The Reluctant Analyst

Dan Bernhardt
University of Illinois
University of Warwick

Chi Wan
University of Massachusetts Boston

Zhijie Xiao
Boston College

Compliance with Data Policy for the Journal of Accounting Research
March 2016

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

   Chi Wan prepared the data and conducted the estimation and related 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.*

   Our sample of analyst recommendations is from the Institutional Brokers Estimate System (I/B/E/S) U.S. Detail file. Each analyst is identified by I/B/E/S with a unique numerical code (analyst masked code). We use this numerical identifier to match an analyst's stock recommendations to his earnings forecasts in the I/B/E/S Detail History file. If an analyst issues multiple recommendations for a firm within a calendar month, we only use the last recommendation.

   Stock return and trading volume related data are collected from CRSP. Firms' accounting and balance-sheet information is extracted quarterly from Compustat. These data are then merged with analyst stock recommendations using firm CUSIP codes. All data are download from the stated sources in 2012.

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).*

Not applicable.

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.*

   Following most studies in the literature, we reverse the I/B/E/S recommendation coding so that more favorable recommendations correspond to larger numbers (i.e., 1=Strong Sell, 2=Sell, 3=Hold, 4=Buy and 5=Strong Buy).
   We exclude recommendations issued by unidentified/anonymous analysts. We exclude analyst-firm pairs with fewer than 20 recommendations (including filled-in reiterations) over the entire sample period. If an analyst issues multiple recommendations for a firm within a calendar month, we only use the last recommendation.
   We further control for industry fixed effects captured by one-digit SIC codes. Appendix B details the sources and definitions of these variables.

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. 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.*

   The accompanying Stata codes take in the I/B/E/S detail file, fill in recommendations by assuming a rating is outstanding for 12 months, and identify analysts who are on the three- or five-tire rating systems. Analyst ratings are then merged with various controls variables (see appendix B for variable definitions), and then the dependent variable (stock recommendations) and all controls, including year and one-digit SIC industry dummies, are fed into the Bayesian estimation. Appendix A provides more details on model estimation.

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.*

   Data and programs will be maintained by at least one author for 6 years.