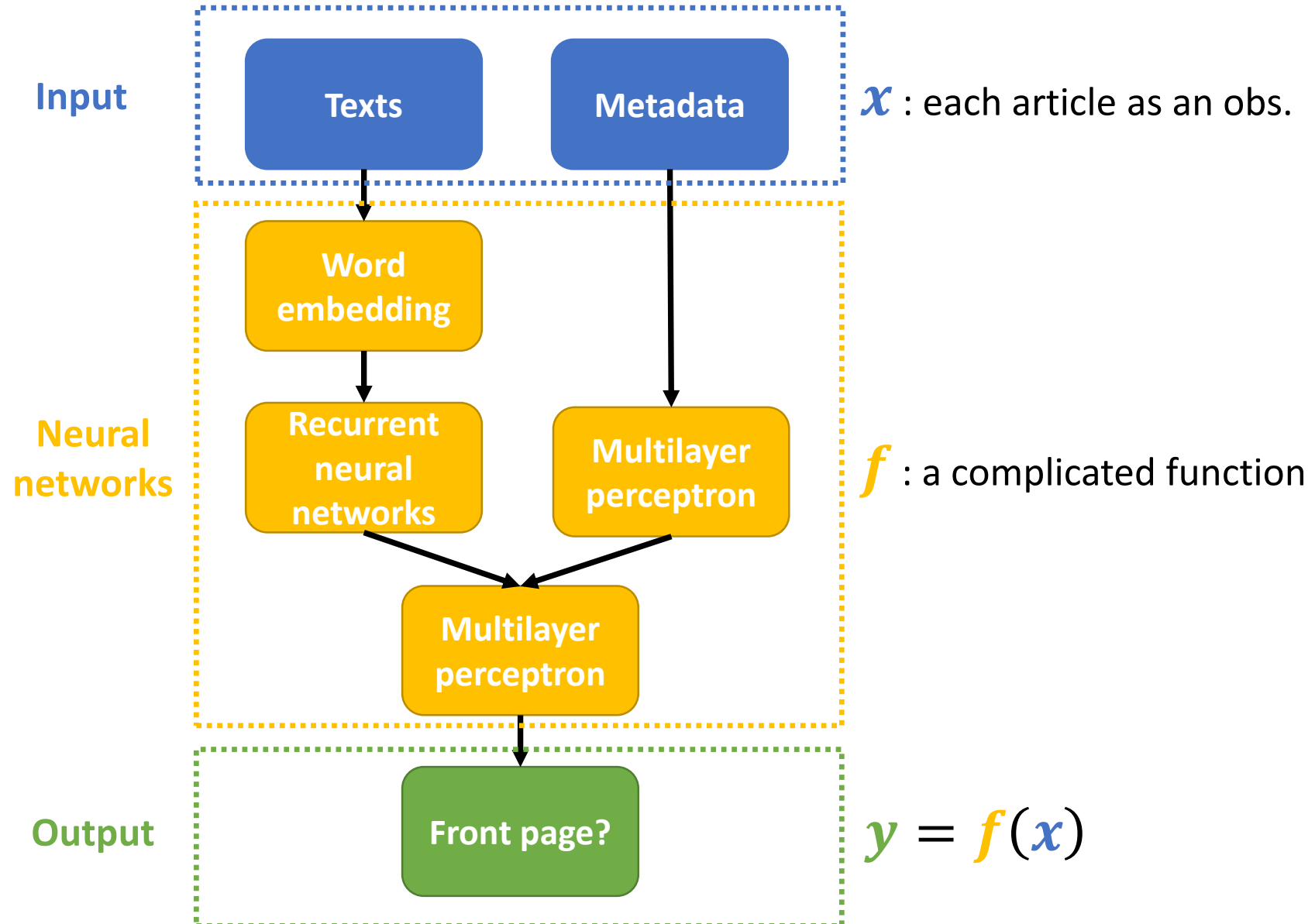
id	date	page	title	body
20181202525	2018-12-31 00:00:00	1.0	今年粮食总产量13158亿斤 农业农村发展取得新成绩，乡村振兴开局良好	本报北京12月30日电 （记者高云才、郁静娴）记者从30日召开的全国农业农村厅局长会议获...
20181202526	2018-12-31 00:00:00	1.0	新时代改革再出发的重要里程碑 ——写在党的十八届三中全会召开五周年之际	本报评论员 改革开放40年来，我们党带领人民绘就了一幅波澜壮阔、气势恢宏的历史画卷。其...
20181202527	2018-12-31 00:00:00	1.0	改革引领中国经济新航向（在习近平新时代中国特色社会主义思想指引下——新时代新作为新篇章） ...	本报记者 许志峰 吴秋余 王珂 林丽福 巨轮向前，离不开航标指引方向。 1978年...
20181202528	2018-12-31 00:00:00	1.0	产量连续六年超1200亿斤，实现“十五连丰” 黑龙江 粮食再丰收，粮仓更稳固	本报哈尔滨12月30日电 （记者方圆）此刻的黑土地，正在冬眠蓄力。今年黑龙江粮食又是一...
20181202529	2018-12-31 00:00:00	1.0	伟大的变革——庆祝改革开放40周年大型展览 现场观众达223万人次 展览将持续到2019年...	据新华社北京12月30日电 “伟大的变革——庆祝改革开放40周年大型展览”自11月13日...
20181202530	2018-12-31 00:00:00	1.0	航线不停、景区不休、生意不闲、游客不断 新疆 旅游无淡季，今冬更红火	本报乌鲁木齐12月30日电 （记者杨明方、阿尔达克）一年之中半年是降雪期，冬季雪深超过3...
20181202531	2018-12-31 00:00:00	1.0	图片报道	徜徉知识海洋，感受阅读乐趣。十二月三十日是元旦小长假的第一天，全国各地的特色文化活动丰富...
20181202532	2018-12-31 00:00:00	2.0	创造新的更大奇迹 ——党的十八届三中全会召开五周年述评	“全面发力、多点突破、蹄疾步稳、纵深推进”——庆祝改革开放40周年大会上，习近平总书记用...
20181202534	2018-12-31 00:00:00	3.0	开启中国特色大国外交新征程（2018年度国际特别报道） ——国务委员兼外交部长王毅回顾20...	本报记者 赵嘉鸣 管克江 杜尚泽 焦翔 王海林 国务委员兼外交部长王毅12月29日接...
20181202535	2018-12-31 00:00:00	3.0	二〇一七年岁末，莫斯科大学的中国留学生们欣喜地收到了习近平主席的回信；一年来，牢记习主席的...	本报驻俄罗斯记者 吴焰 “一年前，习近平主席给留学生回信，鼓励我们胸怀大志，刻苦学习...
20181202536	2018-12-31 00:00:00	3.0	外交部发言人就中美建交40周年发表谈话	本报北京12月30日电 外交部发言人陆慷30日就中美建交40周年发表谈话。 陆慷说，2...

Data structure

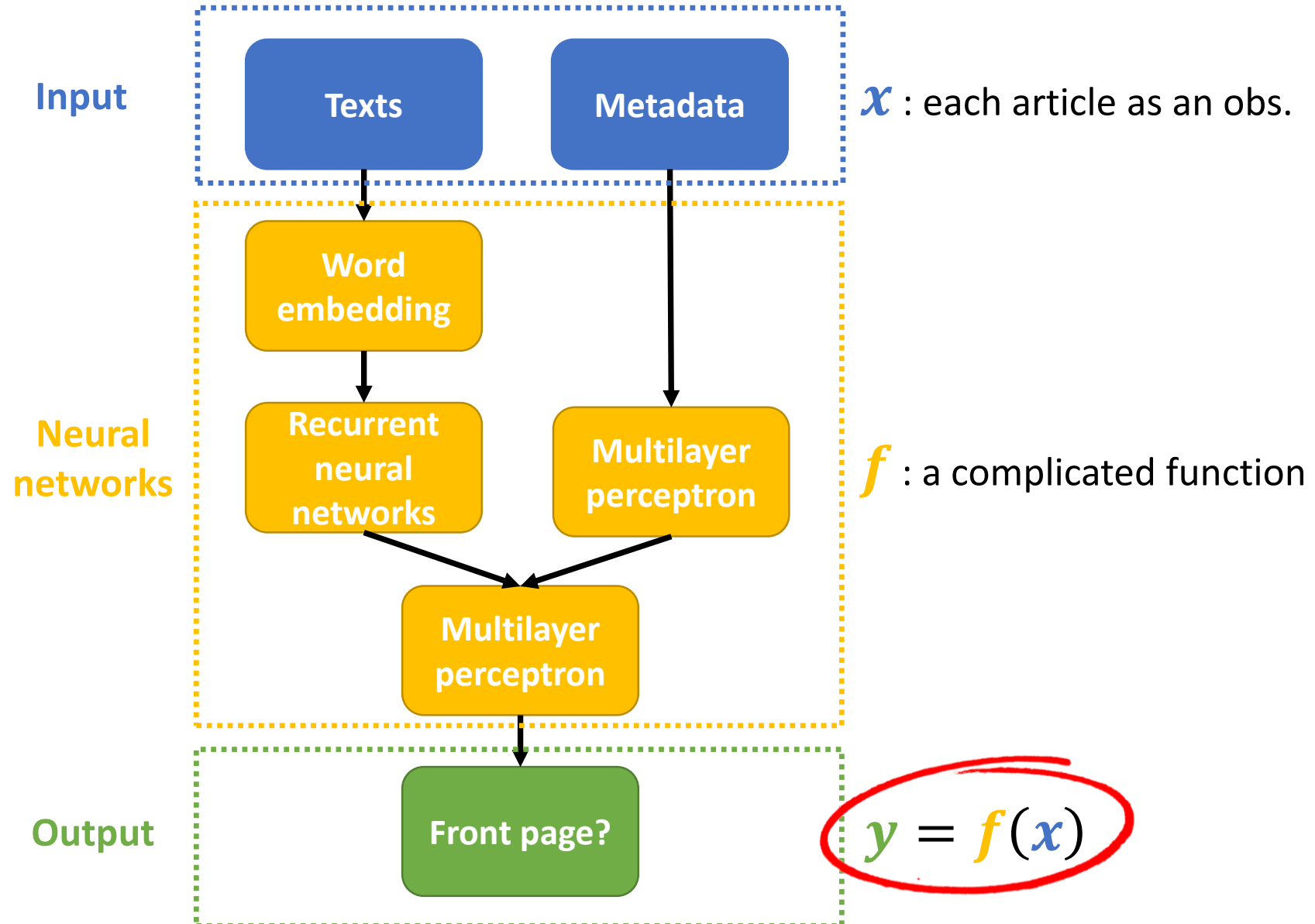
sample_df

id	date	page		title	body
20181202525	2018-12-31 00:00:00	1.0	今年粮食总产量13158亿斤 农业农村发展取得新成绩，乡村振兴开局良好	本报北京12月30日电 （记者高云才、郁静娴）记者从30日召开的全国农业农村厅局长会议获...	
20181202526	2018-12-31 00:00:00	1.0	新时代改革再出发的重要里程碑 ——写在党的十八届三中全会召开五周年之际	本报评论员 改革开放40年来，我们党带领人民绘就了一幅波澜壮阔、气势恢宏的历史画卷。其...	
20181202527	2018-12-31 00:00:00	1.0	改革引领中国经济新航向（在习近平新时代中国特色社会主义思想指引下——新时代新作为新篇章） ...	本报记者 许志峰 吴秋余 王 珂 林丽福 巨轮向前，离不开航标指引方向。 1978年...	
20181202528	2018-12-31 00:00:00	1.0	产量连续六年超1200亿斤，实现“十五连丰” 黑龙江 粮食再丰收，粮仓更稳固	本报哈尔滨12月30日电 （记者方圆）此刻的黑土地，正在冬眠蓄力。今年黑龙江粮食又是一...	
20181202529	2018-12-31 00:00:00	1.0	伟大的变革——庆祝改革开放40周年大型展览 现场观众达223万人次 展览将持续到2019年...	据新华社北京12月30日电 “伟大的变革——庆祝改革开放40周年大型展览”自11月13日...	
20181202530	2018-12-31 00:00:00	1.0	航线不停、景区不休、生意不闲、游客不断 新疆 旅游无淡季，今冬更红火	本报乌鲁木齐12月30日电 （记者杨明方、阿尔达克）一年之中半年是降雪期，冬季雪深超过3...	
20181202531	2018-12-31 00:00:00	1.0		图片报道 徜徉知识海洋，感受阅读乐趣。十二月三十日是元旦小长假的第一天，全国各地的特色文化活动丰富...	
20181202532	2018-12-31 00:00:00	2.0	创造新的更大奇迹 ——党的十八届三中全会召开五周年述评	“全面发力、多点突破、蹄疾步稳、纵深推进”——庆祝改革开放40周年大会上，习近平总书记用...	
20181202534	2018-12-31 00:00:00	3.0	开启中国特色大国外交新征程（2018年度国际特别报道） ——国务委员兼外交部长王毅回顾20...	本报记者 赵嘉鸣 管克江 杜尚泽 焦 翔 王海林 国务委员兼外交部长王毅12月29日接...	
20181202535	2018-12-31 00:00:00	3.0	二〇一七年岁末，莫斯科大学的中国留学生们欣喜地收到了习近平主席的回信；一年来，牢记习主席的...	本报驻俄罗斯记者 吴 焰 “一年前，习近平主席给留学生回信，鼓励我们胸怀大志，刻苦学习...	
20181202536	2018-12-31 00:00:00	3.0	外交部发言人就中美建交40周年发表谈话	本报北京12月30日电 外交部发言人陆慷30日就中美建交40周年发表谈话。 陆慷说，2...	

Modeling



Modeling

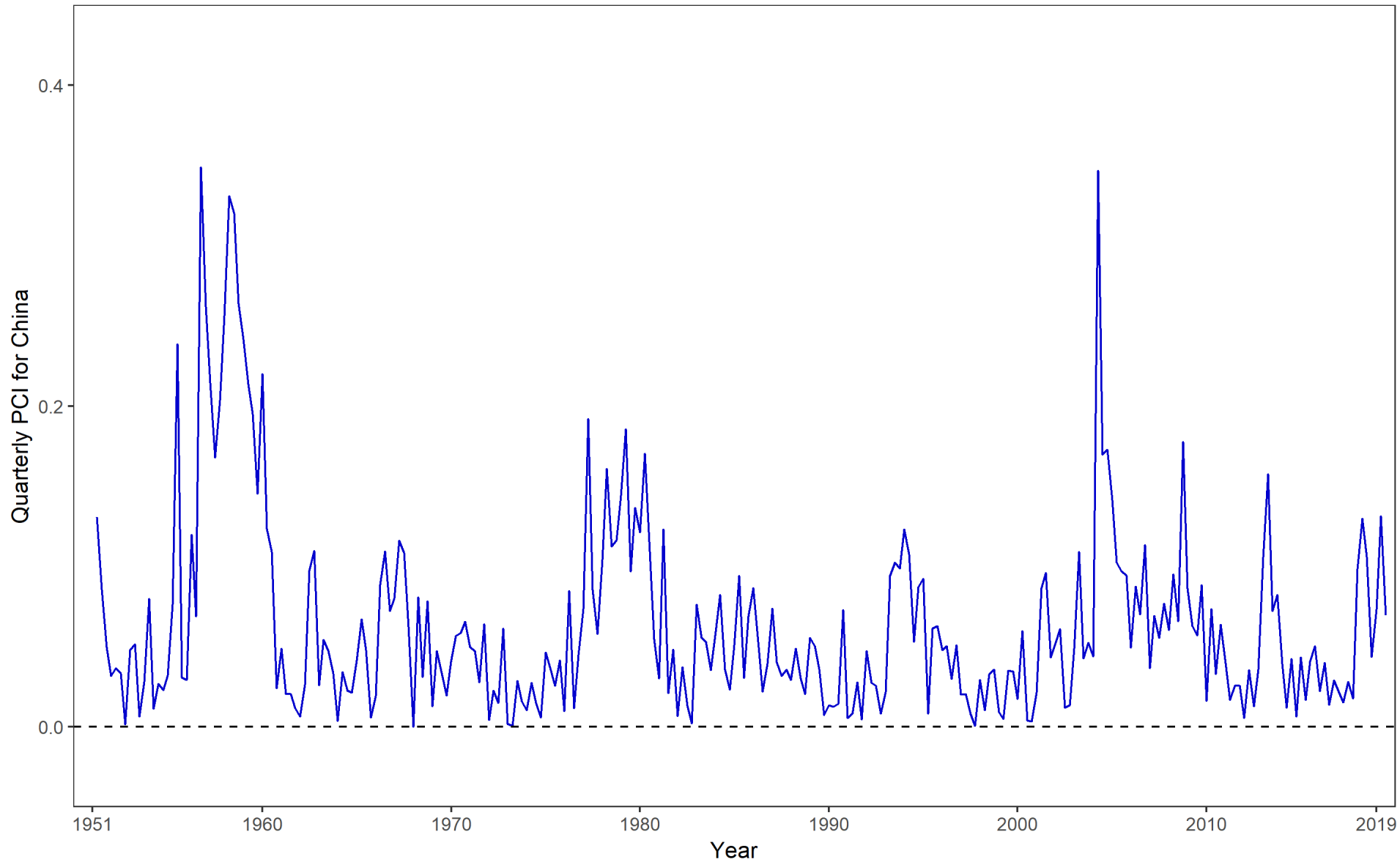


Source of predictive power

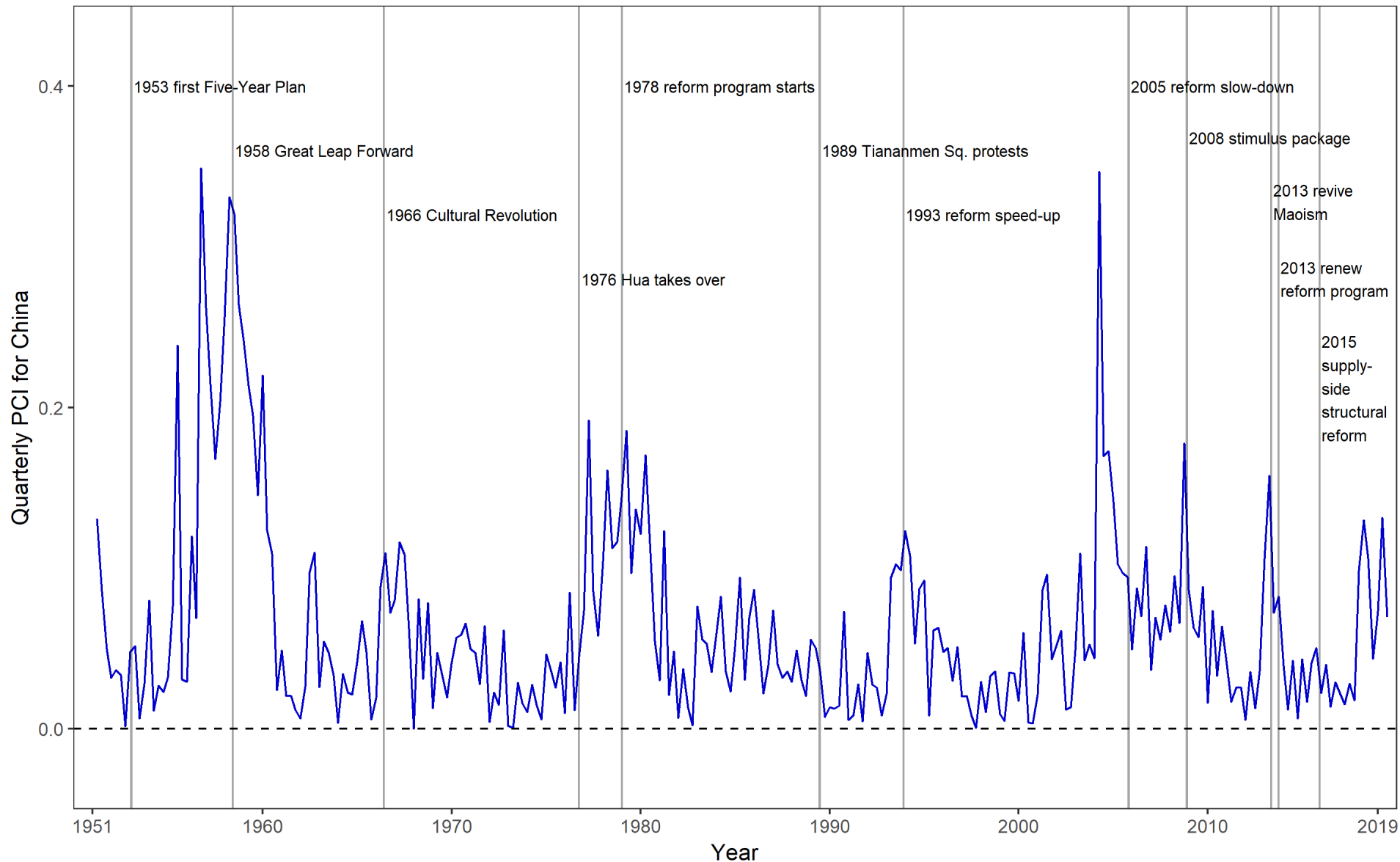
Constructing the PCI

Results

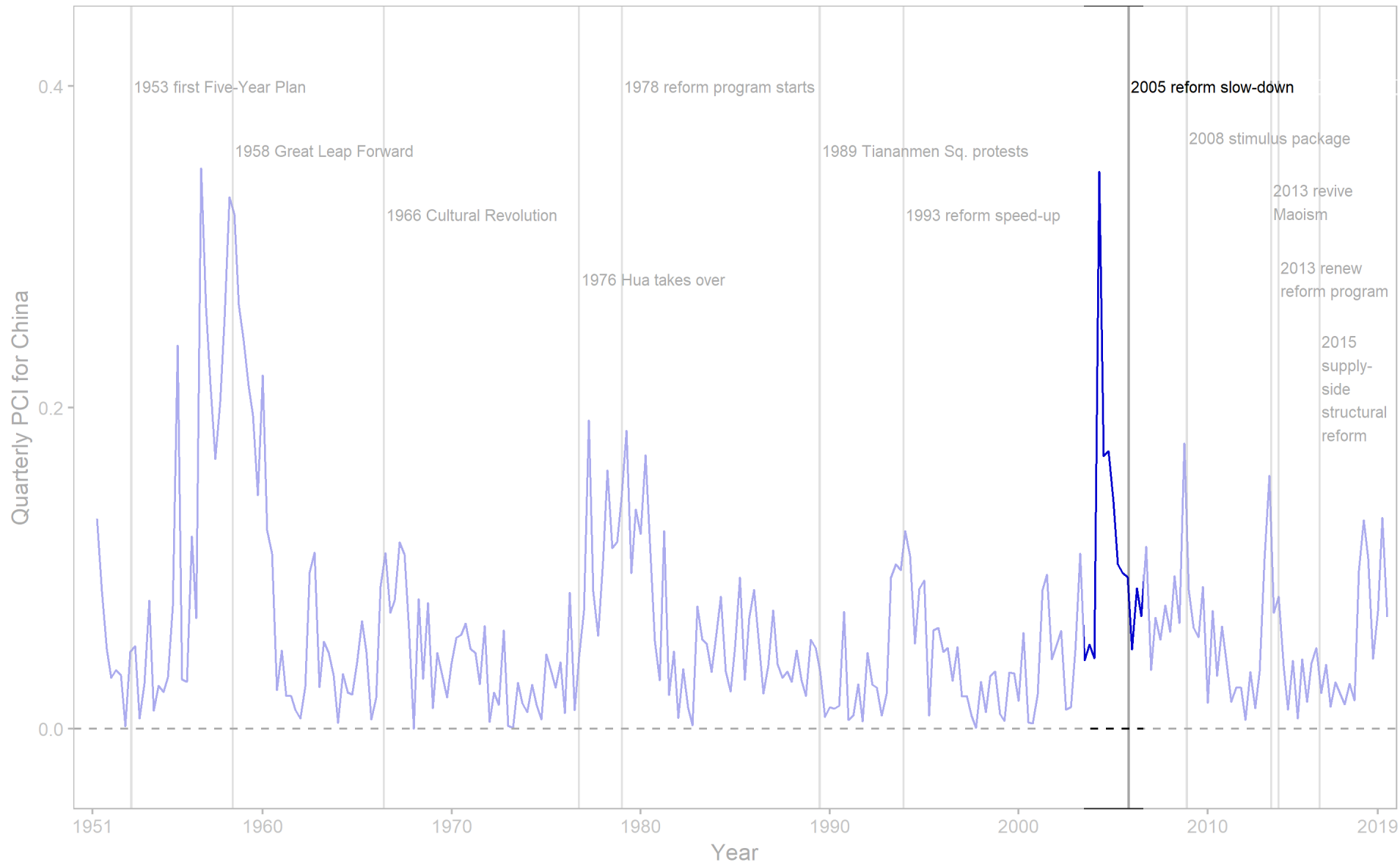
PCI for China



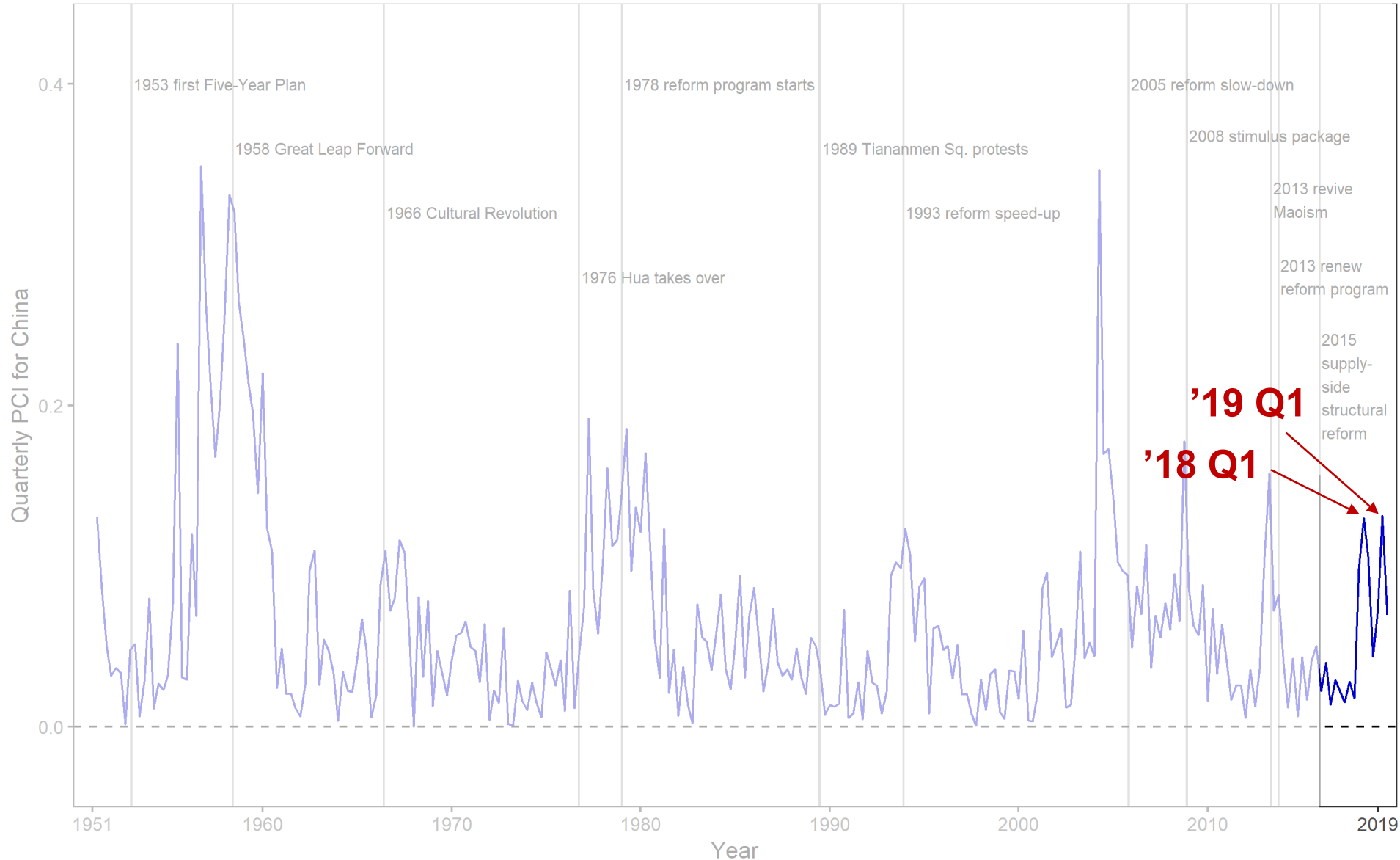
PCI for China — with ground truth



PCI for China — with ground truth



PCI for China — going forward



Understanding changes

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

Understanding changes

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

Content of ***misclassified*** articles has policy substance.

Understanding changes — 2018 Q1 uptick

Understanding changes — 2018 Q1 uptick

- 2018 Q1 spike represents:
 - internally: strengthening party authority;
 - externally: nationalism and global leadership;
 - populist policies to boost political support.

Understanding changes — 2018 Q1 uptick

- 2018 Q1 spike represents:
 - internally: strengthening party authority;
 - externally: nationalism and global leadership;
 - populist policies to boost political support.
- 2019 Q1 uptick: similar patterns.

Understanding changes — 2018 Q1 uptick

- 2018 Q1 spike represents:
 - internally: strengthening party authority;
 - externally: nationalism and global leadership;
 - populist policies to boost political support.
- 2019 Q1 uptick: similar patterns.

⇒ *Curb your enthusiasm for the US-China trade talks!*

Takeaways: many applications

Generally

- PCI for other (ex-)Communist regimes, using:
 - Soviet Union's *Pravda*
 - East Germany's *Neues Deutschland*
 - Cuba's *Granma*
 - North Korea's *Rodong Sinmun*
 - Vietnam's *Nhân Dân*
 - ...

Generally

- PCI for other (ex-)Communist regimes, using:
 - Soviet Union's *Pravda*
 - East Germany's *Neues Deutschland*
 - Cuba's *Granma*
 - North Korea's *Rodong Sinmun*
 - Vietnam's *Nhân Dân*
 - ...
- } Work in progress

More generally

- PCI for other countries:
 - The state may not control the media.
 - But politicians/policymakers still talk!

Even more generally

- Applicability well beyond predicting policies.
- Omitted here. See our [research paper](#).

Interested in DIY?

- Website: policychangeindex.org
- Newsletter: mercatus.org/PCInewsletter
- Repository: github.com/PSLmodels/PCI
- A [simulated example](#) to show how the PCI works.

Interested in DIY?

- Website: policychangeindex.org
- Newsletter: mercatus.org/PCInewsletter
- Repository: github.com/PSLmodels/PCI
- A [simulated example](#) to show how the PCI works.

Questions?