

Buy-Side Analysts and Earnings Conference Calls

Michael J. Jung

Stern School of Business, New York University

M.H. Franco Wong

Rotman School of Management, University of Toronto

X. Frank Zhang

School of Management, Yale University

Compliance with Data Policy for the Journal of Accounting Research

July 2017

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

Franco Wong obtained the raw transcript data from Thomson Reuters StreetEvents. Michael Jung processed the data to identify buy-side analysts and ran the empirical analyses.

2. *A detailed description of how the raw data were obtained or generated, including data sources, the 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 is able to vouch for the stated source of the raw data.*

Franco Wong obtained the raw transcript data from Thomson Reuters StreetEvents in 2009.

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 agreement, any restrictions imposed by the organization on the authors with respect to publishing certain results).*

Conference call transcript data is available commercially from Thomson Reuters. Our agreement with Thomson Reuters StreetEvents prohibits us from making our list of 17,000+ identified buy-side analysts publicly available. Even though we process the raw transcript data to come up with a list of selected participants, companies, and investment firms, our list is still derived from the raw transcript data. Making the list publicly available could subject one of our coauthors to litigation. However, we have shared the list with the editor and reviewers in the review process to show its authenticity. We would be happy to share the list with any other researchers who already have an agreement with Thomson Reuters to access their raw transcript data.

4. *A complete description of the steps necessary to collect and process the data used in the final analyses reported in the paper. For experimental papers, we require information about subject eligibility and/or selection, as well as any exclusion criteria.*

Conference call transcripts were downloaded from the Thomson Reuters StreetEvents website as text files. Each file contained company name, ticker symbol, date, time, corporate executives' full name and title, conference call participants' (analysts') full name and affiliation, transcript of each executive's prepared remarks, transcript of each analyst's question, and transcript of each executive's answer.

A research assistant used a Python script to parse the individual elements of each transcript and export the data into Microsoft Excel. The Python code is contained in the file `qa_parse.docx`.

We cleaned and processed the data manually in the Excel file. Thus, there is no programming code to perform the cleaning and processing. The initial focus was on each conference call participant's (analyst's) affiliation. Because transcribers can misspell affiliations or spell them with slight differences, we sorted and visually inspected the affiliations and manually corrected them or allowed for different variants for a given buy-side investment firm, sell-side brokerage firm, or other institution. For example, "Lord Abbett" is the correct spelling for a large buy-side investment firms, and we corrected for misspellings such as "Lord Abbott," "Lord Abet," and "Lord Abett."

We took four steps to identify participants from buy-side investment firms:

1. We flagged all affiliations that contained keywords that are common in the names of institutional investment firms: "capital," "asset," "fund," "investment," "management," "advisors," "partners," "investors," and combinations such as "capital management" and "asset management." These affiliations were added to the pool of potential buy-side investment firms.
2. We flagged all affiliations known to be a sell-side brokerage firm or investment bank based on data from I/B/E/S. These affiliations were eliminated from the pool of potential buy-side investment firms. There are at least two ways to do this. The detail EPS file has the variable "ESTIMATOR" that contains a brokerage firm code and the detail recommendation file has the variable "ESTIMID" that contains a brokerage firm identifier. Many researchers working in the analyst literature have their own files that map each variable to a full brokerage name (now or defunct) through years of manually searching and matching. Such a file is available from the authors upon request.
3. Using all affiliations flagged from step a) and all affiliations that were not eliminated from step b), we manually matched the affiliations to names of institutional investors in the Thomson Reuters database of Form 13F filings. This step increased the likelihood that an analyst's affiliation is a true institutional investment firm and not a retail or individual investor. When there was a match, we retained the MGRNO variable from the

13F database. Any affiliations that are not matched were eliminated from the pool of potential buy-side investment firms.

4. Of the affiliations that remained in the pool after the above steps, we manually conducted Internet searches to exclude buy-side investment firms that also have sell-side operations. We went to each firm's website (if still active) or its Bloomberg Business online description (which includes inactive firms) to confirm its type. This step is a conservative step and ensures that the analyst is not from the sell-side.

In descriptive statistics, we counted the number of unique buy-side investment firms in our sample. As previously mentioned, firm names may have different variants (e.g., Lord Abbett vs. Lord Abbet & Co.), so we consolidated all variants into one version when computing the number of unique firms.

5. *Prior to final acceptance of the paper, the computer program used to convert the raw data into the dataset used in the analysis plus a brief description that enables other researchers to use this program. Instead of the program, researchers can provide a detailed step-by-step description that enables other researchers to arrive at the same dataset used in the analysis. The purpose of this requirement is to facilitate replication and to help other researchers understand in detail how the sample was formed, including the treatment of outliers, Winsorization, truncation, etc. This programming is in most circumstances not proprietary. However, we recognize that some parts of the data generation process may indeed be proprietary or otherwise cannot be made publicly available. In such cases, the authors should inform the editors upon submission, so that the editors can consider an exemption from this requirement.*

After finalizing the sample of analysts from buy-side investment firms, we exported the data into SAS format to merge with other databases. We used the ticker symbol and conference call date from the transcripts to match to CRSP's PERMNO identifier and then subsequently to Compustat's GVKEY identifier. The file jwz2_201717.sas creates all variables needed for subsequent analyses.

6. *Data and programs should be maintained by at least one author (usually the corresponding author) for at least six years, consistent with National Science Foundation guidelines.*

We agree to this guideline. Franco Wong holds the raw transcript data.