Data README

Data Source

Our dataset covers Chinese A-share listed firms from 2007 to 2022 and is constructed from three sources:

(1) firm-level and financial data from the China Stock Market & Accounting Research (CSMAR) database (https://global.csmar.com);

(2) corporate environmental misconduct data from the CCER™ database provided by Sinofin Financial Information and the China Center for Economic Research at Peking University (see https://www.ccerdata.cn);

(3) manually collected information on the establishment of environmental courts from websites of intermediate people's courts in each prefecture-level city and relevant news reports.

Software and Packages

(1) Software: Stata (SE 19)

(2) Packages

Before running the replication code, install the following packages:

ssc install reghdfe, replace

ssc install coefplot, replace

ssc install psmatch2, replace

ssc install ppmlhdfe, replace

Data Merging and Cleaning

(1) Construction of the baseline firm panel

We begin by extracting all non-financial A-share firms from the CSMAR database. Following standard practice, we exclude firms designated as *ST or ST due to abnormal financial status.

(2) Matching environmental misconduct records

Environmental misconduct data are merged with firm-level observations

using the firm's ticker symbol.

(3) Matching environmental court establishment data

To identify treatment exposure, we match firms to the court-establishment dataset based on the firm's registered office city reported in the CSMAR database.

(4) Cleaning

Continuous variables are winsorized at the 1st and 99th percentiles, and monetary variables are deflated using the GDP deflator.

Variable Construction

(1) Dependent variable

- ✓ Corporate environmental misconduct: the number of reported corporate environmental misconduct cases committed by the firm in year t+1.

(2) Independent variable

- ✓ Environmental court: equal to 1 if an environmental court was established in the city where the firm is located in year t and onwards, and 0 otherwise.

(3) Control variables

- ✓ Firm Size: the natural logarithm of total assets.
- ✓ Leverage: total liabilities divided by total assets.
- ✓ Tobin Q value: the ratio of market value to total assets.
- ✓ Largest shareholder ownership: the shareholding percentage of the largest shareholder.
- ✓ Board size: the natural logarithm of one plus the number of directors on the board.
- ✓ Return on assets: net profits divided by total assets.
- ✓ Firm age: the natural logarithm of the number of years since the firm's initial public offerings plus 1.
- ✓ Ownership: equal to 1 if the firm is state-owned and 0 otherwise.

(4) Variables used in mechanism analyses

- ✓ Environmental attention allocation in the review section: the ratio of Chinese characters in sentences with environmental keywords to the

total number of characters in the review section

- ✓ Environmental attention focus in the review section: the difference in frequency between actual governance keywords and conceptual propaganda keywords, normalized by the total frequency of all environment-related keywords in the review section
- ✓ Environmental attention allocation in the goal section: the ratio of Chinese characters in sentences with environmental keywords to the total number of characters in the goal section
- ✓ Environmental attention focus in the goal section: the difference in frequency between actual governance keywords and conceptual propaganda keywords, normalized by the total frequency of all environment-related keywords in the goal section
- ✓ Environmental expenditure: the ratio of fiscal expenditure on energy conservation and environmental protection to total fiscal expenditure
- ✓ Fiscal pressure: the ratio of fiscal revenue to expenditure
- ✓ Industrial structure: the share of the secondary industry in GDP
- ✓ Economic development: the natural logarithm of per capita GDP
- ✓ Industrial pollution: the natural logarithm of industrial SO2 emissions
- ✓ Total green patent applications: the number of green patent applications in year t+1
- ✓ Independent green invention patent applications: the number of green invention patents applied independently by the firm in year t+1
- ✓ Independent green utility model patent applications: the number of green utility model patents applied independently by the firm in year t+1
- ✓ Joint green invention patent applications: the number of green invention patents applied jointly by the firm in year t+1
- ✓ Joint green utility model patent applications: the number of green utility model patents applied jointly by the firm in year t+1
- ✓ Total independent green patent applications: the number of green patents applied independently by the firm in year t+1
- ✓ Total joint green patent applications: the number of green patents applied jointly by the firm in year t+1
- ✓ Total green invention patent applications: the number of total green invention patents filed by the firm in year t+1
- ✓ Total green utility model patent applications: the number of total green utility model patents filed by the firm in year t+1