Codebook

Thank you for your interest in the background information pertaining to *Navigating Troubled Waters: Measuring Beijing's Use of Military Coercion in the South and East China Seas*, this codebook explains the variables and the encoding used for the creation of the dataset. For a detailed account of the methodology, the specific tests done to this data, or conclusions please see the primary text. For any questions, please contact the author at: Simon.Weiss@tufts.edu.

Date: This column details the date on which the event began. Coded on the American format of DD/MM/YYYY. Start date was defined as the day the first incident involving coercion took place. For military patrols that lasted weeks and involved an individual incident, the date of the event not the start of the patrol was used. In very few instances the start date was estimated from existing news reporting on events.

Country Involved: This column details which country was the primary recipient of the military coercion originating from the PRC. Colloqueall names of countries are used (Taiwan instead of R.O.C., etc.).

Area Involved: Within the disputed maritime borders of each of these countries several different flashpoints exist. This column says in which body of water or contested atoll the event took place in. The names used are Westernized and multiple names are not used for areas with contested naming (Senkaku vs. Diaoyu vs Diaoyutai for instance). This is done to minimize writing in the dataset.

Description: This column provides a short description of the event itself. General includes which part of the PRC military or militarized force was used in the event, and what occurred during the event. For detailed accounts of these events, please read primary source material. For the provocation data, this column rather details the events determined to be provocative for the PRC.

Duration: This is ordinal data is coded on a scale of 1-5 and categorizes the overall length of the dispute in question. Many incidents occurred over multiple days or weeks and this allows for averaging the events across time. Coding was done prior to the collection of data, based on existing work, to avoid P-hacking.

Code	Meaning
1	1 Day
2	1 Week
3	1 Month
4	6 Months
5	>6 Months/Permanent

Magnitude: This ordinal data is coded on a scale of 1-8 representing the highest act that occurred during a specific incident. This data was coded prior to the collection of data based on existing methods to avoid P-hacking. This column allows events to be compared on the strength of the coercion involved between incidents.

Code	Meaning
1	Chased away/ no immediate impact
2	Ship inspected (no damage)
3	Ship damaged (property damaged or stolen)
4	Civilian Ship Sunk
5	Military show of force
6	Targeted with weapons/water canons
7	Civilian lives lost
8	Territory seized
9	War

Source: This includes a stable link to primary source material verifying the existence of the event and the details of that event. For well publicized incidents sources may not identify all aspects of that event but will provide ways to search for other information.

Country (1-4): This column lists the countries involved with the U.S. alliance system that engaged in some form of provocative acts towards the PRC during the study period.