Plan Overview

A Data Management Plan created using HKUL DMPTool

Title: When the Collusive Bargain Breaks: State-Led Restructuring and the Resolution of China's Real Estate Crisis

Creator: Xiaoping Wu

Affiliation: The University of Hong Kong

Template: HKU Template

Project abstract:

Despite a systemic collapse in China's real estate sector, where developers responsible for 40% of home sales have defaulted, a "Lehman Moment" has been conspicuously averted. How did China manage this industry-wide bargaining deadlock? Based on semi-structured interviews with key stakeholders—including judges, government officials, and homebuyers—this study analyzes the institutional origins of the crisis. This paper posits that the crisis is rooted not in simple market failure, but in the inevitable collapse of the dominant informal, collusive bargain — namely, "local state corporatism" and "government-business collusion"—that underpinned the "land finance" model. This collusive contract systematically bypassed formal law (particularly pre-sale fund regulations), endogenously created systemic moral hazard, and externalized this risk onto homebuyers. The formal Enterprise Bankruptcy Law (EBL) proved functionally inert because it is an ex-post remedy attempting to resolve a crisis caused by ex-ante, state-sanctioned opportunism. This paper further argues that, as the "arsonist" in the collusive contract, the local state, acting out of political self-preservation, must become the "firefighter". This solution replaces the old cooperative informal bargain with a new coercive relational governance model. This solution is operationalized through the "Special Task Force" (zhuanban), an administrative body that functions as a de facto "court of equity". The Task Force subverts the legal "Absolute Priority Rule" (APR) , instead enforcing a three-tiered political hierarchy of claims centered on "guaranteed housing delivery" (bao jiaolou). It employs a powerful administrative toolkit— including information dominance , the "shadow of the future", and existential threats against developers —to enforce a "guided compromise", compelling "strategic burden-bearers" like banks and state-affiliated entities to absorb the ultimate losses.

Start date: 06-07-2025

End date: 12-31-2026

Last modified: 11-25-2025

Copyright information:

When the Collusive Bargain Breaks: State-Led Restructuring and the Resolution of China's Real Estate Crisis

Data Collection

What data will you collect or create?

The project will primarily collect and generate qualitative data derived from field research and documentary analysis regarding China's real estate crisis resolution. The data consists of two main categories:

1. Primary Data: Semi-Structured Interviews

Content: The core dataset comprises in-depth, semi-structured interviews with key stakeholders involved in real estate restructuring.

Participants: The target sample size is approximately 100 interviews (with 46 currently completed). The interviewees include five distinct groups: (1) Legal professionals (judges, lawyers, administrators); (2) Government officials (housing, finance, planning departments); (3) Financial creditors (bank and AMC executives); (4) Corporate insiders (developer management); and (5) Homebuyers involved in protests.

Scope: Fieldwork covers key regions including Guangdong, Fujian, Henan, Beijing, and Shanghai.

Data Formats:Researcher's observational notes and memos generated during and after the interviews.

2. Secondary Data: Documentary Evidence

Content: A comprehensive collection of documents used to triangulate and corroborate interview findings.

Types of Documents:

Official Policy Documents: Directives from central and local governments and regulatory bodies (e.g., the "Three Red Lines" policy).

Legal and Corporate Filings: Court rulings/documents from bankruptcy cases, corporate bond prospectuses, annual reports, and stock exchange announcements from defaulted developers.

Internal Materials: Non-public documents provided by interview participants, such as minutes from

"Special Task Force" meetings and internal risk assessment reports.

Media and Industry Reports: In-depth reports from financial media (e.g., Caixin) and industry analysis.

Data Formats: These will largely be stored as digital text files (.pdf, .docx) or image files (.jpg, .png) for archival documents.

How will the data be collected or created?

1. Primary Data Creation: Semi-Structured Interviews

Sampling Strategy: Given the elite and insular nature of the target population (judges, government officials, and developers), the project employs a purposive sampling strategy to identify initial key stakeholders involved in significant default cases. This is followed by snowball sampling, leveraging trust-based referrals to access opaque decision-making circles, such as members of "special task forces".

Collection Protocol: Data will be generated through in-depth, semi-structured interviews. To ensure data consistency, a core set of questions will be asked of all participants to allow for cross-case comparability. Simultaneously, the protocol allows flexibility to probe specific themes relevant to the participant's unique expertise (e.g., specific coercive tactics used by officials or asset-hiding strategies by developers).

Fieldwork Logistics: Interviews are conducted on-site in key geographic locations selected for their specific relevance: Guangdong and Fujian (hubs for distressed private developers), Henan (epicenter of "rotten-tail" building protests), and Beijing/Shanghai (centers for financial policy and corporate headquarters).

2. Secondary Data Collection: Documentary Acquisition

Sourcing: The project will collect existing datasets to corroborate informal narratives. Publicly available data (court filings, bond prospectuses, annual reports) will be downloaded from official repositories and stock exchange databases.

Internal Documents: A distinct subset of data will be collected directly from interview participants. This includes non-public internal materials such as minutes from task force meetings and internal corporate risk assessments.

Triangulation: These documents will be systematically cataloged to verify the timeline of events and the specific legal/financial mechanisms described by interviewees.

Documentation and Metadata

What documentation and metadata will accompany the data?

To ensure the data is interpretable by future researchers and to maintain the integrity of the study, the following documentation and metadata will be generated:

1. Documentation for Primary Data (Interviews)

Given the sensitive nature of the research (involving "extra-legal pressures" and "coercive bargaining"), rigorous documentation is essential to separate identifiable information from the analytical data:

Interview Protocol: A copy of the semi-structured interview guide will be preserved. This includes the "consistent set of core questions" used across all 100 planned interviews to ensure comparability.

Anonymization Codebook: A strictly confidential key (stored separately from the data) will link anonymized Participant IDs (e.g., "Gov-Official-01") to specific individuals. This allows the researcher to track the "snowball sampling" chains without exposing participants to political risk.

Methodological Memos: Field notes and memos documenting the specific context of the interviews (e.g., the "tense standoff" atmosphere in specific locations like Xi'an or Henan) will accompany transcripts to provide necessary contextualization for the raw text.

2. Metadata Standards All digital files will be tagged with standardized metadata to facilitate organization and retrieval without compromising privacy:

For Interview Transcripts:

Participant Category: Tagged according to the five stakeholder groups identified in the study design: (1) Judges/Court Staff, (2) Government Officials, (3) Bank/Financial Personnel, (4) Developer Staff, or (5) Homebuyers.

Geographic Tag: The region of the fieldwork (e.g., Guangdong, Fujian, Henan, Beijing, Shanghai).

Role/Level: Descriptive tags indicating the participant's hierarchy (e.g., "Deputy Mayor level" vs. "District level").

Date of Collection: To correlate findings with external timeline events (e.g., pre- or post- the "Three Red Lines" policy shock).

For Secondary Documents:

Document Type: Classified by source material type: "Official Policy," "Court Filing," "Internal Meeting Minute," or "Media Report" .

Source Origin: e.g., "Supreme People's Court," "Evergrande Group," or "Shaanxi Banking and Insurance Regulatory Commission".

Key Themes: Tags relating to the theoretical framework, such as "Special Task Force," "Guaranteed Delivery," or "Asset Stripping".

Ethics and Legal Compliance

How will you manage any ethical issues?

1. Formal Ethics Approval

I have submitted a formal application for Human Research Ethics approval, which is currently under review. The application reference number is EA250672. All data collection activities will strictly adhere to the conditions and protocols approved by the Ethics Committee prior to commencement.

2. Informed Consent and Participant Safety

Given the study analyzes sensitive political dynamics, including "extra-legal pressures" and "coercive bargaining tactics", protecting participants is the primary ethical obligation. The participant pool includes vulnerable groups (dispossessed homebuyers involved in protests) and political/economic elites (government officials, judges, developers) who face potential professional or legal risks.

Consent Process: Informed consent will be obtained from all interviewees prior to their participation. The consent process will be conducted in strict compliance with the protocols approved by the Ethics Committee (Ref: EA250672), ensuring that all participants are fully informed of the study's nature, their rights, and the voluntary basis of their involvement.

Right to Withdraw: Participants will be clearly informed of their right to withdraw from the study or refuse to answer specific questions regarding sensitive "internal materials" or decision-making processes.

3. Anonymity and De-identification

To mitigate political and social risks, strict anonymity will be maintained for all individuals and specific organizations (where necessary).

De-identification: Direct identifiers (names, specific job titles) will be removed from transcripts and field notes immediately upon processing. Participants will be assigned alphanumeric codes (e.g., "Judge-01", "Developer-03") to distinguish stakeholder roles without revealing identities.

Redaction of Sensitive Locations: While the study covers specific provinces like Guangdong and Henan, specific project names or local task force details that could trace back to a single official will be generalized in the final output to prevent "deductive disclosure."

4. Data Security and Confidentiality

Secure Storage: All raw data, including audio recordings and "internal documents" provided by participants (e.g., meeting minutes), will be stored on password-protected, encrypted drives.

Access Control: Access to the raw, non-anonymized data (such as the key linking codes to names) will be restricted solely to the principal investigator. This is critical for maintaining the "trust" required to secure referrals via the snowball sampling method.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

The project involves both the generation of original research data and the collection of third-party materials. IP rights will be managed as follows:

1. Primary Data: Interviews and Field Notes (IP Creation)

Ownership: Copyright in the transcripts and field notes generated during the "semi-structured interviews with key stakeholders" will be retained by the researcher (and the University, where applicable under institutional policy).

Participant Rights: Through the informed consent process (Ethics Ref: EA250672), participants will grant the researcher a license to use their spoken contributions for the purpose of this study and resulting academic publications.

Anonymization: As the data includes sensitive accounts from "government officials" and "corporate insiders", the raw data containing identifiable IP will be strictly sequestered. Only anonymized extracts will be used in public outputs to prevent any potential claim of breach of confidence or unauthorized disclosure.

2. Secondary Data: Public and Corporate Documents (Third-Party IP)

Public Domain/Open Access: "Official Documents" such as policy directives and "Legal Filings" from courts are generally considered public records. These will be used for analysis and cited according to standard academic conventions.

Copyrighted Media: "Media and Industry Reports" (e.g., Caixin, Caijing) are subject to third-party copyright. These materials will not be reproduced in full. They will be used solely for "triangulation" and analysis under the "Fair Dealing" (or Fair Use) provisions for non-commercial research and criticism.

3. Management of Restricted/Internal Materials

Sensitive IP: The project involves the collection of "Internal Materials" provided by participants, such as "minutes from task force meetings" or "internal risk assessments".

Usage Protocol: It is recognized that the IP for these documents resides with the originating organizations (e.g., the developer or local government). Consequently:

These documents will not be republished or shared in their entirety.
They will be used exclusively for internal analysis and corroboration.
Any reference to these materials in published work will be heavily redacted and summarized to protect the source and avoid infringing on proprietary corporate information.

Storage and Backup

How will the data be stored and backed up during the research? i. e. until stored in the final location (e.g. on your password protected laptop)?

In compliance with the University's Information Security and Data Management (ISDM) Policy and the Policy on the Management of Research Data and Records, the project will implement a rigorous storage and backup protocol. Given that the data involves sensitive political information and internal documents, it is classified as "Confidential/Restricted" and will be handled as follows:

1. Active Storage Strategy (During Fieldwork & Analysis)

Fieldwork Phase: During data collection in Guangdong, Fujian, Henan, Beijing, and Shanghai, data (audio recordings, field notes) will be initially stored on a dedicated, password-protected laptop with full-disk encryption (e.g., BitLocker or FileVault).

Campus Phase: Upon return or when secure internet is available, all data will be transferred to HKU OneDrive for Business (the university-sanctioned cloud storage). This platform provides enterprise-grade security and access logs, unlike personal cloud services (e.g., Google Drive), which are strictly prohibited for this sensitive dataset.
Encryption: All individual files containing identifiable participant information or sensitive internal materials will be further encrypted using AES-256 standard before any network transfer.

2. Backup Strategy (The 3-2-1 Principle)

To prevent data loss due to hardware failure or theft during travel, the project will strictly adhere to the 3-2-1 Backup Principle:

3 Copies: We will maintain three complete copies of the dataset (1 Primary + 2 Backups).
2 Different Media:

Medium A: The primary encrypted laptop hard drive.

Medium B: An external, encrypted solid-state drive (SSD) kept physically separate from the laptop (e.g., in a hotel safe) during fieldwork.

1 Off-site Location: The third copy will be synced to the HKU OneDrive cloud environment. This ensures that even if local equipment is seized or lost during the fieldwork, a secure remote copy remains accessible.

Frequency: Backups to the external drive will occur daily during fieldwork weeks. Cloud synchronization will occur whenever a secure, private network connection is established.

3. Access Control and Security

Authorization: Access to the raw data is strictly limited to the Principal Investigator. No other research assistants or unauthorized personnel will hold decryption keys.

Transmission: Data will never be transmitted via unencrypted email or public Wi-Fi networks.

Retention: In accordance with HKU policy, research data will be retained for 5 years following publication. Long-term preservation will be managed via the HKU DataHub repository upon project completion.

How will you manage access and security?

Given the highly sensitive nature of the data, which includes "internal materials" and interviews regarding "coercive bargaining", access and security will be managed through a strict, multi-layered protocol compliant with HKU's ISDM Policy:

1. Access Control (The Principle of Least Privilege)

Principal Investigator (PI) Only: Full access to the raw, non-anonymized data (including audio recordings and the "Master Key" linking participant codes to identities) will be restricted exclusively to the Principal Investigator.
Authentication: Access to the storage environment (HKU OneDrive) will be protected via HKU Portal UID and mandatory Two-Factor Authentication (2FA) to prevent unauthorized login attempts.
Collaborators/Supervisors: If data sharing is required for supervision or analysis, only fully anonymized transcripts (stripped of names, specific locations, and identifiable internal document references) will be shared. The "Master Key" file will never be shared or transmitted.

2. Technical Security Measures (Encryption & Network)

Data at Rest: All devices (laptops, external SSDs) used during fieldwork in Guangdong, Henan, etc. will utilize Full Disk Encryption (e.g., BitLocker for Windows or FileVault for macOS) with strong, complex passwords. Individual sensitive files (e.g., the Master Key) will be further encrypted using AES-256 standard (e.g., via 7-Zip or Veracrypt).
Data in Transit: Data transfer will strictly occur via HKU VPN (Virtual Private Network) to ensure an encrypted tunnel when accessing university servers. No data will be transmitted over public Wi-Fi or unencrypted email.
Device Security: Automatic screen locking (set to 5 minutes of inactivity) and remote-wipe capabilities will be enabled on all field devices to mitigate risks in case of theft or loss.

3. Physical Security (Fieldwork Protocols)

Device Control: During fieldwork, portable storage devices (USBs/SSDs) containing research data will be kept on the researcher's person or locked in a secure safe when not in use. They will never be left unattended in hotel rooms or public workspaces.

Clean Desk Policy: No physical notes, consent forms, or printed "internal documents" will be left visible or accessible to third parties. Physical documents will be digitized and encrypted at the earliest opportunity, after which the physical copies will be securely destroyed (shredded) or stored in a locked facility if retention is required by Ethics protocols.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

1. Data Selected for Long-Term Retention and Sharing The following data are deemed of high value and will be retained to preserve the scientific record and allow for future verification of the study's findings:

Anonymized Interview Transcripts: These constitute the project's primary contribution, offering rare, "micro-level evidence" of the decision-making processes within the "Special Task Forces". The transcripts capture the distinct perspectives of "judges, government officials, banking executives, and developer management" regarding the "multi-party bargaining collapse". These de-identified texts are essential for understanding the "informal institutions" and "collusive bargains" that replaced formal law during the crisis.

Corroborating Documentary Evidence: The collection of secondary data, including "official policy directives," "court filings," and redacted summaries of "internal materials" (such as meeting minutes) , must be retained. These documents are critical for "triangulating" the informal narratives provided by interviewees and substantiating the existence of the "parallel administrative system".

2. Data Excluded from Retention (To be Destroyed)

While valuable for immediate analysis, certain data subsets pose an unacceptable risk to participants if retained long-term. To adhere to ethical obligations regarding "sensitive political processes" and "existential threats" to stakeholders, the following will not be retained or shared:

Raw Audio Recordings: Due to the risk of voice identification in a politically sensitive context, raw audio files will be destroyed immediately after the accuracy of the transcripts is verified.

Participant Identity Keys: The codebook linking anonymized IDs to specific individuals (e.g.,

specific officials or developers) will be destroyed at the end of the project's mandatory retention period to ensure permanent anonymity.

What is the long-term preservation plan for the dataset?

1.Preservation Strategy (HKU DataHub)

Repository: The final, fully anonymized dataset (transcripts and non-proprietary documents) will be deposited in HKU DataHub, the university's institutional repository. This ensures the data is preserved in a secure, DOI-indexed environment.

Retention Period: In compliance with HKU's Policy on the Management of Research Data and Records, the data will be retained for 5 years following the project's completion or the publication of findings.

File Formats: To ensure long-term accessibility, files will be converted to open, non-proprietary formats for preservation (e.g., .txt or .pdf/A for text, .csv for any quantitative logs).

2. Data Sharing and Access Level

Restricted Access: Given the sensitive nature of the "internal materials" and the potential for "deductive disclosure" of specific local government officials, the dataset deposited in HKU DataHub will be set to "Restricted Access" (or "Mediated Access").

Request Protocol: Future researchers wishing to access the data must submit a request to the Principal Investigator, detailing their intended use and ethical approval. Access will be granted only to bona fide researchers who agree to strict confidentiality terms, ensuring that the "trust" established with original participants is not violated.

Data Sharing

How will you share the data?

1. No Public Access

Due to the political sensitivity of the research, which documents "coercive bargaining tactics" and "existential threats" against stakeholders, the dataset will not be made openly accessible to the public. Unrestricted sharing would compromise the "trust" upon which access to the "elite and

insular population" was conditioned and could expose participants to significant professional or legal risks.

2. Case-by-Case Sharing Protocol Data sharing will be managed through a strict, case-by-case assessment process.

Request Mechanism: Other researchers wishing to verify findings or conduct secondary analysis must submit a formal request to the Principal Investigator.

Cautious Review: Each request will be rigorously vetted. Access will be granted only if the requester can demonstrate a bona fide academic purpose and provide evidence of their own institutional ethics approval.

Legal Agreements: Approved researchers will be required to sign a strict non-disclosure or data use agreement, ensuring they will not attempt to re-identify the "government officials" or "corporate insiders" involved.

3. Limitations on Shared Content Even in approved cases, sharing will be exercised with extreme caution:

Anonymized Transcripts Only: Only fully de-identified transcripts may be shared.

Exclusions: Raw audio recordings and sensitive "internal materials" (e.g., task force meeting minutes) will never be shared due to the high risk of deductive disclosure and the proprietary nature of the documents.

Are any restrictions on data sharing? If yes, Why?

1. Nature of Data Sharing Restrictions Data sharing will be subject to strict restrictions. The dataset will not be made publicly available (Open Access). Instead, a "Mediated Access" model will be employed, where data is shared strictly on a case-by-case basis following a rigorous review of the requester's credentials and intent.

2. Justification for Restricted Access

Political Sensitivity and Participant Safety: The research documents sensitive state behaviors, including "extra-legal pressures" and "coercive bargaining tactics" used against developers and officials. Unrestricted access could expose participants—particularly government officials and corporate executives—to significant professional or legal risks, such as "criminal investigation" or administrative punishment.

Preservation of Trust: Access to this "elite and insular population" was secured solely through "snowball sampling" based on deep personal "trust". Making data publicly available would violate the confidentiality agreements implicit in these referrals and could jeopardize the researcher's professional reputation and future access.

Prevention of Deductive Disclosure: Even with names redacted, specific details regarding a participant's role (e.g., a "Vice Mayor" in a specific "Special Task Force") could allow knowledgeable insiders to deduce identities. Strict control is necessary to prevent this type of "jigsaw identification."

3. Evaluation Criteria for Access Requests Requests for data access will be evaluated by the Principal Investigator against the following strict standards:

Criterion 1: Bona Fide Academic Status

The requester must be a verifiable academic researcher affiliated with a recognized institution. Requests from journalists, commercial analysts, or private intelligence entities will be denied to prevent the data from being used for non-academic purposes that could endanger participants.

Criterion 2: Legitimate Scientific Purpose

The request must articulate a clear, valid scientific objective, such as verifying the study’s findings on the "Great Gridlock" or "state-led restructuring". Broad or undefined requests will be rejected.

Criterion 3: Ethical Reciprocity

The requester must demonstrate adherence to ethical standards equivalent to this project (Ethics Ref: EA250672). They will be required to sign a Data Use Agreement (DUA) explicitly prohibiting any attempt to re-identify participants or contact them.

Criterion 4: Data Minimization

Access will be granted strictly to the minimum subset of data required for the specific inquiry (e.g., only specific anonymized transcripts), excluding any sensitive "internal materials" or "meeting minutes".

Responsibilities and Resources

Who will be responsible for data management?

The PI.

What resources will you require to deliver your plan?

1. Hardware and Physical Storage

Dedicated Field Laptop: A dedicated research laptop enabled with Full Disk Encryption (BitLocker/FileVault) to serve as the primary station for data entry and storage during fieldwork in Guangdong, Henan, and other provinces.

Encrypted External Storage: Two high-capacity, encrypted solid-state drives (SSDs) to facilitate the "3-2-1 backup strategy" (offline backups) while conducting research in locations with unstable internet access.

Audio Recording Equipment: High-fidelity digital voice recorders (non-networked) to capture semi-structured interviews.

2. Software and Digital Infrastructure

Cloud Storage: Access to HKU OneDrive for Business (institutional subscription) to serve as the secure, off-site repository for data synchronization and long-term storage.

Encryption Tools: Utilization of standard encryption software (e.g., Veracrypt or 7-Zip) to apply AES-256 protection to sensitive files, particularly the "internal materials" provided by participants.

Analysis Software (CAQDAS): Licenses for qualitative data analysis software (e.g., NVivo or Atlas.ti) to organize, code, and analyze the large volume of interview transcripts and documentary evidence.

3. Financial and Logistical Resources

Fieldwork Budget: Funding to cover travel and accommodation costs for site visits to key distressed regions (Guangdong, Fujian, Henan, Beijing, Shanghai) to ensure direct access to local stakeholders.

Transcription Costs: Resources (time or funding) allocated for the verbatim transcription of approximately 100 interviews, ensuring accuracy before the destruction of raw audio files.

4. Expertise and Support

IT Support: Consultation with the Faculty's IT or Data Steward to verify the configuration of encryption protocols and secure data transfer channels (VPN) prior to deployment.
