

Analysis: preliminary dataset for stay-at-home requirements

Stay-at-home requirements

Page purpose: This page summarizes our approach to generating a dataset for stay at home requirements for Chinese provinces. It describes the process for generating preliminary datasets and also proposes next steps for more in-depth analysis

Analysis objective: build a dataset of stay-at-home requirements, including granular policy information about the compliance of the policy, the start date of the policy, and when the policy was lifted

Analysis takeaways:

- 90% of the lockdown policy information in Solomon Hsiang's paper come from three sources, two of which are still accessible.
- These raw sources include information at the city and prefecture level for i) complete lockdowns, ii) partial lockdowns, and iii) contained management (internal travel bans if varying stringency)
- Only cities in Hubei province underwent a complete lockdown. Other cities and provinces outside of Hubei underwent partial lockdowns. And many more cities experienced a closed management regime
- To complete this dataset, we should explore the remaining sources in Solomon Hsiang's dataset and also run targeted searches to determine the date the partial lockdowns were lifted in each region

Links:

Code	policy-database / notebooks / explore-china-policy-databases.R
Input data	policy-database / data / external
Output data	policy-database / notebooks / outputs

Table of contents:

- Stay-at-home requirements
- Data sources
- Data cleaning
 - Classification
 - Dataset overlap
- Output
 - solomon-source-policy-cln-20201014.csv
- Next steps
 - Collect the dates of the lifted partial lockdown requirements

Data sources

Solomon Hsiang in his publication [The effect of large-scale anti-contagion policies on the COVID-19 pandemic](#) looks at the impact of three policy types: `home_isolation`, `emergency_declaration`, and `travel_ban_local`. Of these three policy types, we are most concerned with `home_isolation` for our analyses. While Hsiang does not provide the level of detail we need for our analyses, he does provide a dataset of underlying news announcements used for his analysis (available on Github here: https://github.com/Global-Policy-Lab/gpl-covid/blob/master/data/raw/china/CHN_policy_data_sources.csv).

Upon review of this dataset, we identified 3 sources responsible for ~90% of the policies labeled as `home_isolation`. Those are

1. Fang et al., Human mobility restrictions and the spread of the novel coronavirus (see pg 35)
2. Wikipedia page on the COVID-19 pandemic lockdown in Hubei
3. [source not available] <https://github.com/Pratitya/wuhan2020-timeline/>

url	N
https://en.wikipedia.org/wiki/2020_Hubei_lockdowns	78
https://github.com/Pratitya/wuhan2020-timeline/blob/master/%E6%97%B6%E9%97%B4%E7%BA%BFTIMELINE.md	16
https://www.nber.org/papers/w26906.pdf	11
http://news.stcn.com/2020/0210/15633808.shtml ; http://finance.eastmoney.com/a/202002061374896985.html ; http://china.qianlong.com/2020/0315/3844379.shtml ; http://finance.eastmoney.com/a/202002241394722886.html	1

http://www.asahi.com/ajw/articles/13105716	1
http://www.bozhou.gov.cn/News/show/372139.html	1
http://www.caixin.com/2020-02-11/101513995.html	1
https://changsha.ihouse.ifeng.com/news/2020_02_03-52563107_0.shtml	1
https://en.wikipedia.org/wiki/2020_Hubei_lockdowns ; http://news.stcn.com/2020/0210/15633808.shtml	1
https://news.sina.com.cn/c/2020-01-31/doc-iimxyqvy9241769.shtml , https://www.sqnu.edu.cn/info/1013/14613.htm , http://finance.eastmoney.com/a/202002111378815530.html	1
https://www.aljazeera.com/news/2020/02/china-admits-shortcomings-coronavirus-death-toll-hits-425-200203234036932.html	1
https://www.soundofhope.org/post/339913	1
https://xw.qq.com/cmsid/20200213A0J61000	1

Note: We could also consider the work done by He et al., 2020, The short-term impacts of COVID-19 lockdown on urban air pollution in China which also studies lockdowns. In his dataset, he defined a lockdown when all three of the following preventive measures were enforced: (1) prohibition of unnecessary commercial activities in people's daily lives; (2) prohibition of any types of gathering by residents; (3) restrictions on private (vehicle) and public transportation. Following this definition, 95 out of 324 cities were locked down.

Data cleaning

Classification

We extracted the raw data for two of these three sources (the third source is no longer available). Each source includes its own definitions of "stay-at-home", which we have standardized below. In addition to this, the spelling of Chinese provinces and cities was standardized across datasets to streamline our analysis.

Stay-at-home term	Source	Definition	Reclassification
Complete shutdown	Fang et al.	all public transport and private vehicles are banned in the city, all residential buildings are locked down, and residents are not allowed to leave the city	Complete shutdown
Partial shutdown	Fang et al.	majority of the public transportation has been temporarily shut down, checkpoints has been set up to control the inflow population, and surveillance and tighter controls in each neighborhood	Partial shutdown
Checkpoints and quarantine zones	Fang et al.	set up checkpoints and quarantine zones, and public transport maintains normal operation	Closed management
Quarantine	Wikipedia	transport severely restricted, close of trains/planes, shutdown of all nonessential companies. These restrictions were only in place for cities in Hubei province	Complete shutdown
Household-based outdoor restrictions	Wikipedia	Hubei implemented a measure whereby only one person from each household is permitted to go outside for provisions once every two days, except for medical reasons or to work at shops or pharmacies. Many cities, districts, and counties across mainland China implemented similar measures in the days following	Partial shutdown
Closed management	Wikipedia	in most of the areas where this came into effect, villages, communities, and units in most areas would only keep one entrance and exit point open, and each household is allowed limited numbers of entrances and exits. In some places, night-time access is prohibited, effectively a curfew, and in extreme cases, access is prohibited throughout the day. ^[46] People entering and leaving are required to wear masks and receive temperature tests. In some areas, vouchers are issued to the public, with vouchers and valid credentials	Closed management

Dataset overlap

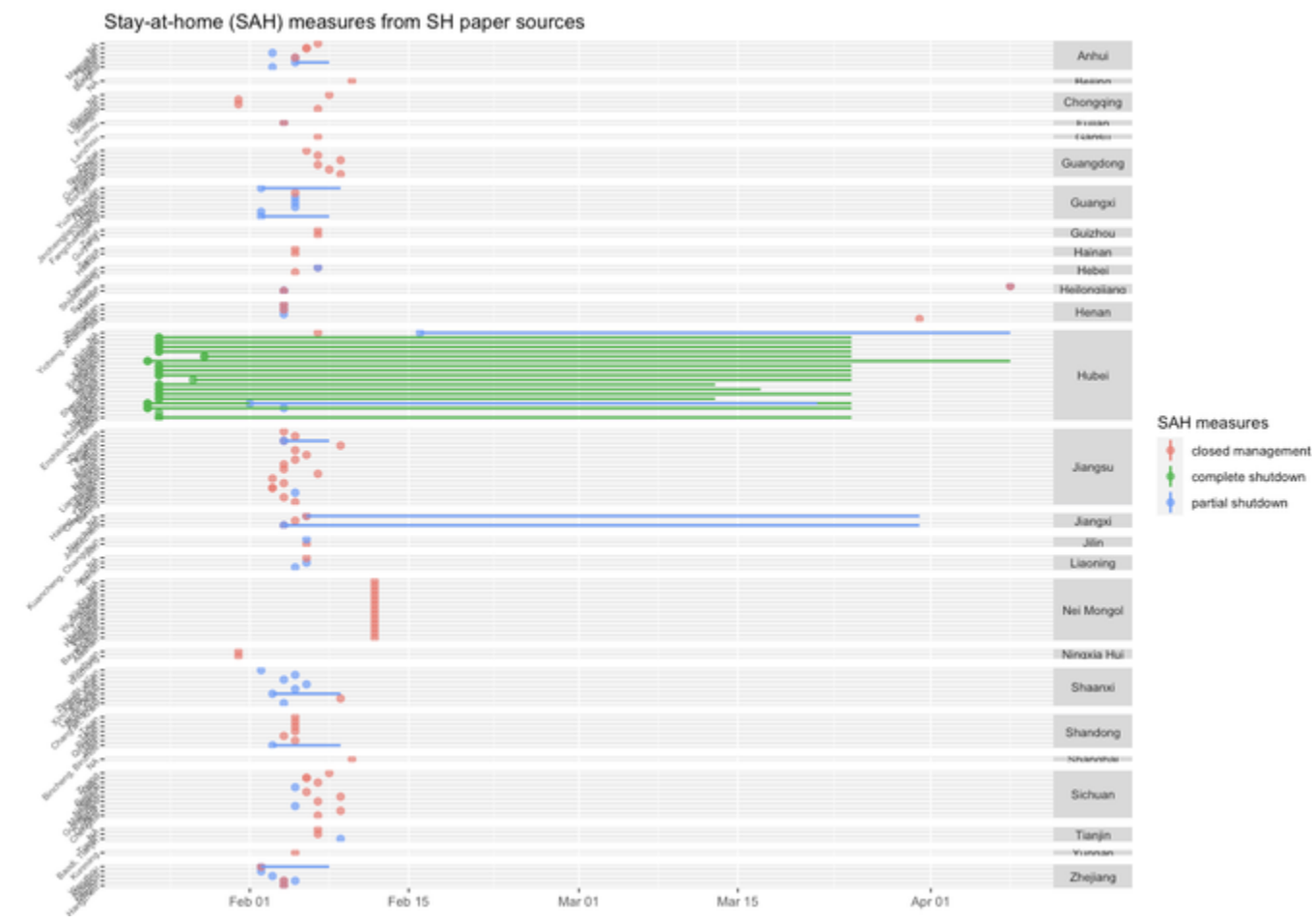
I manually reviewed the resulting combination of these two datasources for contradictory overlap. There are two areas that I reviewed and flagged for follow up

- Contradictory policy classifications:** My deduction is that closed management policies may overlap with shutdown policies, but the two types of shutdown policies, partial and complete may not. No data was removed in this step
- Contradictory date overlaps:** Only some of the collected data includes end dates. In these scenarios, I check to make sure that the end date does not overlap in a contradictory way with the start date of a subsequent policy. There are two records, both for cities in the Hubei province where such an overlap exists. These have not been removed from our analysis

Output

The resulting dataset includes the start dates and some known end dates for the available lockdown policies from the two sources listed above. The below figure illustrates the resulting dataset, which can be accessed in Dropbox here:

:outbox_tray: solomon-source-policy-cln-20201014.csv



Next steps

To feel confident in this dataset, we need two more pieces of information:

:one: First, we could spend more time understanding if this dataset is comprehensive by reviewing the other sources included in Solomon's dataset. Based on our discussion on 15 Oct 2020, we decided that this incremental work would not add significant value to our analysis and will not proceed with it

:two: Second, we want to understand when each home isolation requirement was lifted. Unfortunately not all of the available sources include this information, likely because the paper was published while these policies were still in effect. This exercise could be completed through news review with targeted searches in the above cities or provinces (maybe subset only to those with partial lockdowns)

:two: Collect the dates of the lifted partial lockdown requirements

Background: We have identified the start dates for a series of policy announcements for **partial shutdowns** in China. Partial shutdowns are defined as policies that lead to the majority of the public transportation being temporarily shut down, checkpoints being set up to control the inflow population, and surveillance and tighter controls in each neighborhood. These shutdowns occurred at the province, prefect, and district level. Our data source does not include the end date for these policies.

Region level	n	date_start_earliest	date_start_latest
district	14	2020-02-02	2020-04-08
prefecture	27	2020-02-01	2020-02-07
province	2	2020-02-06	2020-02-16

Objective: We are looking for your support to identify the end date of partial lockdowns in these provinces, prefectures, and districts.

Tasks

1. Divide the list of partial shutdowns between each of you
2. Using available news sources, government announcements, and/or WeChat local subscriptions, identify the end date for each partial lockdown identified in the attached dataset: partial-lockdowns-for-ras-20201015.csv
3. In addition to capturing the date of the end of the partial lockdown, also record the source that supports this date in a column titled "source" in the spreadsheet
4. Review the partial shutdowns collected by the other RA

Note: If, in the process of collecting this information you identify other districts, prefectures, or provinces that underwent partial lockdowns, add these as new rows to the dataset

Timeline: The estimated time for this task is 15 hours (45 identified policies X 20 min per policy for research and verification). Given this time estimate, by when do you think you can complete a rough draft?