Personal Financial Distress, Limited Attention, and Sell-Side Analysts

Compliance with Data Policy for the Journal of Accounting Research

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

Hadiye Aslan handled the data and conducted the analysis.

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 is able to vouch for the stated source of the raw data.*

I hand-collected the names of analysts' and their city of employment from Nelson's Directory of Investment Research, Bloomberg, Factiva, and LinkedIn, LexisNexis Public Records Database as of January 2007. I manually collected the links to analysts' LinkedIn profiles using their full names and city of employment. Then a Perl program is used to parse analysts' LinkedIn profiles, and several analyst attributes, including work history, educational background, job tenure at their current place of employment, the cohort year, and the place of first employment are retrieved from those profiles. My main analysis is based on the analysts' home addresses. Analysts' deed records and home addresses are obtained from CoreLogic, which provides mortgage transactions and property-level information from U.S. deed records as originally recorded at local registries of deeds. For each deed transaction, the dataset provides, among other things, the full names of the seller(s) and buyer(s), the mortgage amount, the mortgage term, the transaction price, the address, and the transaction date, allowing me to create the complete ownership history of each house. Firm characteristics are from Compustat and CRSP. Neighborhood characteristics are extracted from 2000 Decennial Census. Institutional holdings are from Thomson Reuters 13F ownership database. Time series of the three Fama-French factors and the momentum factor are obtained from Kenneth French's data library. Zip code-level home values are from CoreLogic. I vouch for the stated sources of the raw 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 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 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 of instructions and instruments can be provided in the online appendix.*

For further details see the data section of the paper and the online appendix.

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 canconsider an exemption from this requirement. Whenever feasible, authors should also provide the identifiers (e.g., CIK, CUSIP) for their final sample.*

Full names and background information are hand-collected from several biographical sources and then merged with those in deed records. There are no identifiers. SAS was used to perform all statistical analysis, including double clustering of standard errors. The code can be seen in file 'code.pdf'.

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.*

I will maintain all data and programs for at least six years.