Do PCAOB Inspections Improve the Quality of Internal Control Audits?

Data Description Sheet (JAR Data Policy)

Item 1.
Clive Lennox is the author that handled the data and conducted the analysis.

Item 2.
Data for the analysis are obtained from the following three publicly available sources at the times indicated:

i. The PCAOB website at http://pcaobus.org/Pages/default.aspx.
We obtain raw data on the PCAOB inspection reports from the PCAOB website. Audit firm inspection reports for years 2005-2014 were downloaded and coded over several years as they became available to the public. These data are included in the raw data file "ICFR".

ii. Compustat.
The Compustat database was used to collect information about the control variables. These data were downloaded on Dec 4, 2014 and are contained in the raw data files "controls" and "segments".

iii. Audit Analytics at http://www.auditanalytics.com/index.php
The Audit Analytics database was used to collect internal control audit opinions (1/1/2004 - 12/31/2014), other audit data (1/1/2004 - 12/31/2014), and accounting restatements (1/1/2004 - 5/14/2015). The internal control opinions were downloaded on Nov 6, 2014 and are contained in eleven raw data files named "opinion2004" to "opinion2014". All other audit data were downloaded on Nov 6, 2014 and are contained in eleven raw data files named "2004" to "2014". The restatements data were first downloaded on Dec 4, 2014 and subsequently updated though May 14, 2015 and are contained in the raw data file "misstate". The auditor change data were first downloaded on Dec 5, 2014 and are contained in the raw data file "audit-change".

Item 3.
Not applicable.

Item 4.
Please see the attached STATA program named "master_prog". The program is used to import and merge all the raw data files, construct all variables used in the analysis, construct the samples used in the analysis, and compute the results reported in the paper. When opened in Microsoft WORD the program is 29 pages long and the program tasks are divided up as follows:

- Pages 1 to 4:
    - Import and merge the raw data files for internal control opinions and other audit data ("opinion2004" to "opinion2014" and "2004" to "2014").
    - Because accounting restatements can prompt revisions to previously issued clean internal control opinions, we use only the originally-issued internal control opinion in cases where the internal control opinion has been subsequently restated.

    - In cases of audit firm mergers, the “auditorkey” variable from Audit Analytics is transformed to be consistent with the “auditorkey” variable in the “ICFR” file.
    - The resulting data file is named as “opinion”.
- Pages 5 to 10:
    - Import the “ICFR” data file.
    - Each audit firm has a maximum of ten inspection reports. Create ten separate data files (“report1.dta” to “report10.dta”) which correspond to the audit firm’s first inspection, second inspection, …, tenth inspection respectively.
    - Merge the ten inspection files (“report1” to “report10”) with “opinion.dta”.
    - The resulting data file is named as “opinion2”.
- Pages 11 to 19:
    - Create all the inspection variables that correspond to each audit firm’s post-inspection windows and pre-inspection windows.
    - The resulting data file is named as “opinion3”.
- Pages 20 to 24:
    - Import the “controls” file. Create control variables and save the resulting data file with the name “controls1”.
    - Import the “segments” file. Create the business segments, geographic segments, and operating segments variables, which are saved with the names “busseg”, “geoseg”, and “opseg”.
    - Import the “audit-change” file. The maximum number of audit resignations per company is six. Therefore, we create six resignation data files “resign1” to “resign6” with one resignation per company in each file.
    - Merge the “opinion3” data file with the ten previously created data files: “controls1”, “busseg”, “geoseg”, “opseg”, “resign1” to “resign6”.
    - Create the resignations variable, and undertake transformations of the other control variables (e.g., winsorization, logging, scaling, etc…).
    - Merge with “misstate” to obtain the misstatement and restatement variables.
    - Create two variables (“miss” and “miss1”) which are used to identify the sample sizes used in the published tables.
    - The resulting data file is named as “opinion4”.
- Pages 25 to 29:
    - Import the “ICFR” file.
    - Generate the descriptive statistics on PCAOB inspections that are reported in Tables 1 and 2.
    - Import the “opinion4” file.
    - Generate the descriptive statistics on internal control opinions that are reported in Tables 1 and 2.
    - Generate the univariate and regression results that are reported in Tables 3-9.

To facilitate replication, please note that the correspondence between the variable names in the data files and the variable names used in the paper are as follows:

ICFR data file:

- audit firm (name of audit firm)
- auditorkey (audit firm identifier)

- report_no (an identifier indicating whether it is the audit firm's first, second, third, etc. inspection)
- no_reports (total number of inspection reports received by the audit firm through the end of 2013)
- inspect_begin (the date the inspection begins)
- inspect_end (the date the inspection ends)
- inspect_year (the inspection year)
- pcaob_report (the date the inspection report is issued)
- number of audits examined by inspectors. NB: This item is not publicly disclosed by the PCAOB for the annually inspected audit firms prior to 2010.
- number of audits that are found to have significant deficiencies. This variable includes the total number of internal control audit deficiencies and other types of audit deficiencies identified in the inspection report.
- ICFR (internal controls over financial reporting). This shows the total number of audit engagements with internal control audit deficiencies.
- auditors in the 2010-2013 period. This variable is equal to one if the inspected audit firm issued an internal control audit report for fiscal years 2010-2013.

opinion4 data file:

- fyear: the company's fiscal year
- icop: dummy variable indicating whether the company received an adverse internal control opinion (labelled $ICOP_{i,t}$ in the paper)
- big4: dummy variable indicating whether the company is audited by a Big 4 firm
- annual: dummy variable indicating whether the company is audited by an annually inspected audit firm
- misstate: dummy variable indicating whether the company's financial statements are subsequently restated (labelled as $MISSTATE_{i,t}$ in the paper)
- post_icfr_pct: the number of audits by company *i*'s audit firm that are found by PCAOB inspectors to have internal control deficiencies, divided by the total number of audits examined in the inspection, where the PCAOB inspection report is issued at the start of the post-inspection window and prior to the auditor's issuance of the internal control opinion for year *t* (labelled $DEF_IC\%_{i,t}$ in the paper)
- auditorkey: audit firm identifier used by Audit Analytics
- miss: this variable takes the value zero for all observations that have non-missing values for the models reported in Tables 4, 5, 6, 7 and 9.
- cikcode: the company's CIK identifier
- post_non_icfr_pct: the number of audits by company *i*'s audit firm that are found by PCAOB inspectors to have deficiencies unrelated to internal controls, divided by the total number of audits examined in the inspection (labelled $DEF_NOT_IC\%_{i,t}$ in the paper)
- icop_lag: the lagged value of icop (labelled $ICOP_{i,t-1}$ in the paper)
- recent_resign: equals 1 if company *i* announces an auditor resignation during fiscal year *t*; = 0 otherwise (labelled as $RESIGN_{i,t}$ in the paper)
- lnmktval: Natural log of one plus the market value (which equals marketcap/1,000) for company *i* at the end of fiscal year *t* (labelled as $SIZE_{i,t}$ in the paper).

- lnaf: the natural log of audit fees for company *i* in fiscal year *t* (labelled as $Ln(AF)_{i,t}$ in the paper).
- lnnas: the natural log of one plus the non-audit service fees for company *i* in fiscal year *t* (labelled as $Ln(NAS)_{i,t}$ in the paper).
- loss: equals 1 if company *i* reports a loss during fiscal year *t*; = 0 otherwise (labelled as $LOSS_{i,t}$ in the paper).
- wno_busseg: the winsorized number of reported business segments (Compustat Segment File) reported by company *i* in fiscal year *t* (labelled as $SEGMENTS_{i,t}$ in the paper).
- wsalesg: the winsorized annual percentage growth in sales for company *i* in fiscal year *t* (labelled as $GROWTH_{i,t}$ in the paper).
- wxtfin: the winsorized sum of newly issued equity and debt, divided by total assets for company *i* in fiscal year *t* (labelled as $XTFIN_{i,t}$ in the paper).
- acq: equals 1 if company *i* is involved in a merger or acquisition in fiscal year *t*; = 0 otherwise (labelled as $M\&A_{i,t}$ in the paper).
- restructured: equals 1 if company *i* restructures its operations in fiscal year *t* (non-missing values for RCA, RCEPS, RCP, or RCD); = 0 otherwise (labelled as $RESTRUCTURE_{i,t}$ in the paper).
- inventory: inventory divided by total assets for company *i* at the end of fiscal year *t* (labelled as $INVENTORY_{i,t}$ in the paper).
- foreign: equals 1 if company *i* has a non-zero foreign currency translation adjustment in fiscal year *t*; = 0 otherwise (labelled as $FOREIGN_{i,t}$ in the paper).
- aud_1, aud_2, aud_3, aud_4, aud_5, aud_6, aud_7, aud_9: audit firm dummies (these correspond to the eight annually-inspected auditors)
- y_10, y_11, y_12, y_13: year dummies corresponding to the date that the auditor signed the internal control audit report (y_10 = 2010, y_11 = 2011, y_12 = 2012, y_13 = 2013)
- test_pct: equals $DEF_IC\%_{i,t}$ except it only includes internal control deficiencies in which the auditor failed to adequately *test* internal controls (labelled as $DEF_IC_TEST\%_{i,t}$ in the paper).
- materiality_pct: equals $DEF_IC\%_{i,t}$ except it only includes internal control deficiencies in which the auditor failed to adequately *evaluate the materiality* of identified weaknesses (labelled as $DEF_IC_MATERIAL\%_{i,t}$ in the paper).
- accounts_pct: equals $DEF_IC\%_{i,t}$ except it only includes internal control deficiencies in which the auditor failed to adequately test internal controls relating to specific *account balances* (labelled as $DEF_IC_TEST_AC\%_{i,t}$ in the paper)
- it_pct: equals $DEF_IC\%_{i,t}$ except it only includes internal control deficiencies in which the auditor failed to adequately test general or application controls for the company's *IT system* (labelled as $DEF_IC_TEST_IT\%_{i,t}$ in the paper).
- work_others_pct: equals $DEF_IC\%_{i,t}$ except it only includes internal control deficiencies in which the auditor failed to adequately test internal controls due to the auditor's over-reliance on the *work of others* (labelled as $DEF_IC_TEST_OTHERS\%_{i,t}$ in the paper).
- future_restate: equals one if company *i* announces an accounting restatement in the year following the auditor's issuance of the internal control opinion, zero otherwise (labelled as $RESTATE_{i,t+1}$ in the paper).

- preins_pct: equals one if the auditor issues an adverse internal control report to company *i* in year *t*, where year *t* belongs to the pre-inspection window, zero otherwise (labelled as $PRE_ICOP_{i,t}$ in the paper).
- miss1: this variable takes the value zero for all observations that have non-missing values for the models reported in Table 8.
- lnaf_lag: the lagged value of lnaf (labelled as $Ln(AF)_{i,t}$ in the paper).

Item 5.
The computer program ("master_prog.do.txt") used to convert the raw data into the final dataset used in the analysis ("opinion4") is available from JAR. The authors have also provided JAR with all the raw data files (see Item 2 above).

Item 6.
The data and program will be maintained by Clive Lennox for at least six years.