

# Measuring the Prevalence of Earnings Manipulations: A Novel Approach

## *Data Description Sheet*

Nicole Cade, Joshua Gunn, and Alex Vandenberg

August 26, 2024

This document provides information required to comply with the Data and Code Sharing Policy of the *Journal of Accounting Research*.

### **1. Data Handling and Analyses**

Joshua Gunn performed the primary analysis of executives' responses to our standard survey and to our list experiment. All authors have access to all response data.

### **2. Data Generation**

As part of our list experiment instrument, we required 20 pairs of innocuous statements (e.g., *I usually drive to work*) that we could reasonably expect to produce negatively correlated responses when presented to executives. We conducted a pilot study on August 10, 2022 to identify these pairs. In this pilot study, 60 participants recruited from Prolific.com viewed 70 innocuous statements (listed in Appendix B) within a Qualtrics instrument, responded True or False to each statement, and flagged any statement that was unclear or confusing. We required all pilot participants to meet the following criteria: (i) employment status is full-time, (ii) working hours are regular 9-5, (iii) decision-making responsibilities at work are accounting/finance matters, (iv) industry role is junior or middle management, (v) highest education level completed is an undergraduate degree, graduate degree, or doctorate degree, and (vi) approval rate on prior studies is 95% or higher. We downloaded the raw data from our Qualtrics pilot study instrument on August 22, 2022.

After conducting our pilot study and constructing our primary research instruments, we recruited top-level executives from firms listed in the Russell 3000 Index (reconstituted in summer 2022) to complete our primary study. We exclude from consideration any firms where a firm-specific event (e.g., merger, bankruptcy) made recruiting infeasible or where the executive team was substantially similar to that of another firm in the Index. We contacted at least one executive at each of the other 2,886 firms in the Index.

In February 2023, we contacted 200 executives (1 executive per firm) via USPS direct mail and 797 executives (approx. 4 executives per firm) via LinkedIn's paid direct messaging

service, InMail. After conducting this “soft launch,” we contacted three executives at each of the remaining 2,486 firms in our target sample via USPS direct mail. Specifically, we sent our first batch of mailers to primarily CAOs and Controllers in mid-May 2023, our second batch to primarily CFOs in mid-June 2023, and our third batch to primarily CEOs in mid-July 2023. Each mailer included both a paper instrument and a QR code that connects to the online version of the same condition-specific instrument when scanned. Approximately three weeks after each batch is sent, we sent each respective executive a 5 x 7 card. Each card included a brief note to remind executives of the mailer we previously sent and a QR code that connects to the online version of the relevant instrument when scanned.

In total, we collected response data from nine instruments on Qualtrics. Six instruments pertained to our soft launch recruitment period (three condition-specific instruments for each of the two recruitment methods) and three instruments pertained to our primary recruitment period (three condition-specific instruments). Executives contacted via LinkedIn completed the online version of their condition-specific instrument on Qualtrics. Executives contacted via USPS direct mail completed this online version or completed the paper version of their condition-specific instrument and returned it to us using the provided business reply envelope. In instances where an executive returned their instrument to us via mail, we manually submitted their responses in the corresponding condition-specific instrument on Qualtrics. To conduct our analyses, we downloaded the raw data from our Qualtrics instruments on November 19, 2023.

We take two steps to source the population descriptives reported in column 5 of Table 2. First, in February 2024, we used Compustat to identify industry (using SIC codes) and firms’ revenue reported in the most recent fiscal year. Second, in March 2024, we manually searched a variety of outlets (e.g., Yahoo! Finance, firm webpages, SEC filings) to collect the decade founded for a random sample of 500 firms in our target sample.

Sections OA.1-OA.3 and Figure OA.1 provide a detailed account of our data collection procedures. Appendix C separately displays the questions included in two of our three conditions. Each statement about earnings manipulation in Appendix C is presented as a true/false question to executives in our third condition. Appendix D displays the post-experiment questionnaire.

### **3. Proprietary Nature of Data**

The research team is legally bound by a one-way nondisclosure agreement, which stipulates that the research team will not disclose response data “to any other person without [participants’] prior written consent.” We entered into this legally binding agreement (and subsequently shared it with potential participants) to increase the likelihood that executives participate in our study and respond truthfully to the sensitive questions therein.

#### **4. Steps Necessary to Collect and Process Data**

See Item 2 of this document for information about our data collection procedures and participant eligibility criteria.

#### **5. Programs/Codes Used and Brief Description**

The three sets of Stata code file contain the required codes to convert different sets of raw Qualtrics data described in Item 2 above into the final datasets used in the analysis and to produce the results presented in the paper. “List Experiment” converts the raw Qualtrics data and produces the list experiment results presented in Tables 3, 5, OA1, and OA2. “Survey” converts the raw Qualtrics data and produces the direct questioning survey results presented in Tables 3, 5, and OA2. “Demographics” combines the list experiment and direct survey raw data and tabulates various descriptive statistics, including those presented in Table 2. The MS Word file “Stata Code Execution Log” contains the execution logs for all three sets of Stata code. Tables 1 and 4 and Figure 2 were produced manually in excel without the assistance of a statistical package.

#### **6. Data and Code Retention**

All authors will retain the data and code for at least six years.