

Data Description Sheet

Kurt H. Gee managed all data and performed all analyses described in the paper, except as noted.

Appendix C of the paper describes our variable measurement and the data sources for each variable. Except where noted, all data from Compustat, CRSP, and I/B/E/S was updated as of March 1, 2017 by re-extracting data from these databases via WRDS, using SAS. We write programs to extract data directly from WRDS, rather than downloading and saving the data by hand at various dates, so that our sample uses the most up-to-date information from each data source, which would include correction of any data errors.

The exclusions data used in Table 7 was extracted from Compustat as of April 18, 2017.

Manager non-GAAP EPS data

To determine whether managers report non-GAAP EPS in earnings announcements, we download a list of all 8-K filings from 1993 through the beginning of 2013, based on data available on EDGAR as of July 10, 2013. We merged this list with Compustat and download all 8-Ks within four weekdays of the earnings announcement date. The 8-Ks from the list were most recently download on November 3, 2016. From the 8-Ks, we identify earnings announcements based on SEC items numbers, proximity to the earnings announcement date, and if there are ties, the proportion of earnings words. We developed a python script to search the 8-Ks for this information. We require the final 8-K earnings announcement we identify to contain GAAP diluted EPS, somewhere in the filing, in order to include the matching firm-quarter observation in the final dataset.

We developed a python script to search each earnings announcement (hereafter, “filing”) for non-GAAP keywords, and process the resulting output in SAS to construct an indicator for whether the filing contains a non-GAAP EPS disclosure. For earnings announcements that have non-GAAP EPS, research assistants read sentences that we identified as containing non-GAAP EPS numbers and collected the EPS numbers from the sentences. This hand collection took place between April 2015 and February 2018. Appendix B in the paper describes the python script and hand collection procedures in more detail.

To test the accuracy of our non-GAAP EPS data, the authors hand collected non-GAAP EPS data from the 568 observations in the Training Sample presented in Table 2. Each author completed a portion of the hand collection. This hand collection was part of an iterative process while the paper developed. Research assistants collected non-GAAP EPS data for the 211 observations in the Holdout Sample presented in Table 2. This Holdout Sample was collected in January 2014, but this collection did not have the actual non-GAAP EPS numbers. These were collected in October 2016 by research assistants. Originally, the Holdout Sample was 250 observations, but we decided to omit REITs from the sample, based on observations from the Training Sample. This reduced the Holdout Sample to 211 observations.

Sources of I/B/E/S non-GAAP exclusions

The data underlying Table 8 was hand collected by the authors in February 2017, with each author completing a portion of the hand collection.

Comparison of exclusions categories

Benjamin C. Whipple collected the data on exclusion categories presented in Figure 3, in consultation with analyst research reports from ThomsonONE. Per feedback from the review process, some of the collected observations were reclassified or dropped from the sample, so the final data presented in Figure 3 was collected in February 2017.

Code to assemble final dataset used for analyses

The SAS and Python code used to assemble the final dataset is available on the Journal of Accounting Research website. The file “README Documentation.docx” explains all of the code files and how they are used.

Identifiers

The Compustat GVKEY and DATADATE identifiers for the firm-quarter observations used in the paper can be found in the “Identifiers.txt” file on the Journal of Accounting Research website.

Maintenance

Kurt H. Gee will maintain copies of the code and data until at least June 2024.