

Data Description Sheet for
“The Economic Consequences of CEO Compensation Lawsuits”

Xiangnian Kong
Nanjing University

Ke Na
Cheung Kong Graduate School of Business

April 24, 2026

Item 1: A description of which author(s) handled the data and conducted the analyses.

Xiangnian Kong manually collected CEO compensation lawsuit data, downloaded other data, and conducted empirical analyses. Both Ke Na and Xiangnian Kong were involved in research design and interpreting the results.

Item 2: A detailed description of how the raw data were obtained or generated, including data sources, the specific date(s) on which data were downloaded or obtained, and the instrument used to generate the data (e.g., for surveys or experiments).

We collected raw data from several sources as described below:

1. We collected CEO compensation-related derivative lawsuits from the Audit Analytics (AA) litigation database and firms' 10-K and 10-Q filings, supplemented by manual searches from Factiva news search, West Laws, press releases, corporate websites, and Google.
2. We obtained accounting variables from Compustat.
3. We obtained CEO compensation and characteristics from ExecuComp.
4. We obtained stock price and return from the Center for Research in Security Prices (CRSP).
5. We obtained board characteristics, including board independence and board size, from the ISS Directors database, and director co-option data from Professor Lalitha Naveen's research website (<https://sites.temple.edu/lnaveen/data/>).
6. We collected lawsuit filing dates from AA litigation database and 10-K and 10-Q filings and official CEO departure dates from 8-K filings. We then estimated abnormal stock returns surrounding litigation filings and CEO turnover announcements via Event Study by WRDS.
7. We extracted information regarding compensation peers from Incentive Lab, ISS Executive Compensation Peers, and proxy statements.
8. We collected shareholder support levels for Say-on-Pay proposals from firms' Form 8-K filings.
9. We collected data for firms not covered by standard databases from SEC filings.
10. All data were collected between December 2019 and December 2024.

Item 3: If the data are obtained from an organization on a proprietary basis, the authors should privately provide the editors with contact information for a representative of the organization who can confirm data were obtained by the authors. The editors would not make this information publicly available. The authors should also provide information to the editors about the data sharing agreement with the organization (e.g., non-disclosure agreements, and any restrictions imposed by the organization on the authors). In particular, the authors should indicate if an organization or data provider imposes restrictions on the publication of the results, has not given the authors full control of the relevant data, requires that the results must be reviewed or approved prior to public release of the paper or publication.

All data used in the paper are either from commercial databases or manually collected from publicly accessible documents.

Item 4: A complete description of the steps necessary to download, obtain or collect as well as process the data used in the final analyses reported in the paper. For experimental and survey papers, we require information about the instructions and instruments used to generate the data, subject eligibility and/or selection, as well as any exclusion criteria. The full set of instructions and instruments can be provided in the online appendix.

The steps necessary to collect and process the data are provided in Section 3 of the paper.

Item 5: After downloading or obtaining the raw data, all manipulations of the data should be done via computer programs. The code for these manipulations should be included in the code submitted upon acceptance (see below). No manipulations of raw data can take place manually or outside the computer code provided. If compliance with this requirement is not feasible, the authors need to explain and disclose any manipulations of the raw data (e.g., manually created variables or file conversions). When feasible, we also encourage the authors to share the code that downloads the data.

All manipulations of the data were done via computer programs attached.

Item 6: The computer programs (i.e., code) used to (1) convert the raw data into the final dataset used in the analysis, (2) to execute the statistical or econometric analysis, and (3) to generate the tables or to produce the output used in constructing tables of the manuscript. A brief description that enables other researchers to understand and run the code should be provided. The purpose of this requirement is to facilitate replication and to help other researchers understand in detail how the raw data were processed, the final sample was formed, variables were defined, outliers were treated, and which commands were used in the analysis, etc. This code or programming is in most circumstances not proprietary. However, we recognize that some parts of the code or data generation process may be proprietary, including from the authors' perspective. Therefore, instead of disclosing the proprietary portion of the code or program, researchers can provide a detailed step-by-step description of the code or the relevant parts of the code such that it enables other researchers to arrive at the same results that the authors obtained and presented in their manuscript. In such cases, the authors should inform the editors upon initial submission, so that the editors can consider an exemption allowing the step-by-step description. Whenever feasible, authors are required to provide the identifiers (e.g., CIK, CUSIP) for their final sample. Authors should consult our FAQ Sheet on the JAR website for further details.

The computer programs used to convert the raw data, to execute the statistical analysis, and to generate the tables, and a brief description of the computer programs are provided.

Item 7: A comprehensive log file that shows the execution of the entire code. This log file should cover all the steps that convert the raw data into a final dataset and the execution of all statistical and econometric analyses presented in the tables of the manuscript. The portion of the log file that shows proprietary code or data may be masked. In this case, the reader should be referred to the step-by-step description provided as per the requirements in Item 6.

A comprehensive log file showing the execution of the entire code is provided.

Item 8: An assurance that the data and programs will be maintained by at least one author (usually the corresponding author) for at least six years, consistent with National Science Foundation guidelines.

The data and programs will be maintained by Xiangnian Kong for at least six years, in line with NSF guidelines.