

Earnings Management and Investor Protection:

Accounting Reading Group - Assignment III

Melisa Mazaeva

Aug 6, 2024

Abstract

This project uses the TRR 266 Template for Reproducible Empirical Accounting Research (TREAT) to provide an infrastructure for open science-oriented empirical projects. Leveraging external Worldscope data sets on financial data, the repository showcases a reproducible workflow that integrates Python scripts for data analysis. The project's output demonstrates a comprehensive application of skills to replicate and extend the findings from the seminal paper by Leuz, Nanda, and Wysocki (2003), particularly in providing descriptive statistics for the four individual earnings management measures as well as the aggregate earnings management score across various countries. In doing so, it documents and discusses the research design choices made and the variations between the original and reproduced results. This code base, adapted from TREAT, should give you an overview on how the template is supposed to be used for my specific project and how to structure a reproducible empirical project.

1 Research Design Choices and Assumptions

The aim of Assignment III is to replicate a specific empirical table (Table 2 Panel A) from the seminal paper by Leuz, Nanda, and Wysocki (2003). This table involves calculating the EM measures for firms across various countries over a defined period and examining the relationship between these measures and investor protection. The replication process includes data loading, preparation, cleaning, and normalization, followed by the application of statistical methods to compute and interpret financial metrics. For Assignment III, I pulled data from the Worldscope database through WRDS and used the Python programming language to carry out the empirical analysis. Visual Studio Code was used as the Integrated Development Environment (IDE) for writing, debugging, and optimizing the Python code.

The replication is based on data pulled from the Worldscope database, specifically from the `wrds_ws_company` and `wrds_ws_funda` tables, which were merged for the analysis. The first table provides company profile information, including items such as ISIN, Worldscope Identifiers, company name, and the country where the company is domiciled. The latter table contains Fundamentals Annual data at the company-year level, including items such as total assets, net income, and other relevant financial variables.

Following Leuz, Nanda, and Wysocki (2003), I focus the analysis on companies across various countries, ensuring that the data accurately reflects the fiscal years 1990 to 1999 as specified in the original study. The replication aims to mirror the research design as closely as possible with the available data.

In addition, I impose the following assumptions to ensure clarity and consistency where the paper does not provide explicit guidance:

1.1 check assumptions in Fikir and mine

1. The original paper references the November 2000 version of the Worldscope Database. However, the data used for this analysis represents the latest available version, updated in July 2024, with quarterly frequency updates (Wharton Research Data Services 2024a). Due to potential adjustments and updates made to the database since November 2000, there may be differences between the databases that could affect the results. For example, companies may

restate financials after the original reporting period, so that these restatements are reflected in the later database version rather than the historical one. Moreover, the data vendor Refinitiv regularly updates its databases to correct errors and add new information, which may be included in the later data but not in the November 2000 snapshot.

2. While pulling the data for analysis, I encountered negative values for some key financial metrics such as operating income (`item1250`), net income before extraordinary items/preferred dividends (`item1551`), and net income before preferred dividends (`item1651`). The paper by Leuz, Nanda, and Wysocki (2003) does not explicitly specify how to handle negative values in key financial metrics. For the purpose of this replication study, I will include negative values in the analysis. Including these values ensures that the analysis captures the full spectrum of earnings management activities across different countries.
3. The EM measures are based on scaled variables (e.g., operating cash flow scaled by lagged total assets). As such, the currency of the relevant data items should not affect the results as long as the same currency is used consistently for both the numerator and the denominator. This approach ensures comparability across different countries, regardless of their local currencies. Additionally, according to a document by (Thomson Financial 2007, 20), all Worldscope data is consistently reported in the local currency of each firm’s country of domicile, eliminating the need for currency conversions in this project.
- 4.

1.2 delete. mine

5. It is assumed that Penman (2013) restricted the possible P/B values to a range of 0 to 7 to exclude outliers. Hence, extreme P/B values that are negative and very high (greater than 7) are excluded to focus the analysis on firms with more stable and reasonable valuations, reducing the impact of outliers.

By following the steps provided in Section 4 and adhering to the assumptions made, I successfully replicated the analysis and produced the required table. A thorough step-by-step approach, with each step clearly documented, helped to understand and verify the outputs.

2 Replication Steps

2.1 Step 1: Pulling the Data and Managing the Databases

In contrast to Assignment I, where the data was provided externally, Assignment III involves additionally pulling data directly from the Worldscope database, merging relevant tables, and preparing the data for further analysis from raw data to final output.

To compile the dataset, the dynamic and static datasets were merged on the `item6105` identifier, representing the unique Worldscope Permanent ID. WRDS advises using this identifier consistently within Worldscope data since it remains stable over time (Wharton Research Data Services 2024b).

2.2 Step 2: Data Preparation

To prepare the data according to the methodology outlined in the paper, it is essential to follow specific preprocessing steps. Firstly, countries with sufficient firm-year observations must be filtered. Each country should have at least 300 firm-year observations for several key accounting variables, including total assets, sales, net income, and operating income. Secondly, firms with adequate consecutive data must be identified. Each firm must have income statement and balance sheet information for at least three consecutive years, where I focus on the same key variables.

2.3 Step 3: Analysis Implementation and Table Reproduction

In this replication step, implement the calculations for the EM measures as described in the methodology of the paper, ensuring that all statistical methods and groupings are accurately replicated.

References

- Leuz, Christian, Dhananjay Nanda, and Peter D. Wysocki. 2003. “Earnings Management and Investor Protection: An International Comparison.” *Journal of Financial Economics* 69 (3): 505–27. [https://doi.org/10.1016/S0304-405X\(03\)00121-1](https://doi.org/10.1016/S0304-405X(03)00121-1).
- Thomson Financial. 2007. “Worldscope Database: Datatype Definitions Guide.” Issue 6. https://www.tilburguniversity.edu/sites/default/files/download/WorldScopeDatatypeDefinitionsGuide_2.pdf.
- Wharton Research Data Services. 2024a. “Thomson Reuters - Worldscope.” https://wrds-www.wharton.upenn.edu/data-dictionary/tr_worldscope/.
- . 2024b. “WRDS Overview of Worldscope Data.” <https://wrds-www.wharton.upenn.edu/pages/support/manuals-and-overviews/thomson-reuters/worldscope/wrds-overview-worldscope-data/>.