Online Appendix to
“Do PCAOB Inspections Improve the Accuracy of Accounting Estimates?”

Sarah Stuber
Texas A&M University

Chris E. Hogan
Michigan State University

October 2020

TABLE OF CONTENTS

ONLINE APPENDIX A: Data collection procedures

ONLINE APPENDIX B: Examples of findings in PCAOB inspection reports

ONLINE APPENDIX C: Identifying Assumptions

FIGURE C.1: Parallel trend analysis

ONLINE APPENDIX D: Subsample and entropy balancing descriptive statistics

TABLE D.1: Sample descriptive statistics: all banks public and private subsamples
TABLE D.2: Sample descriptive statistics: over-reserved public and private subsamples
TABLE D.3: Sample descriptive statistics: under-reserved public and private subsamples

ONLINE APPENDIX E: Supplemental tables and analyses

TABLE E.1: Correlation Matrix
TABLE E.2: Primary analyses with alternative dependent variables
TABLE E.3: Effect of PCAOB findings on CO, NPA, and ALL levels
TABLE E.4: Public and private bank comparison analysis with bank fixed effects
TABLE E.5: Alternative FINDING_INTENSITY scalar analyses
TABLE E.6: Analyses using unbalanced samples
TABLE E.7: Various clustering
TABLE E.8: Triennially inspected firm descriptive statistics
The data for the calculation of **FINDING_INTENSITY** and **FINDING_INTENSITY_BOTH** are obtained from a manual review of PCAOB inspection reports from 2006-2017. An individual issuer is deemed to have a substantive ALL-related finding if an inspection finding specifically references the ALL and is related to direct testing of the ALL, including the evaluation of either qualitative or quantitative inputs into the ALL calculation. These inputs include items such as evaluation of the loss factors, recalculation of charge-off history, consideration of the charge-off windows and consideration of inputs into asset valuations for determining specific allowances necessary for non-performing assets. An example of a finding considered to be an ALL-related substantive finding is included in Panel B below. An issuer is deemed to have an ALL-related control finding, if there is a specific mention of failure to properly evaluate the design of controls or failure to properly test the controls related to the ALL. Examples of issuers with both control and substantive ALL-related findings is included in Panel C below. Issuers where there were findings noted, but the findings did not mention the ALL calculation or key inputs to the ALL calculation are considered to have non-ALL findings. Due to the clear distinction of control and non-control related findings in the language used by the PCAOB, as well as specific reference to the ALL estimate, classification of the findings involves minimal subjectivity, as seen in the examples below. Each report was initially reviewed and coded by one coauthor for consistency in classification of findings. A second coauthor reviewed a sample of the coded reports for accuracy of classifications.
ONLINE APPENDIX B
Examples of Findings in PCAOB inspection reports

2007 Inspection of McGladrey & Pullen, LLP
April 29, 2008
104-2008-068

Issuer B
In this audit, the Firm failed in the following respects to obtain sufficient competent evidential matter to support its audit opinion –

- In performing procedures related to the ALL –
  - The issuer reduced certain qualitative loss factors used in its ALL calculation in the fourth quarter, notwithstanding indicators of credit deterioration including increasing levels of nonaccrual and impaired loans; high levels of loan policy exceptions; and the ALL at most of its bank subsidiaries that were below peer group averages. **Despite these factors, the Firm failed to evaluate the reasonableness of the qualitative loss factors.**
  - The Firm selected the five largest impaired loans from the total impaired loan population to evaluate whether the recorded amounts exceeded their fair values and whether the issuer properly recognized specific impairment reserves. For certain impaired loans, the issuer used outside specialists to perform appraisals on the underlying collateral, and the Firm used the work of the specialists in its evaluation. However, **the Firm failed to gain an understanding of the methods and assumptions used in the appraisal reports. In addition, the Firm failed to test the valuation of the remaining population of impaired loans.**

Panel C: Substantive and Control Finding Examples

2011 Inspection of Deloitte & Touche LLP
November 28, 2012
104-2012-271

Issuer M

- The Firm failed to perform sufficient procedures to test the effectiveness of ICFR –
  - **The Firm failed to test controls over the completeness and accuracy of reports used in the issuer's controls over the allowance for loan losses ("ALL").**
- **The Firm failed to perform sufficient procedures to test the issuer's ALL.** The Firm evaluated the reasonableness of the issuer's ALL by developing an independent estimate. This estimate was not an adequate test of the ALL, however, as it was based entirely on the issuer's previous year's ALL and the amount of loan charge-offs during the year under audit, and, further, it did not take into account information in the work papers that suggested that the issuer's ALL required further analysis.
Issuer E

- The Firm identified a fraud risk related to the qualitative component of the allowance for loan losses ("ALL"), which represented a significant portion of the total ALL. The Firm failed to perform sufficient procedures with respect to the ALL. Specifically –
  - During the year, the issuer experienced declines in its loan chargeoffs and classified loans, and it decreased the quantitative and specific components of its ALL. The issuer disclosed that the quality of its loan portfolio had improved, noting improved loan underwriting standards, and that economic conditions had improved. The amount of the component of the issuer's ALL that captured non-quantitative and non-specific factors (the qualitative component), however, had increased, which appeared to be inconsistent with the matters described above. The Firm's procedures related to the qualitative component consisted of dual purpose tests. As part of these tests, the Firm selected controls that consisted of management's review of the adjustments from the prior quarters that determined the qualitative component of the ALL. The Firm limited its testing of these controls to inspecting internal and external data and noting that each qualitative adjustment was within the range established by the issuer's policy. The Firm failed to evaluate whether these reviews, or other controls over the qualitative component, took into account all relevant matters, including those described above. In addition, in performing its substantive procedures, including the evaluation of the reasonableness of the qualitative adjustments and the total qualitative component, the Firm failed to take into account the matters described above. (AS No. 5, paragraphs 42 and 44; AU 342, paragraph .11)
As detailed in section 3.3 of the main manuscript, the key identifying assumption for our analysis is that the public (treatment) and private (control) firms share parallel trends in ALL accuracy. As shown in Figure C.1, the levels of $ALL\_QUALITY$ are nearly identical for both the public and the private banks prior to our sample period from 1992-2005. We begin in 1992, as this is the period after FDICIA was passed, which led to greater uniformity in reporting and in the internal control audit requirement for public and private banks. Additionally, this long window allows us to see the trends in ALL accuracy for public and private banks over contractions and expansions in the bank market. This visual test provides comfort that there was not a significant divergence in the trend in ALL accuracy leading up to our sample period that could explain our results.

A second important identifying assumption for a difference-in-differences analysis is that there are no events that coincide with the event of interest that may have a differential effect on the treatment and control group. While the passage of the Sarbanes Oxley Act (SOX), effective in 2004, did have a differential effect on the internal control audits for public and private firms in general, the differential effect of SOX was much less for public and private banks due to FDICIA. Under FDICIA, both public and private banks with assets greater than $500 million were subject to internal control audit requirements from 1992-2005, and those with assets greater than $1 billion are subject to internal control audits throughout our entire sample period. Banks with less than $1 billion in assets are also likely to be below the SOX threshold for requiring audits of internal control, making the effect of SOX minimal for the banks in our sample. Furthermore, the inclusion of $Year \times Region$ fixed effects controls for time varying trends in ALL accuracy at the regional level during our sample period, further mitigating concerns that trend differences between public
and private banks explain our results. Overall, the strengths of using a banking subgroup, combined with several validation tests support the key identifying assumption of parallel trends in ALL accuracy across public and private banks.

Finally, we also perform a supplemental analysis to examine whether there is a relationship between FINDING_INTENSITY and loan portfolio composition and risk characteristics. We regress FINDING_INTENSITY on %MORTGAGE, %C&I, %CONSUMER, %COMMERCIAL, and PORTFOLIO_RISK as well as an interaction of each of these variables with PUBLIC. We find no significant relation between loan portfolio composition and risk variables and FINDING_INTENSITY in general, and no differences in the relation for public vs private banks. Overall, this analysis provides evidence that neither a relationship between underlying loan portfolios and FINDING_INTENSITY nor differences in the relationship between public and private companies explain the results.
FIGURE C.1
Parallel trend analysis

ALL_QUALITY Trend Analysis


Private  Public
Tables D.1, D.2 and D.3 present descriptive statistics for public compared to private bank clients for the sample of all bank-years, over-reserved bank-years and under-reserved bank-years, respectively. As expected, there is a large discrepancy in asset size across these groups. The typical threshold used in assessing normalized differences is 0.25, as differences greater than this may indicate the model is sensitive to specification subjectivity (Imbens and Wooldridge [2009]; Wooldridge [2011]). The normalized difference between the size of public and private banks exceeds 0.25 percent, suggesting that the results of our analysis could be sensitive to research design choices due to limited overlap in bank size. To address these concerns, we perform entropy balancing, balancing public and private clients based on the first three moments of all control variables. In each table, Panel A presents the pre-balanced descriptive statistics, while Panel B presents the descriptive statistics for the entropy-balanced subsample. Normalized differences for the entropy balanced sample are all less than 0.25 threshold. The sample is entropy balanced on three moments, with the exception of the under-reserved sample in Table D.3 which is only balanced on the first moment due to limitations sample size.
ONLINE APPENDIX E
Supplemental tables and analyses

TABLE E.1: Correlation Matrix
TABLE E.2: Primary analyses with alternative dependent variables
TABLE E.3: Effect of PCAOB findings on CO, NPA, and ALL levels
TABLE E.4: Public and private bank comparison analysis with bank fixed effects
TABLE E.5: Alternative FINDING_INTENSITY scalar analyses
TABLE E.6: Analyses using unbalanced samples
TABLE E.7: Analyses using various clustering
TABLE E.8: Triennially inspected firm descriptive statistics