README : Create data

Non-GAAP Reporting and Investment

Charles G. McClure   Anastasia A. Zakolyukina*

20 August 2023

This note describes the code shared in `mz_share/create_data` that creates the final estimation data from raw data. All code is written in `R`.

Proprietary data:

- Equilar provides executive compensation data
- Audit Analytics provides non-GAAP data

Publicly available data:

- Kurt H. Gee's webpage at https://sites.google.com/view/kurthgee/data. Last accessed in October of 2018

Create data

Given the raw data from various sources described in the paper, the shared code should be executed in the following sequence:

non-GAAP data

`create_bcgw.R` imports bcgw data from Bentley et al. JAR (2018).

`create_detail.R` imports quarterly non-GAAP updates data from Audit Analytics.

Preliminary fundamentals data

`create_cpi.R` creates factors for inflation adjustment

`create_prelim.R` creates earnings data "as first reported" from Compustat Preliminary History - Quarterly, US

*https://www.charlesgmcclure.com/, https://www.anastasiazakolyukina.xyz/

Equilar compensation data

`create_executives.R` creates executives data with equity ownership and Edmans et al. myopia measure

Analysts data

`create_ibes_mb.R` creates IBES meet-or-beat data

Linking data

`create_detail_link.R` depends on `{detail}`. Creates cik-gvkey links for AuditAnalytics data

Value data

`create_value.R` depends on `{bcgw}`. Creates value a la' Bradshaw and Sloan (2002)

Intangible assets

`create_intang.R` depends on `{cpi}`. Creates quarterly intangible investment and assets

Non-GAAP Reg G data

`create_detail_set.R` depends on `{detail, detail_link}`. Creates data for non-GAAP Reg G stats by linking AuditAnalytics data to Compustat

Firm fundamentals

`create_firms.R` depends on `{intang, prelim, bcgw, value, cpi}`. Creates data with firms' fundamentals

Merged data

`create_est_merged.R` depends on `{firms, ibes_mb, executives}`. Creates merged estimation data

Estimation data

`create_set.R` creates data for estimation `est_set` from `est_merged`

`create_aa_set.R` creates data for estimation for AuditAnalytics sample from `est_merged`

CIKs for firms in the final estimation sample are in `cik.csv`.

Compute moments

`get_moments.Rmd` computes moments and weight matrices.

`get_pshuffle.Rmd` computes re-shuffling probability for cash flows.

The file `est_set_tar.csv` is pseudo-data for the compute-moments code. This pseudo-data is a random draw of 100 firms from `est_set` sample. To de-identify the firms, `gvkey` and `company_id` were replaced with artificial identifiers. In addition, we dropped the variables not used in `.Rmd` code.