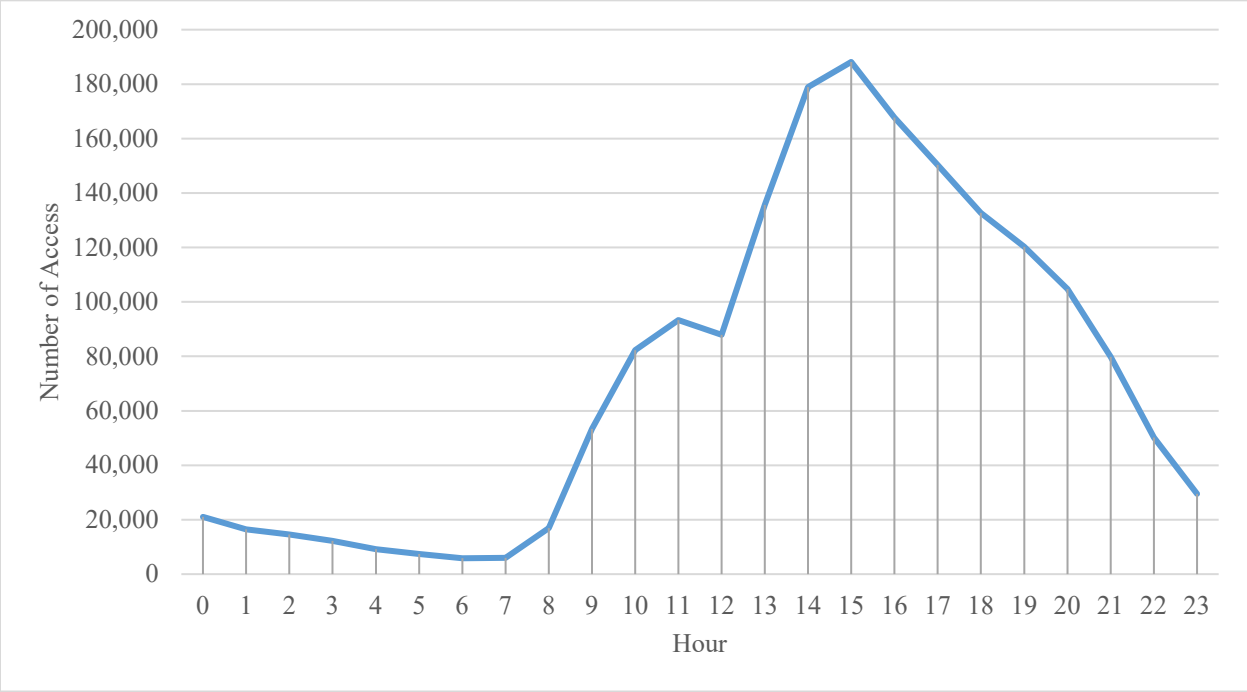
Internet Appendix (IA) For "Deciphering the PCAOB Inspection Process: Evidence and Predictive Insights from Public Data"

Daniel Aobdia, Edward Xuejun Li, K. Ramesh, and Min Shen

Time of day distribution of PCAOB searches

In Figure IA1, we plot the distribution of the page clicks over the time of day. It demonstrates a pattern similar to human activities, with the most intensive search occurring during regular working hours with a dip around lunchtime (Li, Ramesh, and Shen 2011; Li, Lind, Ramesh, and Shen 2023).

FIGURE IA1
Frequency of PCAOB Searches by Time of Day



This figure plots the distribution of PCAOB searches by hours of a day during 2007-2016.

Analysis of PCAOB inspection report lag

We examine at what stage of the inspection the PCAOB relies on SEC filings. Specifically, the time it takes for the PCAOB to release its inspection report following completion of the fieldwork (PCAOB report lag) is the outcome variable of interest. We regress the PCAOB report lag of triennial firm inspections on total searches conducted prior to the inspection fieldwork (*Total Searches*$_{pre\text{-}insp}$), during the inspection fieldwork (*Total Searches*$_{dur\text{-}insp}$), and after the inspection fieldwork (*Total Search*$_{post\text{-}insp}$). We also include as control variables the proportion of inspected audits that are deficient (*Prop. of Deficient Audits*) and the number of audits inspected (*Ln(Audits Inspected)*) as measures of inspection complexity.

Table IA1 shows a negative association between *Total Searches*$_{pre\text{-}insp}$ and report lag, which suggests that more preparation pre-inspection fieldwork helps accelerate the writing and release of the report. We also find a positive association between *Total Search*$_{post\text{-}insp}$ and the report lag, which is consistent with the PCAOB accessing SEC filings for report writing purposes and might indicate that the PCAOB attempts to triangulate its findings with the audit clients' SEC disclosures or incorporates relevant client-specific information in the inspection report. Finally, consistent with heightened inspection complexity, we find that the report lag increases in *Prop. of Deficient Audits* and *Ln(Audits Inspected)*. Taken together, the results in Tables 2 and IA1 validate that the PCAOB indeed relies on SEC filings for the selection of audit engagements, during fieldwork to facilitate its inspection, and post fieldwork to structure its report.

TABLE IA1
Test of PCAOB's Reliance on SEC Filings for Monitoring
Poisson Regression Results of PCAOB Inspection Report Lag on Auditor-Level PCAOB Search of Clients of Triennial Auditors
(1,192 Inspection Reports)

	Total	Cur10K	Prior10K	10Q	8K	REG	CMT	Other
Ln(Total Search$_{pre\text{-}insp}$)	-0.032***	-0.027**	-0.032***	-0.022**	-0.025***	-0.003	-0.032***	-0.012
	(-3.36)	(-2.50)	(-3.31)	(-1.97)	(-2.72)	(-0.21)	(-2.58)	(-0.99)
Ln(Total Search$_{dur\text{-}insp}$)	-0.012	-0.014	-0.031**	-0.006	-0.017	-0.034	-0.003	0.009
	(-1.17)	(-0.96)	(-2.58)	(-0.35)	(-1.50)	(-1.33)	(-0.17)	(0.51)
Ln(Total Search$_{post\text{-}insp}$)	0.127***	0.128***	0.119***	0.089***	0.104***	0.065***	0.055***	0.083***
	(11.50)	(9.42)	(9.92)	(7.60)	(9.98)	(4.57)	(4.34)	(6.72)
Ln(Inspection Length)	0.078***	0.076**	0.089***	0.091***	0.077***	0.089***	0.077**	0.077**
	(2.76)	(2.52)	(2.97)	(2.93)	(2.61)	(2.76)	(2.42)	(2.43)
Prop. of Deficient Audits	0.588***	0.653***	0.657***	0.689***	0.678***	0.769***	0.759***	0.736***
	(14.92)	(17.38)	(17.25)	(18.49)	(18.81)	(22.15)	(21.48)	(21.05)
Ln(Audits Inspected)	0.118***	0.138***	0.152***	0.176***	0.149***	0.212***	0.237***	0.170***
	(3.91)	(4.58)	(5.23)	(5.71)	(4.81)	(7.12)	(7.71)	(5.80)
Prop. NV Clients	0.034	0.078	0.039	0.091*	0.063	0.120**	0.132***	0.119**
	(0.71)	(1.55)	(0.76)	(1.83)	(1.27)	(2.49)	(2.79)	(2.44)
Prop. DE Clients	0.093**	0.102**	0.086*	0.102**	0.099**	0.113**	0.122***	0.105**
	(2.05)	(2.26)	(1.89)	(2.20)	(2.12)	(2.48)	(2.71)	(2.33)
Year Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents Poisson regression results of PCAOB inspection report lag on PCAOB search by SEC filings/forms. We restrict this analysis to inspections on triennial auditors. The dependent variable is the time lag from the end of the inspection fieldwork to the issuance date of the PCAOB inspection report. We exclude weekends and holidays when calculating the time lag. *Ln(Total Search$_{pre\text{-}insp}$)*, *Ln(Total Search$_{dur\text{-}insp}$)*, and *Ln(Total Search$_{post\text{-}insp}$)* are the natural logarithms of PCAOB total hits + 1 before, during, and post the inspection fieldwork, and *Ln(Inspection Length)* is the natural logarithm of PCAOB inspection period in days + 1. We define pre-inspection period as time duration from the beginning of a year through one day prior to the beginning of PCAOB fieldwork. Inspection period is the time duration from the beginning through the end of PCAOB fieldwork. Post-inspection period is time duration from one day after PCAOB fieldwork through the issuance date of the PCAOB inspection report. *Prop. of Deficient Audits* is the number of deficient audit engagements, scaled by the total number of engagements inspected, while *Ln(Audits Inspected)* is the natural logarithm of the total number of engagements inspected by PCAOB. *Prop. NV Clients* and *Prop. DE Clients* are the proportion of audited issuers incorporated in the state of Nevada and Delaware, respectively. We report the estimated Poisson coefficients and include z-statistics based on auditor-clustered standard errors in parentheses. In each regression, we control for calendar year effects. Variable definitions and data sources are detailed in Appendix.

Factor analysis on auditor and client characteristics

We use factor analysis to reduce the dimensionality of the 15 variables measuring auditor-client relationship and traditional client characteristics potentially associated with high-risk audits. Specifically, we run a principal factor analysis on the following variables: *Big4*, *Ln(Auditfees)*, *Ln(Nonauditfees)*, *Integrated Audit*, *LengthRelationship*, *Ln(Asset)*, *High Tech*, *Sales Growth*, *stdCFOat*, *CFOat*, *ForeignPifo*, *nBUSSEG*, *nGEOSEG*, *BTM*, and *Leverage*. We retain the one factor with an eigenvalue greater than one and label this factor as "prominence."[A1] This label underscores the factor's role in capturing the relative prominence and significance of the client, which is likely associated with heightened PCAOB scrutiny. The first factor explains 82% of the total variation in the original 15 variables. The factor loadings are presented in Table IA2.

Table IA2
Principal Factor Analysis of Auditor and Client Characteristics Proxying for High-Risk Engagements
(46,101 Issuer-Years)

	Factor Loading[A2]
Big4	**0.7017**
Ln(Auditfees)	**0.9061**
Ln(Nonauditfees)	**0.5394**
Integrated Audit	**0.6930**
LengthRelationship	**0.4454**
Ln(Asset)	**0.8740**
High Tech	-0.1293
Sales Growth	-0.1177
stdCFOat	**-0.3154**
CFOat	**0.3969**
ForeignPifo	0.2728
nBUSSEG	**0.3433**
nGEOSEG	**0.3026**
BTM	0.1281
Leverage	0.1930
Variance Explained	82.11%

[A1] The eigenvalue drops precipitously from 3.67 for the first factor to 0.78 for the second factor.
[A2] Factor loadings with a magnitude greater than 0.3 are in bold.

Alternative measures of PCAOB search

In Table IA3, we repeat the analysis in Table 3 using PCAOB search volume instead of the number of weeks searched by PCAOB. Our main findings continue to hold for this alternative measure.

TABLE IA3
Test of Predicting Issuer-Level PCAOB Search
Poisson Regression Results of Issuer-Level PCAOB Search Volume on Auditor and Client Attributes and Clients' Corporate Events

Panel A. Regression result – The baseline model

	(1) All	(2) Big4	(3) Next 5	(4) Triennial Insp Years	(5) Triennial Non-Insp Years
Prominence	0.551***	0.862***	0.737***	0.277***	0.230***
	(26.11)	(30.93)	(13.41)	(4.62)	(3.15)
Audchg	0.643***	0.646***	0.594***	0.383***	0.538***
	(16.55)	(11.66)	(7.45)	(4.61)	(5.30)
Restate	0.337***	0.267***	0.388***	0.487***	0.731***
	(10.98)	(7.39)	(4.17)	(5.55)	(6.40)
Bankrupt	0.418***	0.719***	0.587	0.197	-0.188
	(2.70)	(4.47)	(1.32)	(0.31)	(-0.48)
Litigation	0.353***	0.360***	-0.030	0.351*	0.717***
	(6.15)	(6.01)	(-0.19)	(1.74)	(2.76)
Commlett	0.146***	0.105***	0.163***	0.184***	0.208***
	(8.43)	(5.34)	(3.44)	(3.08)	(2.86)
CEO turnover	0.049*	0.033	0.119	0.064	0.022
	(1.65)	(1.00)	(1.45)	(0.63)	(0.21)
CFO turnover	0.166***	0.151***	0.157**	0.190**	0.166*
	(6.32)	(5.27)	(2.06)	(2.00)	(1.68)
Weakness	0.641***	0.596***	0.605***	0.800***	1.214***
	(14.34)	(11.76)	(5.34)	(5.60)	(6.54)
Going Concern	0.492***	0.290***	0.230*	0.189**	0.188*
	(9.31)	(3.45)	(1.66)	(2.22)	(1.70)
NV	0.680***	0.309**	0.318**	0.451***	0.649***
	(10.19)	(2.44)	(2.37)	(4.63)	(5.34)
DE	0.076***	0.091***	0.049	0.003	0.136*
	(2.83)	(3.23)	(0.79)	(0.04)	(1.76)
TARP	0.257***	0.305***	0.021	-0.320**	-0.436***
	(3.88)	(4.49)	(0.23)	(-2.39)	(-3.41)
Non-TARP Bank	0.006	0.223***	-0.130*	-0.390***	-0.342***
	(0.14)	(3.18)	(-1.71)	(-3.90)	(-3.62)
Year effects & December Year End	Yes	Yes	Yes	Yes	Yes
Obs	46,101	28,083	6,285	3,898	7,835
R^2	0.2188	0.2760	0.1821	0.1814	0.1741

(Continued on the next page)

(TABLE IA3 continued)

Panel B. Regression result – The expanded model

	(1) All	(2) Big4	(3) Next 5	(4) Triennial Insp Years	(5) Triennial Non-Insp Years
nonBig4-to-Big4	0.732***	1.053***			
	(7.59)	(10.08)			
Big4-to-Big4	0.351***	0.501***			
	(5.50)	(7.94)			
Big4-to-nonBig4	0.624***		0.574***	-0.054	0.010
	(7.60)		(5.49)	(-0.33)	(0.05)
nonBig4-to-nonBig4	0.797***		0.654***	0.430***	0.594***
	(12.25)		(5.15)	(4.82)	(5.48)
⋮					
Restate_accounting	0.318***	0.261***	0.280***	0.480***	0.747***
	(10.00)	(7.13)	(2.88)	(5.25)	(6.35)
Restate_fraud	0.681***	0.486**	1.231***	0.141	0.061
	(4.15)	(2.46)	(3.12)	(0.55)	(0.12)
Restate_cler	-0.010	0.003	0.528*	0.146	-0.211
	(-0.08)	(0.02)	(1.86)	(0.31)	(-0.97)
⋮					
TARP	0.235	0.356**	-0.483	-0.383	0.261
	(1.61)	(2.30)	(-1.39)	(-1.47)	(0.84)
TARP 2008	0.558***	0.613***	0.972**	0.311	-0.444
	(4.48)	(4.28)	(2.35)	(0.61)	(-1.08)
TARP 2009	0.535***	0.527***	0.922***	0.136	-0.082
	(4.42)	(3.92)	(2.64)	(0.44)	(-0.15)
TARP 2010	0.020	0.044	0.544	0.256	-0.992**
	(0.13)	(0.23)	(1.36)	(0.64)	(-2.44)
TARP 2011	0.037	0.027	0.533	-0.017	-1.364***
	(0.26)	(0.16)	(1.18)	(-0.04)	(-4.06)
TARP 2012	0.097	0.004	0.520	0.106	-0.730*
	(0.59)	(0.02)	(1.11)	(0.30)	(-1.67)
TARP 2013	-0.600***	-0.618**	-0.578	0.640	-0.921*
	(-3.06)	(-2.54)	(-1.31)	(1.09)	(-1.93)
TARP 2014	-0.396**	-0.546***	0.133	-0.406	-1.053*
	(-2.37)	(-2.76)	(0.30)	(-1.22)	(-1.85)
TARP 2015	0.012	-0.175	0.713*	0.290	-0.367
	(0.08)	(-0.93)	(1.76)	(0.97)	(-0.97)
TARP 2016	-0.195	-0.341*	0.404	-0.206	-0.514
	(-1.21)	(-1.72)	(1.13)	(-0.63)	(-1.60)
⋮					
Other variables	Yes	Yes	Yes	Yes	Yes
Year effects & December Year End	Yes	Yes	Yes	Yes	Yes
Issuer-year obs	46,101	28,083	6,285	3,898	7,835
R^2	0.2237	0.2836	0.1877	0.1816	0.1723

(Continued on the next page)

(TABLE IA3 continued)

*, **, and *** indicate two-tailed statistical significance of 10%, 5% and 1% levels, respectively.

Panel A and Panel B present the Poisson regression results of the baseline model and expanded model, respectively. For the expanded model, we create separate indicators for auditor changes between Big 4 and non-Big 4 firms, for restatements by types, and for TARP and non-TARP firms by years. For the sake of brevity, in Appendix B we report only the estimated Poisson coefficients of these variables. For each model, we report regression results for (1) all issuer-years, (2) issuer-years audited by Big 4 auditors, (3) issuer-years audited by Next 5 auditors, (4) issuers audited by triennial auditors during PCAOB inspection years, and (5) issuers audited by triennial auditors during no-inspection years. We regress the number of PCAOB searches of all SEC filings/forms on determinants. We report the estimated Poisson coefficients and include z-statistics based on issuer-clustered standard errors in parentheses. R^2 is calculated by squaring the correlation between the predicted PCAOB search, estimated from the Poisson regression, and the actual PCAOB search. In each regression, we control for calendar year effects and December year end. All variables are defined in the Appendix.

Comparison of determinants in our models versus those in Christensen, Newton, and Wilkins (2022)

We follow Aobdia (2018) in identifying traditional client characteristics potentially associated with high-risk audits, which are included in his inspection model. Assuming that Christensen et al. (2022) obtained data from a Big-4 or Next-5 firm, in Table IA4 we compare the determinants of our inspection models for Big-4 clients in Table 6 with their Model (3) in Table 3 (hereafter, CNW model). They include two variables that they obtained from the audit firm (*PastPCAOBInspect* and *PastPartnerMatters*), with *PastPCAOBInspect* being statistically significant at the 0.01 level. Firm characteristics in CNW model but not included in our model are *Intangible*, *AR_Inventory*, *Dacc*, *Loss*, and *Issue*. Only *AR_Inventory* and *Dacc* are statistically significant at the 0.05 and 0.01 level, respectively. Auditor characteristics in CNW model but not included in our model are *IndustryExpert, OfficeSize*, and *NewLeadPartner*, with only *OfficeSize* being statistically significant at the 0.05 level.

Consistent with Aobdia (2018) and given our focus on a broader cross-section beyond manufacturing and retail firms, we do not include *AR_Inventory*, and *Dacc*. However, we include 12 client characteristics not present in the CNW model, with 10 being statistically significant in the group of Big 4 clients. Among variables common to both studies, key variables such as material weakness, restatements, and size are highly significant only in our estimations. They report an AUC of 0.78, but our estimation model for the 2016 KPMG sample has an AUC of 0.961. Although they are not directly comparable, given the predictive nature of the models in both studies, we believe we have a robust set of predictors.

Table IA4
A Comparison of the Determinants between the Table 6 Inspection Model and Table 3 Model (3) of Christensen, Newton, and Wilkins (2022)

	CNW Model (3), Table 3		Inspection Model, Table 6	
	Sign	Significant	Sign	Significant
Size	+	No	+	Yes
Leverage	+	Yes	+	Yes
Segments	-	No	-	Yes
Intangible	+	No		
AR_Inventory	+	Yes		
Dacc	+	Yes		
CashFlow	+	Yes	+/-	No
Loss	+	No		
Issue	-	No		
MTB[*]	-	No	-	Yes
AccFiler	+	Yes	+	No
MW	-	No	+	Yes
PastPCAOBInspect	-	Yes		
RestateAnnounce	+	No	+	Yes
GC	+	No	+	No
AuditFees	+	Yes	-	Yes
IndustryExpert	-	No		
OfficeSize	-	Yes		
PastPartnerMatters	+	No		
DecYE	+	No	-	Yes
FirstYear[*]	+	No	+	Yes
NewLeadPartner	-	No		
NonAuditFees	+	No	+	Yes
Audchg			+	Yes
Bankrupt			+	Yes
Litigation			+	Yes
Commlett			-	Yes
CEO turnover			-	No
CFO turnover			+	Yes
NV			+/-	No
DE			-	Yes
High Tech			-	Yes
Sales Growth			+	Yes
stdCFOat			-	Yes
ForeignPifo			-	Yes

[*]Our measures corresponding to *MTB* and *FirstYear* are *BTM* and *LengthRelationship*, respectively, which are supposed to have the opposite signs. We therefore flip the signs of coefficients for these two variables in our baseline model.

Additional proxies for PCAOB proactivity in selecting non-compliant audits

In Table IA5, we include additional variables proxying for PCAOB proactivity in selecting inspected engagements that are more likely to be non-compliant. Beyond the baseline model determinants, such as whether the issuer is incorporated in Nevada (*NV*) or participated in the TARP program (*TARP*), we add the following variables: an indicator for combined audits of financial statements and internal controls (*Integrated Aud*it),[A3] an indicator for firms in the top quintile of book-to-market ratio representing high-value stocks (*BTMQ5*), Piotroski's (2000) F-score as a measure of a firm's financial health (*F-score*), an interaction term between *BTMQ5* and the demeaned *F-score* (the *F-score* is demeaned in the interaction term for ease of interpretation), an indicator for earnings surprises within the range [0, 3] cents (*MOB0_3*), and a loss avoidance indicator for EPS within the range [0, 3] cents (*EPS0_3*).[A4]

As expected, *Integrated Audit* is positively associated with PCAOB search intensity across the full sample and all auditor groups. Integrated audits must abide by additional PCAOB standards on the audit of internal controls over financial reporting (i.e., AS 2201) and thus are more likely to be non-compliant. For the full sample and Big 4 clients, we find that PCAOB tends to pay less attention to issuers with strong financial health (higher *F-score*), evidenced by a negative coefficient on *F-score*, particularly those classified as value stocks (see Piotroski 2000), evidenced by a negative coefficient on the *F-score* interaction with *BTMQ5*. Additionally, there is evidence of increased PCAOB scrutiny for issuers that just meet or beat earnings benchmarks, suggesting a focus on potential earnings management. However, this pattern is mostly evident in the full sample and the Big 4 group. Importantly, our main inferences drawn from the original baseline model remain robust. In addition, the increase in R-squared compared with Panel A of Table 3 is very modest.

[A3] Given *Integrated Audit* is among the variables included to construct the factor score (*Prominence*) used in our primary analysis (Table 3), we re-estimate the factor score variable for this analysis after excluding *Integrated Audit*.

[A4] Piotroski's (2000) F-score is calculated as the sum of indicators derived from the level or change in financial characteristics or comparing the level of two characteristics at the firm-year level. It is not impacted by inter-industry differences and does not rely on variables specific to certain industries or business models (e.g., number of days of receivables outstanding in Beneish's M-score). As a result, the F-score is straightforward to apply across a broad cross-section of firms.

TABLE IA5
Poisson Regression Results of Baseline Model with Additional Determinants Proxying for Non-Compliant Audits

	(1) All	(2) Big4	(3) Next 5	(4) Triennial Insp Years	(5) Triennial Non-Insp Years
Prominence[a]	0.509***	0.705***	0.550***	0.125***	0.153***
	(35.57)	(50.87)	(12.09)	(3.03)	(2.84)
Audchg	0.436***	0.465***	0.363***	0.260***	0.349***
	(19.90)	(13.70)	(8.24)	(5.63)	(6.55)
Restate	0.218***	0.178***	0.245***	0.317***	0.516***
	(12.99)	(9.16)	(4.61)	(5.99)	(8.72)
Bankrupt	0.091	0.280**	0.528***	0.468	-0.352
	(0.90)	(2.54)	(3.02)	(0.88)	(-1.15)
Litigation	0.226***	0.207***	0.088	0.430***	0.411**
	(6.49)	(5.65)	(0.93)	(3.79)	(2.32)
Commlett	0.079***	0.058***	0.083***	0.071**	0.077*
	(8.05)	(5.15)	(3.00)	(2.02)	(1.93)
CEO turnover	0.015	0.008	0.037	0.019	0.052
	(0.94)	(0.42)	(0.74)	(0.36)	(0.93)
CFO turnover	0.069***	0.058***	0.060	0.060	0.160***
	(4.82)	(3.69)	(1.36)	(1.15)	(3.01)
Weakness	0.345***	0.312***	0.288***	0.330***	0.607***
	(13.62)	(10.27)	(4.84)	(3.96)	(6.22)
Going Concern	0.381***	0.183***	0.213***	0.159***	0.144**
	(12.54)	(3.82)	(2.87)	(3.12)	(2.54)
NV	0.456***	0.184**	0.311***	0.379***	0.358***
	(11.45)	(2.57)	(3.50)	(6.25)	(5.28)
DE	0.054***	0.059***	0.079**	0.053	0.066
	(3.54)	(3.69)	(2.25)	(1.18)	(1.35)
TARP	0.227***	0.269***	0.055	-0.192**	-0.214***
	(5.91)	(7.00)	(0.91)	(-2.46)	(-2.67)
Non-TARP Bank	0.056**	0.215***	0.031	-0.256***	-0.220***
	(2.09)	(5.31)	(0.60)	(-4.30)	(-3.73)
Integrated Audit	0.089***	0.082***	0.292***	0.327***	0.342***
	(3.97)	(3.11)	(6.42)	(6.31)	(5.01)
BTMQ5	0.002	0.038**	-0.011	0.017	-0.042
	(0.14)	(2.14)	(-0.32)	(0.40)	(-0.84)
F-score	-0.027***	-0.030***	-0.000	-0.004	-0.005
	(-6.79)	(-6.62)	(-0.02)	(-0.34)	(-0.38)
BTMQ5 × F-score	-0.015*	-0.033***	-0.031	0.006	-0.013
	(-1.83)	(-3.03)	(-1.62)	(0.28)	(-0.57)
MOB0_3	0.026**	0.028**	-0.000	0.003	0.004
	(2.10)	(2.05)	(-0.01)	(0.07)	(0.07)
EPS0_3	0.141***	0.079	0.017	0.007	-0.005
	(3.16)	(1.07)	(0.13)	(0.09)	(-0.06)

(Continued on the next page)

(TABLE IA5 continued)

Year effects & December Year End	Yes	Yes	Yes	Yes	Yes
Obs	46,039	28,045	6,281	3,895	7,818
R^2	0.4587	0.5051	0.4051	0.3383	0.2953

*, **, and *** indicate two-tailed statistical significance of 10%, 5% and 1% levels, respectively.
[a] Prominence is extracted from a principal factor analysis excluding *Integrated Audit*. The factor loadings on the remaining 14 variables are qualitatively similar and we continue to retain only one factor with an eigenvalue greater than one.

This table presents the Poisson regression results of the number of PCAOB search weeks of all SEC filings/forms on the baseline model determinants and additional variables that proxy for non-compliant audits: *Integrated Audit* is an indicator for combined audit of financial statements and internal controls. *BTMQ5* is an indicator for the top quintile of book-to-market ratio. *F-score* is a rank variable ranging from 0 to 9 based on nine indicators: (1) positive return on assets (ROA), calculated as net income before extraordinary items scaled by beginning-of-the-year total assets, in the current year, (2) positive operating cash flow in the current year, (3) positive change in ROA in the current year; (4) positive accrual, calculated as operating cash flow minus net income, in the current year, (5) a decrease in leverage, calculated as long-term debt relative to average total assets, in the current year compared to the previous year, (6) an increase in liquidity, calculated as current assets over current liabilities, in the current year compared to the previous year, (7) no common stock issuance in the current year, (8) an increase in gross margin, calculated as gross profit scaled by sales, in the current year compared to the previous year, and (9) an increase in asset turnover, calculated as sales relative to beginning-of-the-year total assets, in the current year compared to the previous year. The higher *F-score* a firm receives, the better its financial condition is. *MOB0_3* is an indicator for earnings surprise between 0 and 3 cents, where earnings surprise is defined as the actual EPS minus the latest mean EPS forecast by analysts prior to earnings announcement, both are obtained from unadjusted I/B/E/S databases. If any I/B/E/S variable is missing, we replace with change in EPS obtained from Compustat. *EPS0_3* is an indicator for unadjusted EPS between 0 and 3 cents. We report the estimated Poisson coefficients and include z-statistics based on issuer-clustered standard errors in parentheses. R^2 is calculated by squaring the correlation between the predicted PCAOB search, estimated from the Poisson regression, and the actual PCAOB search. In each regression, we control for calendar year effects and December year end. All other variables and data sources are detailed in the Appendix.

Average marginal effects of Table 6 PCAOB inspection model

TABLE IA6
Estimating PCAOB Inspection of Big 4 Clients
Average Marginal Effects of Probit Regressions of Predicting Inspected Big 4 Clients (Excluding 2016 KPMG Clients) at Different Cutoffs

	(1) Fractional Reg. of Pr (Insp)	(2) Probit of Pr (Insp) ≥ 7%	(3) Probit of Pr (Insp) ≥ 8%	(4) Probit of Pr (Insp) ≥ 9%	(5) Probit of Pr (Insp) ≥ 10%	(6) Probit of Pr (Insp) ≥ 11%
Auditor Change	0.023***	0.049***	0.051***	0.051***	0.050***	0.046***
	(7.89)	(6.44)	(7.03)	(7.33)	(7.41)	(6.76)
Restate	0.023***	0.069***	0.067***	0.065***	0.062***	0.057***
	(13.03)	(15.59)	(15.55)	(15.63)	(15.23)	(14.45)
Bankrupt	0.018*	0.020	0.023	0.026	0.030	0.030
	(1.72)	(0.54)	(0.67)	(0.81)	(0.97)	(1.00)
Litigation	0.015***	0.029***	0.030***	0.026***	0.024***	0.023***
	(4.30)	(3.61)	(3.87)	(3.59)	(3.27)	(3.17)
Commlett_dis	-0.009***	-0.028***	-0.026***	-0.023***	-0.023***	-0.021***
	(-7.93)	(-10.23)	(-9.75)	(-9.08)	(-8.99)	(-8.56)
CEO turnover	-0.004**	-0.012**	-0.013***	-0.013***	-0.013***	-0.012***
	(-2.08)	(-2.51)	(-2.75)	(-2.82)	(-2.89)	(-2.85)
CFO turnover	0.026***	0.079***	0.074***	0.070***	0.065***	0.062***
	(20.06)	(22.81)	(22.22)	(21.67)	(20.49)	(20.12)
Weakness	0.019***	0.049***	0.047***	0.044***	0.041***	0.041***
	(5.84)	(6.34)	(6.31)	(6.09)	(5.72)	(5.93)
Going Concern	0.007	0.025*	0.022	0.023	0.020	0.022
	(1.21)	(1.65)	(1.48)	(1.60)	(1.42)	(1.64)
Ln(Auditfees)	-0.004***	-0.016***	-0.015***	-0.014***	-0.013***	-0.013***
	(-2.75)	(-4.05)	(-3.91)	(-3.57)	(-3.40)	(-3.24)
Ln(Nonauditfees)	0.031***	0.079***	0.076***	0.072***	0.069***	0.067***
	(12.40)	(9.37)	(9.02)	(8.47)	(8.05)	(7.71)
Integrated Audit	0.005	0.012	0.011	0.013*	0.011	0.013*
	(1.50)	(1.44)	(1.46)	(1.69)	(1.39)	(1.72)
Auditor Tenure	-0.005***	-0.020***	-0.018***	-0.016***	-0.015***	-0.014***
	(-5.04)	(-9.09)	(-8.55)	(-7.99)	(-7.68)	(-7.31)
Ln(Asset)	0.030***	0.080***	0.076***	0.072***	0.069***	0.065***
	(15.99)	(23.87)	(23.06)	(21.87)	(21.16)	(19.83)
High Tech	-0.020***	-0.059***	-0.055***	-0.052***	-0.050***	-0.049***
	(-8.18)	(-12.15)	(-11.43)	(-10.82)	(-10.59)	(-10.33)
Sales Growth	0.004***	0.012***	0.011***	0.010***	0.010***	0.009***
	(7.19)	(7.48)	(7.17)	(6.97)	(6.62)	(6.08)
stdCFOat	-0.020**	-0.085***	-0.077***	-0.068***	-0.058***	-0.049***
	(-2.50)	(-5.33)	(-4.84)	(-4.47)	(-4.08)	(-3.67)
CFOat	-0.004**	-0.013***	-0.013***	-0.013***	-0.012***	-0.011***
	(-2.43)	(-4.55)	(-4.58)	(-4.70)	(-4.42)	(-4.50)
ForeignPifo	-0.027***	-0.075***	-0.070***	-0.065***	-0.061***	-0.056***
	(-9.29)	(-11.99)	(-11.65)	(-11.17)	(-10.76)	(-10.33)

(Continued on the next page)

(TABLE IA6 continued)

nBUSSEG	-0.007***	-0.020***	-0.018***	-0.016***	-0.015***	-0.014***
	(-6.53)	(-9.55)	(-8.93)	(-8.22)	(-7.99)	(-7.55)
nGEOSEG	-0.017***	-0.040***	-0.038***	-0.036***	-0.034***	-0.033***
	(-9.27)	(-11.09)	(-10.57)	(-10.29)	(-9.91)	(-9.75)
BTM	0.006***	0.011***	0.010***	0.009***	0.009***	0.008***
	(7.97)	(6.25)	(5.91)	(5.75)	(5.15)	(5.17)
Leverage	0.005***	0.010***	0.010***	0.009***	0.008***	0.008***
	(4.66)	(4.20)	(4.28)	(3.98)	(3.75)	(3.84)
NV	0.007*	0.018	0.017	0.015	0.014	0.015
	(1.95)	(1.51)	(1.55)	(1.40)	(1.36)	(1.56)
DE	-0.025***	-0.075***	-0.068***	-0.063***	-0.060***	-0.056***
	(-15.90)	(-21.34)	(-19.80)	(-18.45)	(-17.52)	(-16.83)
DecYe	-0.034***	-0.086***	-0.082***	-0.078***	-0.073***	-0.070***
	(-18.00)	(-21.85)	(-21.00)	(-20.15)	(-19.12)	(-18.40)
Observations	27,469	27,469	27,469	27,469	27,469	27,469
Predicted Inspections (%)		8.75%	8.10%	7.52%	7.11%	6.71%
Predicted Inspections (n)		2,404	2,225	2,065	1,953	1,842
Pseudo R2	0.3156	0.4581	0.4516	0.4468	0.4417	0.4327
AUC		0.9356	0.9349	0.9347	0.9342	0.9327

*, **, and ** indicate two-tailed statistical significance of 10%, 5%, and 1% levels, respectively.

Column (1) presents the fractional probit regression results of predicted inspection on observable corporate events and auditor/client characteristics of 27,469 Big 4 clients other than KPMG clients in 2016. Columns (2)-(6) present the probit regression results of predicted inspections at various cutoffs. The predicted inspection is estimated from Model (3) of Table 5. We standardize all count and continuous variables cross-sectionally within the Big 4 sample to have a mean zero and a standard deviation of one for each year. We report the average marginal effects of estimated coefficients and include z-statistics based on issuer-clustered standard errors in parentheses. Variable definitions and data sources are detailed in Appendix.

Analysis of PCAOB enforcement

Although only 379 out of 46,101 firm-years are potentially influenced by the enforcement investigations, the PCAOB is likely to examine SEC filings during the period when the enforcement is underway. To measure its economic significance, we re-estimate our full-sample baseline and expanded Poisson regressions (Table 3) by including an indicator (*enforce_p*) for the period from the year following the related violation period to the year when the enforcement report was released. The coefficients of *enforce_p* are reported in Table IA7. Panel A shows that, after controlling for the determinants of screening and inspection in Table 3, we find that the overall count of searched weeks is higher by over 200% (100 * (exp(1.156) – 1) on average per year for the 379 enforcement cases in our sample in the baseline model. The inferences are similar in the expanded model and when we use the level of search intensity instead of the count of weeks searched (Panels B–D). We also report the results for the current 10K, non-current 10K, 10Q, and 8K. Note that the current 10-K and non-current 10-K are defined with respect to the inspection years. To the extent the PCAOB is searching the filings related to subjects under enforcement proceedings, then we should find larger marginal effects for non-current 10-Ks, given those 10-Ks are likely related to the prior inspections currently under enforcement. This is exactly what we find. Overall, although enforcement impacts only a very small subset of our sample, the PCAOB continues to use the SEC filings when an enforcement action is underway.

TABLE IA7
Analysis of Enforcement Actions
Poisson Regression Results of Issuer-Level PCAOB Search on Indicator for Enforcement Period
(46,101 Issuer-Years)

Panel A. The baseline model of the number of PCAOB searched weeks

	Total	Cur10K	NonCur10K	10K	10Q	8K
enforce_p	1.156***	0.971***	1.549***	1.201***	1.184***	1.525***
	(11.97)	(8.92)	(12.92)	(11.53)	(11.29)	(14.69)
Control	Yes	Yes	Yes	Yes	Yes	Yes

Panel B. The expanded model of the number of PCAOB searched weeks

	Total	Cur10K	NonCur10K	10K	10Q	8K
enforce_p	1.173***	0.997***	1.555***	1.218***	1.207***	1.523***
	(12.63)	(9.52)	(13.41)	(12.00)	(11.51)	(15.38)
Control	Yes	Yes	Yes	Yes	Yes	Yes

Panel C. The baseline model of the number of PCAOB search hits

	Total	Cur10K	NonCur10K	10K	10Q	8K
enforce_p	1.537***	1.012***	1.775***	1.477***	1.322***	1.805***
	(13.50)	(8.28)	(12.64)	(12.05)	(8.72)	(13.80)
Control	Yes	Yes	Yes	Yes	Yes	Yes

Panel D. The expanded model of the number of PCAOB search hits

	Total	Cur10K	NonCur10K	10K	10Q	8K
enforce_p	1.545***	1.043***	1.785***	1.497***	1.335***	1.795***
	(14.20)	(9.13)	(13.19)	(12.78)	(9.02)	(13.84)
Control	Yes	Yes	Yes	Yes	Yes	Yes

*, **, and *** indicate two-tailed statistical significance of 10%, 5% and 1% levels, respectively.

This table presents Poisson regression results of PCAOB searches on enforcement actions. We identify issuers subject to enforcement actions using PCAOB enforcement reports. Panels A and B show the results of the baseline model and expanded model, respectively, of the number of PCAOB searched weeks, while Panels C and D show the results of the baseline model and expanded model, respectively, of PCAOB search hits. For the sake of brevity, we only report the result of *enforce_p*, which is an indicator for the time period from the year following the violation period through the year when the PCAOB enforcement report is released. We report the estimated Poisson coefficients and include z-statistics based on issuer-clustered standard errors in parentheses. All other variable definitions and data sources are detailed in Appendix.

Restatement analysis by form types

In Table IA8, we repeat Table 7 Panel B and Panel C analysis using the number of weeks different types of SEC filings are searched by PCAOB. We report results for the current 10K, non-current 10K, 10Q, 8K, registration form, and comment letter. Our main findings continue to hold for all filing types.

Table IA8
Prediction of Restate Fiscal Years by SEC Form Type
(42,115 Issuer-Years)

	(1) Cur10K	(2) NonCur10K	(3) 10Q	(4) 8K	(5) REG	(6) CMT
Panel A. Probit regression results of Table 7, Model (2)						
counthits_week	0.004***	0.005***	0.007***	0.006***	0.011***	0.011***
	(5.53)	(5.59)	(7.12)	(5.18)	(3.71)	(5.15)
Area Under Curve	0.5698	0.5703	0.5745	0.5697	0.5634	0.5665
Pseudo R^2	0.0078	0.0075	0.0089	0.0072	0.0066	0.0071
Panel B. Probit regression results of Table 7, Model (3)						
Pred counthits_week	0.011***	0.021***	0.019***	0.025***	0.105***	0.086***
	(7.57)	(8.31)	(7.57)	(7.52)	(5.55)	(9.02)
Resid counthits_week	0.001**	0.003***	0.004***	0.004***	0.008**	0.006***
	(2.00)	(2.76)	(3.60)	(2.93)	(2.44)	(2.62)
F test of Pred and Resid	40.05***	49.63***	31.83***	39.15***	25.77***	68.72***
Area Under Curve	0.5857	0.5861	0.5883	0.5807	0.5747	0.5907
Pseudo R^2	0.0109	0.0109	0.0116	0.0096	0.0085	0.0120

*, **, and *** indicate two-tailed statistical significance of 10%, 5% and 1% levels, respectively.

This table presents the probit regression results of Table 7 Models (2) and (3) by SEC form type. For ease of interpretation, we report the average marginal effects of estimated coefficients and include z-statistics based on issuer-clustered standard errors in parentheses. In each regression, we control for year effects and December year-end. For each type of SEC filing, *Pred counthits_week* and *Resid counthits_week* are predicted value and residual value, respectively, both estimated from a Poisson regression of *counthits_week* of a specific SEC form on the variables of the expanded model. All other variables and data sources are detailed in Appendix.

Out-of-sample restatement analysis

We conduct an out-of-sample analysis to determine whether data publicly available to the PCAOB has greater predictive power on deficient audits than PCAOB searches. We regress *F_restate* on the variables from our expanded model in Table 3 using data of four-year estimation periods that end one year prior to each test year during 2012–2016.[A5] In the out-of-sample analysis, we exclude any restatements announced during or one year after the estimation window to ensure that the PCAOB had the same public data as in the estimation model. We then extrapolate the probability of misstatement, *Pr(misstate)*, for each issuer in the test year during 2012–2016, by applying the estimated coefficients from the estimation window.

In Table IA9 we compare the predictive power of *Estimated Pr(misstate)*, a proxy for publicly available information, with *counthits_week*, a proxy for the PCAOB's information set. Both *Estimated Pr(misstate)* and *counthits_week* generally predict misstatements in Panels A and B, for all types of auditors. However, the explanatory power of *Estimated Pr(misstate)* is generally higher, which suggests that PCAOB searches underperform a prediction model based on past public data. Nevertheless, based on Panel C, both *Estimated Pr(misstate)* and *counthits_week* load when included in the same regression. Overall, these results suggest that, based on public data available to the PCAOB at the time of monitoring, the PCAOB's risk-based approach could have targeted deficient audits with more precision.

[A5] For example, for the test year 2012, the regression parameters are estimated using the sample period 2007–2010.

TABLE IA9
Additional Analyses of Restated Fiscal Years
Average Marginal Effects of Probit Regression Results

	Total (19,665 issuer-years)	**Big 4** (11,984 issuer-years)	**Next 5** (2,869 issuer-years)	**Triennial** (4,812 issuer-years)
Mean of DV	6.37%	7.55%	6.27%	3.49%
Panel A. Probit regression of misstatement on estimated probability of misstatement from four-year estimation window				
Pr(misstate)	0.590***	0.712***	0.664***	0.293***
	(8.84)	(6.89)	(3.22)	(3.17)
Year Effects & December Year End	Yes	Yes	Yes	Yes
Area Under Curve	0.6141	0.6122	0.6146	0.5982
Pseudo R^2	0.0179	0.0218	0.0114	0.0111
Panel B. Probit regression of misstatement on PCAOB search				
Counthits_week	0.003***	0.003***	0.003**	0.003***
	(6.08)	(3.51)	(2.42)	(3.82)
Year Effects & December Year End	Yes	Yes	Yes	Yes
Area Under Curve	0.5866	0.5912	0.5667	0.6057
Pseudo R^2	0.0106	0.0137	0.0046	0.0163
Panel C. Probit regression of misstatement on estimated probability of misstatement and PCAOB search				
Pr(misstate)	0.527***	0.668***	0.605***	0.241**
	(7.76)	(6.47)	(2.91)	(2.54)
Counthits_week	0.002***	0.002**	0.002*	0.003***
	(3.56)	(2.04)	(1.70)	(3.46)
Year Effects & December Year End	Yes	Yes	Yes	Yes
Area Under Curve	0.6157	0.6145	0.6179	0.6291
Pseudo R^2	0.0195	0.0226	0.0135	0.0198

*, **, and *** indicate two-tailed statistical significance of 10%, 5% and 1% levels, respectively.

This table presents the probit regression results of whether an issuer's financial statement is misstated during 2012–2016 on (1) the estimated probability of misstatement; (2) PCAOB search; and (3) the estimated probability of misstatement and PCAOB search. For each year in the examination window (e.g., year 2012), we first run the misstatement prediction model using the expanded model variables over a four-year estimation window ending one year prior to the test year (e.g., 2007-2010. We exclude any restatements announced during or one year after the estimation window to ensure that the PCAOB had the same public data as in the estimation model. We then estimate the probability of misstatement *Pr(misstate)* in the test year by applying the estimated coefficients. For ease of interpretation, we report the average marginal effects of estimated coefficients and include z-statistics based on issuer-clustered standard errors in parentheses. To avoid confounding effects, we restrict our sample in this analysis to restated fiscal years with restatements announced after the end of PCAOB search year t. *Counthits_week* is the number of PCAOB search weeks in a year. All other variable definitions and data sources are detailed in Appendix.