Facial Structure and Achievement Drive: Evidence from Financial Analysts

Xianjie He
Shanghai University of Finance and Economics

Huifang Yin
Shanghai University of Finance and Economics

Yachang Zeng
Nanyang Technological University

Huai Zhang
Nanyang Technological University

Hailong Zhao
Southwestern University of Finance and Economics

Compliance with Data Policy for the *Journal of Accounting Research*

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

Hailong Zhao, Huifang Yin and Xianjie He handled the data collection and conducted the analyses for the enclosed manuscript titled "Facial Structure and Achievement Drive: Evidence from Financial Analysts".

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.*

1) We provided detailed sample selection procedure in Table 1 and data source(s) of each variable in the Appendix (and in Online Appendix 15 for the newly introduced variables in the online appendix). The initial analyst-level sample and the analysts' license ID photos were downloaded in July 2015 from the website of Securities Association of China (SAC). We then draw data from Chinese Brokerage and Analyst System (CBAS) in December 2015 to obtain these analysts' annual earnings forecasts between 2004 and 2014 to form the initial sample of analyst forecasts at the analyst-firm-year level.[1] We manually link analysts' *fWHR* with their IDs in CBAS by analysts' names and their work experiences.

2) Our main independent variable is $D(fWHR>median)_i$, an indicator variable equal to one if analyst *i*'s facial width-to-height ratio (*fWHR*) is above the median in the analyst-level sample, and zero otherwise. *fWHR* is measured from the analysts' license ID photos collected from the SAC website, using the *ImageJ* software provided by the National Institutes of Health (Rasband 2015). Three Research Assistants (RAs), Jiahui Fan, Jingyi Wu and Xue Lai, are involved in measuring *fWHR* from the photos. We confirm that the RAs are not aware of the hypotheses at the time of data collection.

3) The analysts' annual earnings forecasts and the actual earnings that are used to construct the scaled forecast accuracy variable (*Accuracy*), the corporate site visit data (*Sitevisit*), and the data of analysts' star status (*Star*) were downloaded in December 2015 from CBAS.[2] The stock price data for the construction of *Accuracy* are downloaded in December 2015 from CSMAR. The stock

[1] CBAS is an academic database on Chinese brokerage firm and financial analyst research. CBAS contains data on analysts' forecasts and recommendations. It also provides data on analyst and brokerage house characteristics, which are gathered from the brokerage firms' annual reports posted in the SAC website. The database is available on the platform of CNRDS (Chinese Research Data Services).

[2] CBAS also provides (1) the listed firms' actual earnings reported in firms' annual reports and (2) analysts' corporate site visit records disclosed by the firms listed on China's Shenzhen Stock Exchange.

recommendation profitability variable (*BHAR[0,2]* and *LagBHAR[0,2]*) is calculated using the stock recommendation data from CBAS and the stock return data from CSMAR, which are downloaded in December 2015. To construct the variable of investment-banking generating ability (*Investment_Banking*), we identify the initial public offerings or the seasoned equity offerings of the firms underwritten by the investment banks that employ the analysts (in our initial sample downloaded from the SAC website and CBAS) in the same year. The data of the initial public offerings and the seasoned equity offerings are downloaded from the CSMAR database in December 2015. Other dependent variables including *Bold*, *Consistency*, *LFR*, *Optimism*, *Overconfidence*, and *Walkdown* are measured using the earnings forecast data downloaded from CBAS in December 2015.

4) The control variables for the analyst-level and forecast-level characteristics, including *Dayselap*, *Horizon*, *Frequency*, *Brokersize*, *Experience*, *Companies*, *Industries*, *Brokern*, *Coverage*, *BrokerAssets*, *NumAnalyst*, *Bias*, *and First-mover* are calculated using the earnings forecast data downloaded from CBAS in December 2015. *AnalystAge*, *Postgraduate*, and *Title* are measured using the analyst demographic information and work experience downloaded from CBAS in December 2015. *Strongbuy* and *Recomms* are constructed using stock recommendation downloaded from CBAS in December 2015.

5) We construct *Distance* by obtaining the firm location data and the brokerage firm location data, respectively, from CSMAR in December 2015 and from WIND and iFIND in July 2016. Data of the *Drating* variable are collected from the website of Shenzhen Stock Exchange in July 2016. The other firm-level and year-level control variables, including *Tangidum*, *HHIdum*, *SOE*, *MV*, *ROA*, *FirmAge*, *BM*, *Leverage*, *CF_Volatility*, *R&D*, *MKTSentiment*, *Instown*, *ROA_indadj*, *ROE*, *ROE_indadj*, *BHAR*, *BHAR_mktadj*, *Sales Growth*, *Earnings Growth*, *RetVolatility*, *Cash*, and *Accrual* are calculated using data from CSMAR downloaded in December 2015.

6) Hailong Zhao, Huifang Yin and Xianjie He 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.*

   Section 3 of the manuscript has described the steps to collect and process the data in details.

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.*

   We use SAS to convert the raw data and perform all analyses. The SAS code file titled "HYZZZ JAR_2018" uses the input datasets and yields the final sample. The Excel file titled "HYZZZ_2018 firm identifier" includes the firm IDs of final sample analyzed in this study. The Excel file titled "analyst_fwhr" presents the names and fWHR of the 1,193 male Chinese analysts examined in the manuscript.

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.*

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

Reference:

Rasband, W. 2015. ImageJ. *U. S. National Institutes of Health, Bethesda, Maryland, USA*: //imagej.nih.gov/ij/.