

Public Disclosure of Private Meetings: Does Observing Peers' Information Acquisition Affect Analysts' Attention Allocation?

Yi Ru

Renmin University of China
ruyi@rmba.ruc.edu.cn

Ronghuo Zheng

The University of Texas at Austin
Ronghuo.Zheng@mcombs.utexas.edu

Yuan Zou

Harvard Business School and D³ Institute
yzou@hbs.edu

Compliance with Data Policy for the Journal of Accounting Research

January 2025

The following memo provides information requested under the *Journal of Accounting Research* (JAR) Data Policy.¹

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

One of the authors, Yi Ru handled the data and conducted the analyses. All authors are involved in the research design and interpreting the results.

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 recommend that more than one author can vouch for the stated source of the raw data.*

We obtained the data from a variety of sources, including China Stock Market & Accounting Research Database (CSMAR), Chinese Research Data Services (CNRDS), RESSET, and Corporate Site Visit Database (CSVD) developed by Datago Technology Limited. In addition, for site visits before 2012, we also hand collect site visit records from firm semi-annual and quarterly reports that are not disclosed in the annual reports.

CSMAR data

We use the data on corporate site visits after July 2012 from CSMAR, which collects the records of site visits from the *HudongYi* platform. We also use the data on firm characteristics and analyst characteristics from CSMAR. We downloaded the CSMAR data in July 2019. In addition, we use the data from CSMAR combined with analyst reports and web search to construct the indicator

¹ <https://onlinelibrary.wiley.com/page/journal/1475679x/homepage/forauthors.html>

variable to capture whether an analyst is connected to a firm. We collected data to construct this variable from August 2021 to November 2021.

CNRDS data and RESSET data

We use analyst forecast data from three databases: CSMAR, CNRDS, and RESSET. We downloaded CNRDS and RESSET data from June 2019 to August 2020.

CSVD data

We use the data on corporate site visits before July 2012 from CSVD, which collects the records of site visits disclosed in firms' periodic reports. We downloaded the CSVD data in August 2020. In addition, for site visits before 2012, we also hand collect site visit records from firm semi-annual and quarterly reports that are not disclosed in the annual reports.

All authors vouch for the stated source of the raw data. All authors have access to the data.

- 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 of the data came from publicly available sources. There are no restrictions on the use of the data we have gathered.

- 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.*

We described all relevant steps in Section 2.3 of the manuscript, including an appendix with variable definitions.

- 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.*

After downloading and obtaining the raw data, all manipulations of the data are done via computer programs. We have provided the code for the manipulation of raw data in the Stata program (code_RZZ.do file).

- 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.*

We have provided the programs and a brief description of how to use them in the code_RZZ.do file. Additionally, we have provided two files containing the identifiers for our main sample as follows:

- identifier_firm.dta: This file contains a list of identifiers for the 814 firms in our main sample, using *id* as firm identifier;
- identifier_analyst.dta: This file includes a list of identifiers for the 4,003 analysts in our main sample, with *brokern* representing brokers and *author* denoting analyst names.

- 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.*

We have provided the log file that shows a comprehensive execution of the entire code in the log_file_RZZ.smcl file.

- 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 authors will maintain the data and programs for at least six years.