Targeting the Corps: The Effectiveness of Humanitarian Sanctions in China

A Data Management Plan created using DMPonline.be

Creators: Kam Pui Tsang, n.n. n.n.

Affiliation: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Principal Investigator: n.n. n.n.

Data Manager: Kam Pui Tsang

Project Administrator: Kam Pui Tsang

Grant number / URL: GPUCL/22/042

ID: 201470

Start date: 01-10-2022

End date: 30-09-2026

Project abstract:

The proposed PhD project seeks to quantify the effectiveness of humanitarian sanctions in curbing alleged human right abuses by targeting the economic interests of political elites. We will study the recent Global Magnistksy Act sanction regime imposed on the Xinjiang Production and Construction Corps (XPCC) in China in September 2020 by the US and later joined by the EU and the UK. The sanction regime imposed financial sanctions on firms directly or indirectly owned by the XPCC and individuals linked to these firms and was followed by an import ban on specific products produced by these firms. The supervisory team combines a mix of expertise in the political economy of sanctions, international trade, and development economics, together with extensive research experience on China. Such a complementarity among the supervisory team paves the way for significant contributions along several lines of inquiry and on different output dimensions. By exploiting detailed financial and accounting firm-level data, the project will quantify the ability of such sanctions to cause economic harm on targeted firms (compared to non-targeted firms). Next, using firm-level custom transactions and product-level UN Comtrade data we will assess the trade consequences of the sanction regime. Finally, innovative methods of natural language processing will be used to study how sanctions affect the reporting by state-controlled news media with respect to the Xinjiang’s minorities or government interventions. Such analysis has the potential to indirectly explore how sanctions influence policy decisions of China’s elites.

Last modified: 25-08-2023

Targeting the Corps: The Effectiveness of Humanitarian Sanctions in China

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
China Stock Market & Accounting Research Database (CSMAR)	Chinese Stock market data at the firm level	Reuse existing data	Digital	Observational	.dta (Stata)	< 100 GB	*not applicable*
UNComtrade Data	Trade data at the product (HS6) level	Reuse existing data	Digital	Observational	.dta (Stata)	< 100 GB	*not applicable*
China's Customs data	Chinese customs data at the product level	Reuse existing data	Digital	Observational	.dta (Stata)	< 100 GB	*not applicable*

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

China Stock Market & Accounting Research Database: https://cn.gtadata.com/
UNComtrade Data:https://comtrade.un.org/
China's Customs data:http://stats.customs.gov.cn/

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- No

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

Kampui Tsang, a contributor of this project, is affiliated with Hong Kong Baptist University. He has access to and permission to use the China Stock Market & Accounting Research Database. However, the unauthorized copying, exploitation, and dissemination of this data are strictly prohibited under the User License Agreement between the data provider, GTA Information Technology Co., and Hong Kong Baptist University.
Kampui Tsang has registered for a free account on United Nations COMTRADE and utilized the API provided by UNComtrade to download the data. It should be noted that the usage and distribution of this data is subject to restrictions and requires the prior written permission of the United Nations.
Chinese Customs data is publicly accessible.

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

The data providers for the three mentioned datasets offer comprehensive documentation on their data products, which includes the compilation process and the codebooks for variables.
For our project, we will predominantly utilize the statistical software STATA to handle data cleaning and perform data analysis. To enhance communication among collaborators and others, we will incorporate comments in our program files (also known as "do files") for each step. These program files will enable us and others to: (i) reconstruct the dataset from the raw data files, (ii) replicate our estimation results, and (iii) generate the corresponding tables and figures as presented in our paper.
Since we will be generating multiple program files, we will create corresponding README files. These README files will serve to document the purpose of each program file and specify the order in which they should be executed. Additionally, we will provide an auxiliary codebook that documents the definitions of the variables used in our analysis.

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:

If no, please specify (where appropriate per dataset or data type) which metadata will be created:
We will create and provide README files, a list of the programs that have been created and utilized, as well as an auxiliary codebook for variables.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- OneDrive (KU Leuven)

All datasets will be stored on OneDrive. Given that all three datasets have sizes below 100 GB, 2 TB of storage space within the OneDrive environment should provide more than enough capacity for data storage and backup.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

Files stored in OneDrive benefit from automatic version management, which serves as a backup mechanism.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- No (explain solution below)

As mentioned, given that all three datasets have sizes below 100 GB, 2 TB of storage space within the OneDrive environment should provide more than enough capacity for data storage and backup.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

To maintain data security, the datasets are stored on OneDrive in a manner that ensures restricted access. Only individuals with the shared link and encryption key can access the datasets. This ensures that only authorized collaborators have the necessary credentials to access the folder containing the datasets.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

Costs are negligible.

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

All datasets will be stored on OneDrive. If the OneDrive service is terminated, the datasets will be transferred to Professor Van Biesebroeck's desktop computer. This computer is mirrored on a network drive, ensuring that a backup copy of the datasets is securely stored.

Where will these data be archived (stored and curated for the long-term)?

- Other (specify below)

As specified earlier, all data will be stored on Professor Van Biesebroeck's desktop computer until his expected retirement in the year 2040. This ensures that the data remains accessible and securely stored throughout his tenure.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

Costs are negligible.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- No (closed access)

If access is restricted, please specify who will be able to access the data and under what conditions.

Only collaborators on the project can access the data.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- Yes, intellectual property rights

As previously stated, dissemination of all three datasets is prohibited under corresponding user licence agreements.

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other (specify below)

Not applicable.

When will the data be made available?

- Other (specify below)

Not applicable**.**

Which data usage licenses are you going to provide?

If none, please explain why.

- Other (specify below)

As previously stated, dissemination of all three datasets is prohibited under corresponding user licence agreements.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- No

What are the expected costs for data sharing? How will these costs be covered?

Not applicable.

Responsibilities

Who will manage data documentation and metadata during the research project?

Prof. Jo Van Biesebroeck will have overall responsibility. He will instruct all collaborators how to document the statistical programs generated during this project
No new data will be generated.

Who will manage data storage and backup during the research project?

Prof. Van Biesebroeck will be responsible for all data stored on OneDrive.

Who will manage data preservation and sharing?

After completion of the project, responsibility will revert to Prof. Jo Van Biesebroeck.

Who will update and implement this DMP?

The PI bears the end responsibility of updating & implementing this DMP.