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.
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;
  - ...
  - *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
Propaganda often precedes policies.

Detect changes in newspaper’s priorities

≈

Predict changes in gov’t policies

Source of predictive power

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

Machine learning
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

Articles in previous 5 years

Model a front-page classifier

Test
The framework

Articles in previous 5 years

Articles in next quarter

Model a front-page classifier

Test

“Forecast”
The framework

Policy Change Index = Test performance − “Forecast” performance

PCI ≫ 0 ⇒ Structural difference in the newspaper’s priorities
<table>
<thead>
<tr>
<th>id</th>
<th>date</th>
<th>page</th>
<th>title</th>
<th>body</th>
</tr>
</thead>
<tbody>
<tr>
<td>20181202525</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>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</td>
<td>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...</td>
</tr>
<tr>
<td>20181202526</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>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</td>
<td>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</td>
</tr>
</tbody>
</table>
### Data structure

<table>
<thead>
<tr>
<th>id</th>
<th>date</th>
<th>page</th>
<th>title</th>
<th>body</th>
</tr>
</thead>
<tbody>
<tr>
<td>20181202525</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>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</td>
<td>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会...</td>
</tr>
<tr>
<td>20181202526</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>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</td>
<td>本报评论员 改革开放40年来，我们党带领人民绘就了幅波澜壮阔、气势恢宏的历史画卷。其...</td>
</tr>
<tr>
<td>20181202527</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>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） ...</td>
<td>本报记者 许志峰 吴秋余 王 珂 林丽卿 巨轮向前，离不开航标指引方向。1978年...</td>
</tr>
<tr>
<td>20181202528</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>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</td>
<td>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...</td>
</tr>
<tr>
<td>20181202529</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>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</td>
<td>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...</td>
</tr>
<tr>
<td>20181202530</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>航线不停、景区不休、生意不闲、游客不断 新疆 旅游无淡季，今冬更火</td>
<td>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...</td>
</tr>
<tr>
<td>20181202531</td>
<td>2018-12-31</td>
<td>1.0</td>
<td>图片报道</td>
<td>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...</td>
</tr>
<tr>
<td>20181202532</td>
<td>2018-12-31</td>
<td>2.0</td>
<td>创造新的更大奇迹 ——党的十八届三中全会召开五周年述评</td>
<td>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...</td>
</tr>
<tr>
<td>20181202533</td>
<td>2018-12-31</td>
<td>3.0</td>
<td>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...</td>
<td>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...</td>
</tr>
<tr>
<td>20181202534</td>
<td>2018-12-31</td>
<td>3.0</td>
<td>二〇一七年岁末，莫斯科大学的中国留学生纷纷收到了习近平主席的回信 ——2017年，牢记习主席的...</td>
<td>本报驻俄罗斯记者 吴 焜 “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...</td>
</tr>
<tr>
<td>20181202535</td>
<td>2018-12-31</td>
<td>3.0</td>
<td>外交部发言人就中美建交40周年发表谈话</td>
<td>本报记者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...</td>
</tr>
</tbody>
</table>
Modeling

Input
- Texts
- Metadata

Neural networks
- Word embedding
- Recurrent neural networks
- Multilayer perceptron
- Multilayer perceptron

Output
- Front page?

\[ x : \text{each article as an obs.} \]
\[ f : \text{a complicated function} \]

\[ y = f(x) \]
Modeling

\( \mathbf{x} : \) each article as an obs.

\( f : \) a complicated function

\[ y = f(\mathbf{x}) \]
Source of predictive power

Constructing the PCI

Results
PCI for China
PCI for China — with ground truth

- 1953 first Five-Year Plan
- 1958 Great Leap Forward
- 1966 Cultural Revolution
- 1976 Hua takes over
- 1978 reform program starts
- 1989 Tiananmen Sq. protests
- 1993 reform speed-up
- 2005 reform slow-down
- 2008 stimulus package
- 2013 revive Maoism
- 2013 renew reform program
- 2015 supply-side structural reform
PCI for China — going forward

- 1953 first Five-Year Plan
- 1958 Great Leap Forward
- 1966 Cultural Revolution
- 1976 Hua takes over
- 1978 reform program starts
- 1989 Tiananmen Sq. protests
- 1993 reform speed-up
- 2005 reform slow-down
- 2008 stimulus package
- 2013 revive Maoism
- 2013 renew reform program
- 2015 supply-side structural reform

'18 Q1
'19 Q1
## Understanding changes

<table>
<thead>
<tr>
<th>Front page?</th>
<th>Classified on front page?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

- **False positives**: When a change is not on the front page but is classified as such.
- **False negatives**: When a change is on the front page but is not classified as such.
Understanding changes

<table>
<thead>
<tr>
<th>Front page?</th>
<th>Classified on front page?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

- False positives: √
- 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?