Earnings Management and Investor Protection:

Accounting Reading Group - Assignment III

Melisa Mazaeva

Aug 11, 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 Worldscope data on financial information, this 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 four individual earnings management measures, as well as an 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, provides an overview of how the template can be used for this specific project and how to structure a reproducible empirical research project.

Table of contents

1 List of Abbreviations

CFO: Operating Cash Flow

EM: Earnings Management

IDE: Integrated Development Environment

UK: United Kingdom

US: United States

WRDS: Wharton Research Data Services

2   Introduction

The aim of this paper is to illustrate the use of open science tools in empirical accounting research.

The original study by Leuz, Nanda, and Wysocki (2003) examines systematic differences in earnings management (EM) across 31 countries and proposes that these differences are influenced by the level of investor protection. Specifically, the paper argues that stronger investor protection reduces earnings management by limiting insiders' ability to acquire private control benefits.

This paper presents the replication of Table 2 Panel A from the seminal study by Leuz, Nanda, and Wysocki (2003). The original table provides descriptive statistics for four individual earnings management measures and the aggregate earnings management score across various countries, highlighting significant differences in EM practices between Continental Europe, Asia, and Anglo-American countries.

The project uses data from the Worldscope Database via Wharton Research Data Services (WRDS), focusing on fiscal years 1990 to 1999 and covering multiple countries. The replication aims to follow the research design by Leuz, Nanda, and Wysocki (2003) as closely as possible with the available data. It employs four individual earnings management measures: earnings smoothing using accruals (EM1), the correlation between changes in accruals and cash flows (EM2), the magnitude of accruals (EM3), and the ratio of small profits to small losses (EM4). These measures are combined to create an aggregate earnings management score, providing a comprehensive overview of EM practices.

The analysis includes pulling and filtering the relevant financial data, identifying relevant companies and countries based on the authors' requirements, calculating the EM measures for each country, and examining the relationship between each country's aggregate earnings management score and its ranking. Explicit assumptions were made whenever the paper by Leuz, Nanda, and Wysocki (2003) was unclear on how to proceed. The complete computation code for EM measures is available in `code/python/do_analysis.py` for detailed review if necessary. Code snippets will be presented only in exceptional cases to highlight specific points.

The replicated table is shown below, followed by the research design assumptions in Section 3, documentation of the replication steps in Section 4, a detailed comparison in Section 5, and

concluding remarks in Section 6.

Table 1 provides descriptive statistics for the four individual earnings management measures as well as the aggregate earnings management score, where countries are sorted in descending order based on their aggregate score. The discussion on replication findings is presented in Section 5.

[Table 1 about here.]

3 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 (Wharton Research Data Services 2024a). 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 (Wharton Research Data Services 2024b).

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 by Leuz, Nanda, and Wysocki (2003) does not provide explicit guidance:

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 2024c). 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 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. The original paper outlines key terms that will be used in this project to ensure consistency and accuracy in the replication. Leuz, Nanda, and Wysocki (2003) define earnings management as the manipulation of a firm's reported economic performance by insiders to deceive certain stakeholders or to affect contractual outcomes. The authors describe investor protection as a key institutional factor that limits insiders' acquisition of private control benefits, thereby reducing their incentives to manage accounting earnings by ensuring strong and well-enforced rights for outside investors. Finally, private control benefits are the advantages that insiders can gain from controlling a firm, which can include financial gains or other benefits not shared with other stakeholders (Leuz, Nanda, and Wysocki 2003).

3. While pulling the data for analysis, I encountered negative values for some key financial metrics such as operating income (`item1250`) 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 included negative values in the analysis. Including these values ensures that the analysis captures the full spectrum of earnings management activities across different countries.

4. Another potential source of discrepancies between the original and replicated tables may be the choice of variables pulled from Worldscope. For example, in the `wrds_ws_funda table`, both `item1151` and `item4051` are labeled as "DEPRECIATION, DEPLETION AND AMORTIZATION" (Wharton Research Data Services 2024b). I chose `item1151` from the Income Statement rather than `item4051` from the Cash Flow Statement, based on the *Worldscope Balancing Model - Industrials* by Wharton Research Data Services (2024d). Since the authors

do not specify the choice of variables used from the database, this could cause differences in the results.

5. 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.

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.

4   Replication Steps

4.1   Step 1: Pulling the Data and Managing the Databases

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

To ensure data integrity, I filtered out rows with empty `item6105` (Worldscope Permanent ID) values, as this identifier is crucial for firm/year-level filtering in the data preparation step. A total of 125 observations from the dynamic data and 10,306 observations from the static data were removed. The dynamic and static datasets were then merged on the `item6105` identifier, representing the unique Worldscope Permanent ID. According to Wharton Research Data Services (2024e), this identifier should be used consistently within Worldscope data, as it remains stable over time. An inner join was used for this merge to ensure that only complete and consistent data from both tables were retained.

Additionally, a filter was applied to retain only company rows. This was achieved by selecting rows where the `item6100` field equals 'C', indicating that the Worldscope Identifier represents a company. This step ensures that the analysis includes only company data, excluding averages,

exchange rates, securities, or stock indices, as advised by Wharton Research Data Services (2024e).

In line with the methodology used by Leuz, Nanda, and Wysocki (2003), financial institutions were removed from the analysis based on their SIC codes. This was done by filtering out rows where the `item7021` identifier, representing the SIC code, falls within the range of 6000 to 6999. This ensures that the dataset focuses solely on non-financial companies.

Further filtering ensured that only data from the 31 countries specified in the paper were included. These steps were applied to reduce the dataset size and improve the workflow.

Moreover, the configuration file utilized additional refinements as proposed by Wharton Research Data Services (2024e). Specifically, using "A" (Annual) in the `freq` variable ensured that the data represents financial information reported on an annual basis, consistent with the methodology outlined in the paper by Leuz, Nanda, and Wysocki (2003). This excludes data reported on a current, daily, or quarterly basis. After retaining only relevant company data, filtering out financial institutions, and focusing on the specified countries, the processed data was saved to a CSV file in the `data/pulled` directory, as specified in configuration file `config/pull_data_cfg`.

4.2 Step 2: Data Preparation

To verify the pulled data, I checked the dataset for duplicate firm-year observations based on the combination of the Worldscope Permanent ID (`item6105`) and the fiscal year (`year_`). I confirmed that no duplicates were present, ensuring the accuracy of the data for further analysis. As the original study does not specify the net income measure, I used `item1651` (Net Income before Preferred Dividends) as the variable for net income. This choice aligns with the final net income figure reported in the income statement, based on *Worldscope Balancing Model - Industrials* by Wharton Research Data Services (2024d).

To prepare the data sample in line with the methodology outlined in the paper, it was essential to follow the requirements set by Leuz, Nanda, and Wysocki (2003). First, countries with sufficient firm-year observations were filtered. Each country needed at least 300 firm-year observations for key accounting variables, including total assets, sales, net income, and operating income. In this step, no countries were eliminated, as all countries met the requirement, which aligns with the paper's overview of countries (31 countries in total).

In the second filtration step, firms with adequate consecutive data were identified. Each

firm had to have income statement and balance sheet information for at least three consecutive years, with all key accounting variables mentioned above present. If a firm had at least three consecutive years of complete data at any point, all its data entries were retained in the final dataset. Therefore, only those countries and firms that met these criteria for all specified variables were retained in the dataset.

During the preparation step, 8,265 firms and 20,521 firm-year observations (all due to the second filtration step) were dropped, resulting in a final dataset with 18,040 non-financial firms and 123,469 firm-year observations. The differences in the numbers between the prepared dataset and the figures mentioned in the paper (70,955 firm-year observations and 8,616 non-financial firms) could be due to assumptions outlined in Section 3, such as variations in initial datasets, data updates, and filtering criteria. However, the original paper might have included additional data cleaning steps not explicitly mentioned, such as handling outliers, specific industry exclusions, or other criteria, which could affect the final counts. ## table 2 correctly referenced? Given that the number of observations in this project is significantly higher than in the original study, Table 1 from Leuz, Nanda, and Wysocki (2003) was partially replicated (focusing only on columns related to countries and firm-year observations) to highlight specific discrepancies that could arise for certain countries. The results in the Table 2 below show a higher number of firm-year observations for almost all countries compared to the paper by Leuz, Nanda, and Wysocki (2003), likely due to database updates leading to a more comprehensive dataset. The trend in the number of firm-year observations across countries is consistent with the original study, with larger capital markets like Japan, the United Kingdom, and the United States having the highest numbers. The distribution of observations in my dataset shows higher mean and median values compared to the paper by Leuz, Nanda, and Wysocki (2003), indicating a greater overall number of firm-year observations, while the minimum values are relatively close, suggesting that the overall distribution of observations across countries remains broadly similar.

However, there are significant discrepancies, particularly for the United States, which has 41,585 observations compared to 3,792 in the original study, possibly due to more comprehensive data or changes in dataset scope. Leuz, Nanda, and Wysocki (2003) focused only on S&P 500 companies for the U.S., while this analysis uses a more comprehensive dataset from the Worldscope database, which includes detailed financial statement data and profile data on public companies

globally. This broader scope results in a higher number of firm-year observations for the United States. Despite the differences in absolute numbers, the relative trends across countries align with the paper.

[Table 2 about here.]

4.3 Step 3: Analysis Implementation and Table Reproduction

In this replication step, the calculation of the EM measures is implemented according to the methodology described by Leuz, Nanda, and Wysocki (2003). The study employs four distinct proxies that capture different dimensions of earnings management, such as earnings smoothing and accrual manipulations. These measures are used to assess the pervasiveness of earnings management across countries. This step ensures that all statistical methods and groupings are accurately replicated, following the paper's approach to capture the extent to which corporate insiders use accounting discretion to obscure their firm's true economic performance.

4.3.1 Smoothing Reported Operating Earnings Using Accruals (EM1)

EM1 is a measure that assesses how much insiders (e.g., company executives) smooth or stabilize reported earnings by manipulating accruals. The concept is that firms may attempt to reduce fluctuations in their reported operating income to present a more stable financial performance to investors. This measure reflects the degree to which insiders use accounting discretion to smooth reported earnings, with lower values of EM1 suggesting more extensive earnings smoothing.

To calculate this measure:

1. Compute Accruals by using changes in various components of a firm's balance sheet and income statement, such as total current assets, cash, current liabilities, short-term debt, taxes payable, and depreciation:

10

$$\text{Accruals}_{it} = (\Delta \text{CA}_{it} - \Delta \text{Cash}_{it}) - (\Delta \text{CL}_{it} - \Delta \text{STD}_{it} - \Delta \text{TP}_{it}) - \text{Dep}_{it}$$

where:

$\Delta \text{CA}_{it} = $ Change in total current assets

$\Delta \text{Cash}_{it} = $ Change in cash and cash equivalents

$\Delta \text{CL}_{it} = $ Change in total current liabilities

$\Delta \text{STD}_{it} = $ Change in short-term debt included in current liabilities

$\Delta \text{TP}_{it} = $ Change in income taxes payable

$\text{Dep}_{it} = $ Depreciation and amortization expense for firm i in year t

2. Derive Operating Cash Flow (CFO) that is indirectly derived from operating income minus accruals:

$$\text{CFO}_{it} = \text{Operating Income}_{it} - \text{Accruals}_{it}$$

3. Calculate EM1, which is determined as the country's median ratio of the firm-level standard deviations of operating income and operating cash flow. Both standard deviations are scaled by lagged total assets to control for differences in firm size:

$$\text{EM1} = \text{Median} \left(\frac{\frac{\sigma(\text{Operating Income}_{it})}{\text{Total Assets}_{i(t-1)}}}{\frac{\sigma(\text{Operating Cash Flow}_{it})}{\text{Total Assets}_{i(t-1)}}} \right)$$

4.3.2 Smoothing and the Correlation Between Changes in Accounting Accruals and Operating Cash Flows (EM2)

EM2 is designed to capture the extent to which insiders use accounting discretion to smooth reported earnings by examining the correlation between changes in accruals and changes in operating cash flows. A stronger negative correlation between these variables indicates that accruals are being used to buffer the impact of cash flow fluctuations, potentially masking the true economic performance of the firm.

To calculate this measure:

1. Compute Accruals and CFO using same Formulas used in Section 4.3.1.

2. Calculate the change in Accruals and CFO for each firm-year:

$$\Delta\text{Accruals}_{it} = \text{Accruals}_{it} - \text{Accruals}_{i(t-1)}$$

$$\Delta\text{CFO}_{it} = \text{CFO}_{it} - \text{CFO}_{i(t-1)}$$

3. Scale both changes by lagged total assets:

$$\frac{\Delta\text{Accruals}_{it}}{\text{Total Assets}_{i(t-1)}}, \quad \frac{\Delta\text{CFO}_{it}}{\text{Total Assets}_{i(t-1)}}$$

4. Calculate EM2, which is computed as the Spearman correlation between these scaled changes, aggregated at the country level over the pooled set of firms within each country, meaning that the correlation reflects the aggregated relationship between changes in accruals and changes in CFO across all firms in that country:

$$\text{EM2} = \rho\left(\frac{\Delta\text{Accruals}_{it}}{\text{Total Assets}_{i(t-1)}}, \frac{\Delta\text{CFO}_{it}}{\text{Total Assets}_{i(t-1)}}\right)$$

4.3.3 Discretion in Reported Earnings: The Magnitude of Accruals (EM3)

EM3 is a measure that obtains the degree of discretion that insiders exercise when reporting earnings, particularly when it comes to misstating a firm's economic performance. It focuses on the magnitude of accruals as a proxy for earnings management. Insiders can use accruals to overstate or understate earnings, thereby achieving certain financial targets or presenting a skewed view of the firm's performance. Higher scores of EM3 indicate more earnings management, as larger magnitudes of accruals relative to cash flow from operations suggest greater discretion in financial reporting.

To calculate this measure:

1. Compute Accruals and CFO using same Formulas used in Section 4.3.1.
2. Calculate the absolute values of accruals and CFO for each firm-year.
3. EM3 is calculated as the country-level median ratio of the absolute value of firms' accruals to the absolute value of their cash flow from operations (CFO). This ratio is used as it scales the magnitude of accruals to a measure of the firm's size and performance, allowing for comparison

across firms of different sizes and performance levels:

$$\text{EM3} = \text{Median}\left(\frac{|\text{Accruals}_{it}|}{|\text{CFO}_{it}|}\right)$$

4.3.4 Discretion in Reported Earnings: Small Loss Avoidance (EM4)

EM4 captures the extent to which firms manage their reported earnings to avoid small losses. This is an important measure that quantifies the propensity of firms to avoid small losses through earnings management. By calculating the ratio of small profits to small losses, it provides a clear indication of how widespread this practice is within a particular country. A higher EM4 value suggests that firms in a given country are more likely to use accounting discretion to convert small losses into small profits, reflecting a tendency toward earnings management. Note that I used `item1651` (Net Income before Preferred Dividends) as the variable for net earnings, as mentioned in Section 4.2.

EM4 is calculated as the ratio of the number of small profits to the number of small losses for each country, where small profits and small losses are determined by scaling net earnings by lagged total assets:

$$\text{EM4} = \frac{\text{Small Profits}}{\text{Small Losses}}$$

where:

$$\text{Small Profits} = \frac{\text{Net Earnings}_{it}}{\text{Total Assets}_{it-1}} \in [0, 0.01]$$

$$\text{Small Losses} = \frac{\text{Net Earnings}_{it}}{\text{Total Assets}_{it-1}} \in [-0.01, 0)$$

4.3.5 Aggregate Measure of Earnings Management

The aggregate earnings management score is a summary measure that captures the overall level of earnings management in each country. This measure is calculated by ranking countries on each of the four individual earnings management measures (EM1-EM4), where a higher rank indicates a greater degree of earnings management. The final aggregate score is the average of these ranks, providing a comprehensive view of a country's earnings management practices. The countries are sorted in descending order based on their aggregate score.

To calculate this measure:

$$\text{Aggregate Earnings Management Score}_{\text{country}} = \frac{\text{Rank}_{\text{EM1}} + \text{Rank}_{\text{EM2}} + \text{Rank}_{\text{EM3}} + \text{Rank}_{\text{EM4}}}{4}$$

5 Results

The comparison between the calculated EM1 values and those reported in the original paper reveals a strong alignment, with only minor differences observed across most countries. The overall pattern of EM1 values is consistent, effectively capturing the underlying trends in earnings management across different countries. Any differences, such as those for the Philippines, Taiwan, and Norway, are relatively small and likely due to variations in data sources, updates, or slight methodological differences, as outlined in Section 3. The summary statistics (mean, median, min, max, and standard deviation) also closely match those in the original study, further supporting the robustness of the replication process. The results indicate that Anglo-American countries tend to have higher EM1 values, suggesting less earnings smoothing compared to Continental Europe and Asia, consistent with the conclusion of Leuz, Nanda, and Wysocki (2003) that earnings are smoother in Continental Europe and Asia than in Anglo-American countries after controlling for the volatility of cash flows.

The replication findings regarding EM2 indicate that earnings smoothing is more pronounced in countries like Greece and Japan, while countries like Canada and the United States exhibit less pervasive smoothing behavior, in line with the findings of Leuz, Nanda, and Wysocki (2003). Despite some differences in magnitude, the overall trends are consistent, reinforcing the idea that earnings smoothing varies significantly across different countries. In fact, these small variations do not significantly alter the overall conclusions. Specifically, the minimum and maximum EM2 values in the replication closely match the original findings, with Greece showing the most earnings smoothing (min: -0.942 vs. -0.928) and Norway the least (max: -0.786 vs. -0.722). This consistency reinforces the accuracy of the replication.

The replicated EM3 results are very similar to those of the original study, with close alignment in summary statistics and consistent identification of South Africa and Germany as the countries with the minimum and maximum EM3 values. Additionally, the replication confirms that

14

the magnitude of firms' accruals relative to their operating cash flows is smaller in the U.K. and the U.S. compared to Austria, Germany, and South Korea, consistent with the original findings.

The replicated results show significant regional differences in EM4 values, reflecting varying degrees of loss avoidance. Spain has the highest EM4 value, highlighting strong loss avoidance among European firms. Norway, with the lowest EM4 value, corresponds to the minimum value in the original paper, confirming less aggressive loss avoidance. High EM4 values in Asian countries like Indonesia (5.667) and India (6.000) further support the trend that European and Asian firms exhibit greater loss avoidance than their Anglo-American counterparts, such as the United States (1.912) and Canada (2.402), which have lower EM4 values. These results align with the original study's conclusions on regional loss avoidance behavior.

The comparison between the original and replicated aggregate earnings management scores reveals a strong overall alignment with consistent key trends. In the original study, Austria and Greece are at the top of the rankings with aggregate scores of 28.3, indicating the highest levels of earnings management. In the replication, Greece and Austria remain at the top with slightly adjusted scores of 27.5 and 27.2, respectively. Taiwan and South Korea also rank high, mirroring their positions in the original study. While some differences in specific rankings and scores exist, such as the United States consistently having the lowest level of earnings management (2.0 in the original and 3.5 in the replication), these variations are minor and do not undermine the overall consistency.

In summary, the replication effectively captures the original study's key findings, confirming that European and Asian countries generally exhibit higher levels of earnings management than their Anglo-American counterparts. The strong similarity between the original and replicated results underscores the robustness and reliability of the findings.

6 Conclusion

The project effectively demonstrates the use of an open science and collaborative workflow for analyzing earnings management measures across various countries. By following a systematic approach and utilizing the TRR 266 Template for Reproducible Empirical Accounting Research, I successfully replicated an empirical table from Leuz, Nanda, and Wysocki (2003) and provided

insights into trends in earnings management practices, which, similar to the original study, vary significantly across countries.

Leuz, Nanda, and Wysocki (2003) document systematic differences in the level of earnings management across 31 countries, suggesting that insiders use earnings management (EM) to conceal firm performance and protect their private control benefits. The study further argues that stronger investor protection reduces earnings management by limiting insiders' ability to secure private benefits. By conducting a descriptive cluster analysis following the descriptive statistics replicated in this project, the paper identifies groupings of countries with similar institutional characteristics, demonstrating that earnings management varies systematically across these institutional clusters.

The replicated results reveal significant differences in earnings management practices across countries, with consistent patterns in the relative magnitudes of the individual measures. These differences are particularly notable in areas such as loss avoidance and earnings smoothing. In line with the original study, the replication suggests that firms in countries with more developed equity markets, dispersed ownership structures, strong investor rights, and robust legal enforcement tend to engage in less earnings management. This consistent finding across both the original and replicated studies underscores the critical role of the institutional environment in shaping corporate earnings management behaviors on an international scale.

The replication analysis also reveals that the individual earnings management measures are highly correlated, resulting in similar rankings across the four measures and the aggregate earnings management score. This consistency supports the original study's conclusion that a single underlying factor likely represents the four individual measures. As in the original results, the replicated country rankings based on the aggregate score show high ranks for countries like Greece, Austria, and South Korea, and lower ranks for countries such as Australia, the U.K., and the U.S. This reinforces the robustness of the aggregate measure in capturing the overall level of earnings management across different countries.

This final assignment required the comprehensive application of all the skills learned and feedback received from previous assignments, making it a more complex project workflow. In the future, this repository can be cloned or forked (if made public) to kickstart further projects on earnings management measures analysis. Thanks for reading!

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.

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. "Table: WRDS - Worldscope - Company Information (Wrds_ws_company)." https://wrds-www.wharton.upenn.edu/data-dictionary/tr_worldscope/wrds_ws_company/.

———. 2024b. "Table: WRDS - Worldscope - Fundamentals Annual (Wrds_ws_funda)." https://wrds-www.wharton.upenn.edu/data-dictionary/tr_worldscope/wrds_ws_funda/.

———. 2024c. "Thomson Reuters - Worldscope." https://wrds-www.wharton.upenn.edu/data-dictionary/tr_worldscope/.

———. 2024d. "Worldscope Balancing Model - Industrials." https://wrds-www.wharton.upenn.edu/documents/526/Worldscope_Balancing_Model_-_Industrials.xls.

———. 2024e. "WRDS Overview of Worldscope Data." https://wrds-www.wharton.upenn.edu/pages/support/manuals-and-overviews/thomson-reuters/worldscope/wrds-overview-worldscope-data/.

item6026	EM1	EM2	EM3	EM4	Aggregate_EM_Score
GREECE	0.365	-0.942	0.764	3.375	27.5
AUSTRIA	0.350	-0.933	0.833	3.263	27.2
TAIWAN	0.302	-0.937	0.838	2.593	26.2
KOREA (SOUTH)	0.363	-0.933	0.706	3.205	24.1
JAPAN	0.430	-0.930	0.602	3.992	23.1
BELGIUM	0.442	-0.910	0.709	3.308	22.5
PORTUGAL	0.367	-0.922	0.746	2.474	22.2
ITALY	0.461	-0.916	0.686	3.245	21.2
PHILIPPINES	0.430	-0.855	0.745	3.500	21.1
THAILAND	0.452	-0.906	0.738	3.043	20.8
SINGAPORE	0.350	-0.914	0.713	2.034	20.6
INDONESIA	0.502	-0.896	0.695	5.667	20.4
GERMANY	0.501	-0.882	0.855	2.825	18.8
SPAIN	0.476	-0.894	0.543	7.333	18.8
HONG KONG	0.466	-0.898	0.635	3.214	18.5
SWITZERLAND	0.483	-0.893	0.582	5.077	18.2
DENMARK	0.516	-0.926	0.557	3.227	18.0
MALAYSIA	0.449	-0.900	0.663	2.400	17.2
PAKISTAN	0.517	-0.911	0.552	2.643	15.2
FRANCE	0.507	-0.896	0.626	2.559	14.6
INDIA	0.573	-0.880	0.529	6.000	14.4
NETHERLANDS	0.554	-0.884	0.516	4.929	14.0
SWEDEN	0.602	-0.850	0.533	2.565	9.0
CANADA	0.620	-0.830	0.538	2.402	7.4
NORWAY	0.573	-0.786	0.627	1.314	7.1
FINLAND	0.631	-0.849	0.539	2.119	7.0
UNITED KINGDOM	0.624	-0.857	0.435	2.071	6.2
AUSTRALIA	0.617	-0.830	0.470	1.844	4.9
IRELAND	0.713	-0.853	0.402	1.500	3.8
UNITED STATES	0.732	-0.789	0.525	1.912	3.5
SOUTH AFRICA	0.669	-0.823	0.326	1.364	2.2
Mean	0.504	-0.885	0.620	3.129	
Median	0.501	-0.896	0.626	2.825	
Std	0.111	0.042	0.130	1.399	
Min	0.302	-0.942	0.326	1.314	
Max	0.732	-0.786	0.855	7.333	

Table 1: Replicated table - Country scores for earnings management measures (Sorted by aggregate earnings management)

Country	# Firm-years
AUSTRALIA	2099
AUSTRIA	686
BELGIUM	932
CANADA	4215
DENMARK	1326
FINLAND	1058
FRANCE	5515
GERMANY	5307
GREECE	1034
HONG KONG	1862
INDIA	2068
INDONESIA	978
IRELAND	625
ITALY	1591
JAPAN	20528
KOREA (SOUTH)	2038
MALAYSIA	2283
NETHERLANDS	1782
NORWAY	1210
PAKISTAN	598
PHILIPPINES	574
PORTUGAL	577
SINGAPORE	1291
SOUTH AFRICA	1334
SPAIN	1221
SWEDEN	1766
SWITZERLAND	1553
TAIWAN	1259
THAILAND	1664
UNITED KINGDOM	12910
UNITED STATES	41585
Mean	3983
Median	1553
Min	574
Max	41585

Table 2: Replicated table - Descriptive statistics: Number of firm-year observations per country