

Using Numbers to Study Mainland China

Chao-Yo Cheng

(Birkbeck, University of London)



Introduction: Using numbers to study mainland China

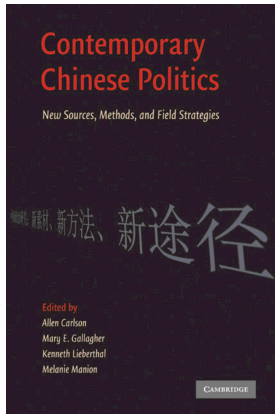
- ▶ Why numbers matter: A paradigm shift (?)
- ▶ Where data comes from
- ▶ What you need to work with data
- ▶ Examples from past and current research
- ▶ Concluding remarks and tips



Why numbers matter: A paradigm shift?

- ▶ The field of (mainland) China studies has changed significantly in the past decade as a result of **methodological/technological advances**, (new) **data availability** and **geopolitics**
- ▶ Quant techniques are essential to **establish patterns** for theory building and policy impact evaluation
- ▶ Using numbers also encourage us to rethink the purpose of mainland China studies – government and/or governance?
- ▶ The use of computational and quantitative data/methods by no means undermine the importance/value of traditional approaches – **multi-method/data triangulation** is now the norm





Manion, Melanie (Duke), Kenneth Lieberthal (Michigan), Mary E. Gallagher (Michigan), and Allen Carlson (Cornell). 2010. *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*. New York, NY: Cambridge University Press.



Where data comes from

- ▶ Surveys (e.g., Chinese General Social Survey)
- ▶ Admin data and census (e.g., statistical yearbooks)
- ▶ Archives and historical documents, such as
 - Gazetteers
 - Culture and history materials
- ▶ Digital and "big" data (e.g., Weibo and Baidu Index)



北京大学开放研究数据平台

重用数据，再现科研。
提升引用，跟踪影响。

开放数据
为您的研究点赞！

创建我的数据空间

或 浏览平台数据

<https://opendata.pku.edu.cn/>



法律法规

标题 法律法规

开始检索

高级检索
法宝订阅

同义词 | 精确 | 模糊 | 网络 | 同条 | 同段 | 同句 | 法规变迁 | 法宝联想 | 公报索引 | 法宝引证码

中央法规

地方法规

立法资料

立法计划

中外条约

外国法规

香港法规

澳门法规

台湾地区法规

法律动态

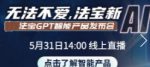
合同范本

法律文书

境外法律信息资源指引

首页 > 法律法规 > 中央法规

本库说明



相关提示

本月新颁 (786)

本月生效 (816)

本月时效变更 (95)

效力位阶

+ 法律 (3176)

+ 行政法规 (11562)

监察法规 (1)

+ 司法解释 (8841)

+ 部门规章 (347778)

+ 军事法规规章 (279)

方针政策

法律动态

案例报道

文章推荐

我的订阅

1. 中共中央、国务院印发《国家水网建设规划纲要》

现行有效 / 2023.05.25公布 / 2023.05.25施行

2. 中共中央办公厅、国务院办公厅印发《关于推进基本养老保险体系建设的意见》

现行有效 / 2023.05.21公布 / 2023.05.21施行

3. 中共中央办公厅、国务院办公厅印发《关于全面加强新形势下森林草原防火工作的意见》

现行有效 / 2023.04.20公布 / 2023.04.20施行

4. 中共中央印发《中央党内法规制定工作规划纲要(2023 - 2027年)》

现行有效 / 2023.04.18公布 / 2023.04.18施行


5. 中共中央发出关于学习《习近平著作选读》第一卷、第二卷的通知

现行有效 / 2023.04.10公布 / 2023.04.10施行

加载更多

<https://www.pkulaw.net/>





China Data Online

Welcome **Guest**, Your IP is: 114.36.26.217

[Home](#)
[Data Products](#)
[Dictionary](#)
[Support](#)
[Contact](#)
[Q&A](#)
[Citations](#)
[Workshops](#)
[Subscription](#)
[Login](#)

SPATIAL DATA

- China Geo-Explorer II
- China Geo-Explorer I
- China Map Library

CHINA STATISTICS

- National Statistics
- Provincial Statistics
- City Statistics
- County Statistics(1997-)
- Industrial Data
- Statistics Datasheets
- Statistics Charts





CENSUS DATA

- Census Maps
- All Census Data
- Economic Census 2004
- Industrial Census 1995
- Census 1982
- Census 1982(10%)
- Census 1990
- Census 1995(1%)
- Province 2000
- County 2000
- Census 2005(1%)
- Census Data Search

FREE CHINA MAPS

- 2000 Population Census
- Pop&Env(1990-1999)
- Pop&Env(2000)
- Atlas of Industrial Census

Workshops and Seminars

The Summer Leadership Workshop on Spatiotemporal Innovation

1726 Cambridge Street, Cambridge MA 02139
July 10-14 2023

2023-02-14
The Summer Leadership Workshop on Spatiotemporal Innovation

[More...](#)

Latest China Statistical News

- Consumer Prices in February 2023(2023/03/20)
- The Gross Imports and Exports Amounts 6.18 Trillion Yuan in the First Two Months of 2023(2023/03/20)
- Total Retail Sales of Consumer Goods from January to February 2023(2023/03/20)

<https://www.china-data-online.com/>





<https://index.baidu.com/>



What you need to work with data

- ▶ Training in different **research methods**
- ▶ A grasp of potential **data sources**
- ▶ Solid and deep **field experiences** and/or **subject knowledge**



What you need to work with data

- ▶ Training in different **research methods**
 - Qualitative: Case study (Bayesian and machine learning), ethnography and (elite) interviews
 - Quantitative: Econometrics and regression (frequentist and Bayesian), experiments and causal inference
 - Computational: Text-as-data, social network analysis, geographic data science and ChatGPT-assisted research
- ▶ A grasp of potential **data sources**
- ▶ Solid and deep **field experiences** and/or **subject knowledge**



What you need to work with data

- ▶ Training of different **research methods**
- ▶ A grasp of potential **data sources**
 - Databases
 - Libraries and archives
 - Survey firms and research centres
 - API and web scraping
 - Replication packs
- ▶ Solid and deep **field experiences** and/or **subject knowledge**



What you need to work with data

- ▶ Training in different **research methods**
- ▶ A grasp of potential **data sources**
- ▶ Solid and deep **field experiences** (e.g., politics of data generation and release) and/or **subject knowledge** (e.g., political selection and government finances)



Past and current research

- ▶ Rethinking the politics of ethnic local autonomy (book project)
 - Poverty alleviation and local state-building in peripheral provinces (*JJPS* 2021)
 - Ethnic empowerment and environmental governance
 - Dual descriptive representation of ethnic cadres
- ▶ (Re)mapping the information network of political elites using computational approaches
- ▶ Using court records to study administrative litigation, with **Haibo He** (Tsinghua) and **Chao Ma** (UIBE)
- ▶ Multilevel agenda setting in times of crisis: A text-as-data approach, with **Tao Lin** (University of Washington)



Rethinking the politics of ethnic local autonomy



Rethinking the politics of ethnic local autonomy

- ▶ Why do the Communist Party of China grant **ethnic local autonomy**?
- ▶ When and where will the Center designate **ethnic autonomous territories** (EATs)?
- ▶ How can local (policy/administrative) decentralization contribute to nation-state building?



Rethinking the politics of ethnic local autonomy

- ▶ **Archival research** in the US, Hong Kong and mainland China
- ▶ **Elite interviews:**
Conversations with Han and non-Han cadres at the central and local levels across different provinces
- ▶ **Process-tracing and QCA:**
Structured process tracing the designation of ethnic autonomous prefectures



Hoover Institution Archives
(Palo Alto, California)



University of California, Berkeley
(Berkeley, California)



The Chinese University of Hong Kong
(Hong Kong)



Fudan University
(Shanghai)



Minzu University of China
(Beijing)



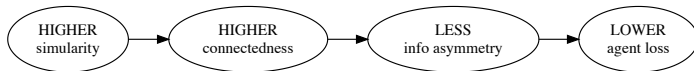
The National Library of China
(Beijing)



Rethinking the politics of ethnic local autonomy

- ▶ **Supervised machine learning and social network analysis:** Capturing the *connectedness* and *network structure* among/between central and local cadres based on an original biographical dataset of Chinese political elites (1949-2003)

- Between the pair of cadres, I create the indices of **connectedness** based on their **biographical similarity**



- The same connectedness measure can be used to construct the information network of Politburo members and measure the degree of **central leadership fragmentation** by **community detection**
- ▶ **Multivariate regression:** Modeling the number of active EATs as a function of different network structural indicators



Rethinking the politics of ethnic local autonomy

- Ethnic local autonomy acts as an **institution of agent control**

	Results
(1) Central leadership fragmentation (Modularity of the Politburo network)	+ EATs
(2) Local leadership embeddedness (Similarity b/w provincial Party secretary and governor)	+ EATs
(3) Central-local connectedness (Similarity b/w central leader and provincial Party secretary)	– EATs
(4) Central fragmentation + Local embeddedness	+ EATs
(5) Central fragmentation + Central-local connectedness	– EATs
(6) Central-local connectedness + Local embeddedness	– EATs



Rethinking the politics of ethnic local autonomy

- ▶ EATs act as **enclaves** that constrain the power of provincial elites
 - Sub-provincial EATs are **more** likely to pass and revise autonomous regulations when the provincial elites are potentially defiant
 - Ruling non-Han cadres in sub-provincial EATs are **less** likely to have worked at the provincial level prior to their appointments



(Re)mapping the information network of political elites



(Re)mapping the information network of political elites

- ▶ The focus on "high-politics" in the literature on Communist regimes: The absence of independent, strong formal institutions has pushed researchers have to focus on prominent political elites in authoritarian regimes
- ▶ **Factionalism** is one of the most researched topics in the Chinese politics literature (e.g., Nathan 1973; Tsou 1976; Pye 1981; Dittmer 2002; Huang 2006; Bo 2007; Shih et al 2012)
- ▶ The power struggles and cooperation between different factions or "informal" groups have been considered as the main driving forces for policy and political changes in the country



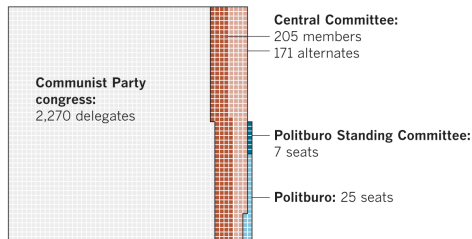
(Re)mapping the information network of political elites

- ▶ Create the timeline of key life events of each cadre using public biographic information
 - Demographic characteristics (e.g., birth year and provincial origin)
 - Educational background (e.g., colleges and universities)
 - Political career (e.g., involvement in key events between 1921 and 1949; Party and government posts after 1949)
- ▶ Create the (decomposed) **attribute vector** of a cadre based on her/her timeline and measure the distance/similarity between a pair of attribute vectors in the multidimensional space
 - Time: Are the cadres of interest both alive in a given year?
 - Location: Are they both at the same location in a given year?
 - Activity: Are they working in the same department/unit in a given year?



(Re)mapping the information network of political elites

China's Communist Party congress (as of 2012)



Source: Los Angeles Times reporting

Angelica Quintero / @latimesgraphics

- ▶ Use **variable selection** (elastic net regularization and other methods) to identify the key biographical variables in CC of each year when measuring biographical similarity
- ▶ Use **community detection** to assess the degree to which a Politburo network can be divided



Using court records to study admin litigation



Using court records to study admin litigation

The release of **China Judgements Online** (<https://wenshu.court.gov.cn/>) made a crucial breakthrough in the empirical legal studies in China (Ma, Yu and He 2016; Liebman et al 2020).

The screenshot shows the homepage of the China Judgements Online website. At the top is the Chinese judicial emblem. Below it, the title '中国裁判文书网' (China Judgements Online) is displayed. A navigation bar includes links for '首页' (Home), '刑事案件' (Criminal Cases), '民事案件' (Civil Cases), '行政案件' (Administrative Cases), '赔偿案件' (Compensation Cases), '执行案件' (Enforcement Cases), '其他案件' (Other Cases), and '民族语言文书' (Ethnic Language Documents). A search bar with the placeholder '高级检索' (Advanced Search) and the text '输入案由、关键词、法院、当事人、律师' (Enter case type, keywords, court, parties, lawyer) is present, along with a '搜索' (Search) button and a help icon. Below the search bar, the slogan '开放 动态 透明 便民' (Open, Dynamic, Transparent, Convenient) is shown. The bottom section features three circular icons representing statistics: '今日新增' (New today), '文书总量' (Total documents), and '访问总量' (Total visits). Each icon is accompanied by a list of document types and their counts.

Category	Sub-category	Count
今日新增 (New today)	民事文书 (Civil documents)	86,473,907篇
	刑事文书 (Criminal documents)	137,334篇
文书总量 (Total documents)	刑事文书 (Criminal documents)	9,996,704篇
	执行文书 (Enforcement documents)	40,340,403篇
访问总量 (Total visits)	行政文书 (Administrative documents)	3,032,867篇
	其他文书 (Other documents)	1,134,802篇



Using court records to study admin litigation

- ▶ **Question:** Are the local courts more likely stand with the citizens following the reforms?
 - Railway transport courts (*China Review* 2022)
 - Case registration system
 - Elevated and off-site trials
- ▶ **Extensions**
 - Who do the citizens go to the courts? A national **conjoint** experiment
 - Local dynamics of judicial decision-making: A **text-as-data** approach



Concluding remarks

- ▶ Find out **who you are** and **what you want to be and do** – your decision may depend on where you are (or going back to history)
- ▶ Curate the skills you need – be **selectively thorough**
- ▶ Find those who can support and collaborate with you – engage in **interdisciplinary knowledge co-production**
- ▶ Read broadly – **go beyond political science and (mainland) China** to gain inspiration and a broader comparative perspective



Thank you!

c.cheng@bbk.ac.uk



Rethinking the politics of ethnic local autonomy

- ▶ Going **cross-national**: Post-WWII authoritarian regimes are more likely to introduce regional autonomy to ethnic minorities when their ruling power is contested within the inner circle
- ▶ Going back to **history**: Post-civil war *decentralization* and *state power consolidation* in imperial and Republican China
 - Western Han dynasty (202 BC–220 AD): Kingdoms and marquis states
 - The Kuomintang (KMT) regime (1928-1949 AD): Countering recalcitrant provincial warlords through the county-level self-government movement

