

How objective is peer review?

Evidence from self-regulation of the accounting profession

Divya Anantharaman
Ph.D candidate

Columbia Business School
Uris Hall #6Q
3022 Broadway
New York NY 10027

First draft: July 1, 2007
This draft: Dec 4, 2007.

Abstract

I test whether peer reviews performed under the AICPA's Peer Review Program were objective indicators of audit quality. Using PCAOB inspection opinions of audit firms as a neutral, independent benchmark of audit quality, I find that firms that themselves review other firms consistently receive more favourable opinions from peer reviews than from PCAOB reviews. Firms with peer reviewers who are less likely to be independent (or more likely to be friendly), such as smaller, non-competing firms and firms likely to be 'peer review specialists', also receive significantly more favourable opinions from peer review than from the PCAOB. Also, while positive peer reviews from relatively independent reviewers are negatively related to the likelihood of Accounting and Auditing Enforcement Releases, positive peer reviews from more friendly reviewers are not. These findings suggest that not all peer reviews were objective or rigorous, and that some firms were able to control peer review outcomes by choosing friendly reviewers. However, it is unclear whether audit committees impounded the independence of peer review opinions when using these opinions in decisions to hire or fire auditors.

Keywords: Peer review, Self-regulation, Sarbanes-Oxley Act, Public Company Accounting Oversight Board

* I am indebted to Gilles Hillary and Clive Lennox for making the peer review data available. I am very grateful to Stephen Penman and Clive Lennox for invaluable support and guidance. Many thanks to Bob Colson for sharing his experience and insights. I would also like to thank Jiri Novak, Bjorn Jorgensen, Sudhakar Balachandran, Gil Sadka, participants in the PhD workshop at Columbia Business School, and an anonymous reviewer at the AAA 2008 Auditing Midyear Section Conference for useful comments.

I. INTRODUCTION

The Sarbanes-Oxley Act of 2002 created the Public Company Accounting Oversight Board (PCAOB). The PCAOB is an independent, non-profit organization with the mission of overseeing auditors of public companies in order to protect the interests of investors. The PCAOB's duties include registration of accounting firms that perform audits of public companies trading in U.S. securities markets and periodic inspections of registered accounting firms.

The establishment of the PCAOB and the introduction of statutory inspections of audit firms marks a significant change in the regulatory framework of the accounting profession in the United States. The accounting profession in the United States has historically been self-regulated, with the AICPA's Peer Review Program forming the cornerstone of the self-regulatory framework. Under the peer review system, audit firms would typically find another firm to perform a review of their quality control systems and performance of audit engagements. The reviewer would issue a peer review report expressing an opinion (clean, unmodified, modified or adverse) on the reviewee's performance of its audit engagements and on its quality control systems.

The peer review system has long been criticized for lacking independence and rigour. Very few peer reviews concluded with a 'modified' or 'adverse' opinion. Former SEC Chairman Williams stated in Congressional testimony in 2002 that

“peer review in its present form has become too incestuous. A system needs to be established which is independent of the accounting profession” (Williams, 2002).

The Public Oversight Board in a white paper on regulatory reform of the audit profession, stated:

“Monitoring of firms’ accounting and auditing practices by the peer review system has come to be viewed as ineffective, as a diagnostic or remedial tool. The process has lost credibility because it has come to be perceived as *clubby* and not sufficiently rigorous” (POB, 2002).

In a similar vein, Mason (2005) gives the following anecdote:

“For years, the Big Eight, Big Six and then the Big Five peer reviewed each other, exchanging peer review fees reputed to be in the neighborhood of \$500,000. Despite scandals at Lincoln Savings and Loan (audited by Arthur Young), Baptist Foundation of Arizona (Andersen), CUC (Ernst & Young) and Adelphia Communications (Deloitte & Touche), the peer review reports always concluded with a commendation for an "unmodified" or clean opinion.”

Concerns about the objectivity of peer reviews have revolved around the independence of reviewers. In the AICPA’s Peer Review Program, firms have the ability to choose their own reviewers. This leaves open the possibility for firms to select friendly reviewers (Grumet, 2005). If firms are permitted to pick their reviewers, they could, as suggested by Matusiak (1979), exercise control over the resulting reports, creating a “back-scratching” result.

Despite these concerns about the independence of peer reviews, empirical evidence on these criticisms has been scarce. The independence, or objectivity, of a peer review is difficult to assess as there exists no normative benchmark on the actual number of problems existing at a firm or the ‘right’ number of critical findings to be expected from a peer review (Wallace, 1991).

The objective of this study is to test whether peer reviews under the erstwhile system of self-regulation were *objective indicators of audit quality*. I assess the objectivity of peer reviews using the shift from self-regulation to statutory regulation to construct a natural experiment. I obtain the sample of audit firms that have *both* a peer review on record in the last AICPA public file *and* a PCAOB inspection report available from the first round of PCAOB inspections that

have occurred till date. For each of these firms, I compare the outcome from the peer review process with the outcome from the PCAOB review process . Assuming that the PCAOB reviews all firms with the same level of rigour and that audit firms are unable to exert influence over the PCAOB's inspection reports, I then use the difference between each firm's last available peer review result and its immediately subsequent PCAOB result to capture any bias in the peer review outcome. In this sense, I use an audit firm's PCAOB rating as a neutral, independent benchmark measure of its quality. I then construct proxies for the extent of influence of the firm being reviewed and for the 'friendliness' of the reviewer, and test whether the measure of bias in peer reviews varies systematically with the proxies for the independence of the review.

Among the sample of firms that have both a peer review and a PCAOB review, I find that several reviewee and reviewer characteristics¹ systematically predict the bias in peer reviews. I find that firms that themselves perform reviews on other firms tend to receive peer reviews that are significantly more favourable than PCAOB reviews. I also find that firms that are reviewed by another firm of their own choosing (as opposed to an AICPA-appointed review team), and firms reviewed by non-competitors (who are more likely to be friendly) receive peer reviews that are favourable compared to PCAOB reviews. I also find that firms reviewed by smaller firms and by firms that rely more on peer reviews for their revenues, tend to receive peer reviews that are favourable compared to PCAOB reviews. I also combine these various reviewer characteristics² into a single summary measure that is expected to capture how 'friendly' the reviewer is likely to be. In summary, I find that firms reviewed by reviewers who might be expected to be 'friendly'

¹ Throughout this paper, I use 'reviewer' to mean the firm performing the peer review, and 'reviewee' to mean the firm that is undergoing the peer review.

² (i) Is the reviewer another firm or an AICPA-appointed team; (ii) Is a reviewer a non-competitor; (iii) Is the reviewer a smaller firm, and (iv) Is the reviewer a 'peer review specialist'

tend to consistently receive peer reviews that are more favourable than the PCAOB reviews received by these same firms.

I verify these results using Accounting and Auditing Enforcement Releases (AAERs) as an alternative proxy for audit quality. If Accounting and Auditing Enforcement Releases issued against an audit firm or its issuer clients are potential indicators of audit failure, then if peer reviews are objective, audit firms that receive clean peer reviews should be less likely to have AAERs than firms that do not receive clean peer reviews³. I find this to be true *only* in the sub-sample of firms that have been reviewed by relatively rigorous reviewers. Within the sub-sample of firms with relatively friendly reviewers, firms that clean peer reviews are *not* significantly less likely to have AAERs than firms that do not have clean peer reviews. This supports the conclusion that certain firms obtained reviews that were not independent or rigorous.

I interpret this evidence as supporting critics of peer review who alleged that the lack of independence of peer reviews distorted peer review results for several firms. This evidence is relevant for several reasons. First of all, it helps to understand one of the reasons for the apparent failure of the self-regulatory system for public accounting firms in the United States. Auditors perform an integral ‘gatekeeper’ role in our corporate governance system, and were an important component in the accounting and governance failures that led to Sarbanes-Oxley [Coffee (2002), Edwards (2004)]. The question of ‘*Who should audit the auditors?*’ becomes important in light of these events. Secondly, self-regulatory peer review still exists in various forms in accounting, as pointed out by Hillary and Lennox (2005). Audit firms in the U.S. with only privately-held clients continue to be self-regulated by peer review. Accounting professions in various countries

³ Hillary and Lennox (2005) confirm this to be true for the entire sample of firms with peer reviews.

are still self-regulated to different degrees. This evidence should be useful in evaluating the effectiveness of peer review in these settings.

II. INSTITUTIONAL BACKGROUND

A. The old self-regulatory framework

The self-regulatory structure of the accounting profession that existed prior to the Sarbanes-Oxley Act revolved around the AICPA's Division for CPA firms and the Public Oversight Board. A timeline of self-regulation and its collapse is provided in Figure 1.

AICPA Division for CPA Firms

The impetus for creation of the formal self-regulatory framework for the accounting profession arose from several highly public audit failures in the 1970s. The cases involved fraudulent financial reporting and illegal acts such as bribes, political payoffs and kickbacks. Hearings at the United States Senate and the House of Representatives ensued, following which the AICPA created its Division for CPA firms, consisting in turn, of the SEC Practice Section (for firms auditing SEC registrants) and the Private Companies Practice Section (for firms auditing privately-held clients).

The Division for CPA firms was intended to serve as the primary vehicle for self-regulation. The objective of the SEC Practice Section (SECPS) was to improve the quality of firms auditing clients that filed statements with the SEC. The distinguishing requirement of membership in the

SECPS was submission to a mandatory peer review once every three years. All AICPA members who audited publicly-held companies were required to belong to the SECPS.

The Public Oversight Board (POB)

The Public Oversight Board was an independent body created to oversee the peer review, quality control and other activities of the SECPS. The POB was funded by SECPS member firm dues. It represented the public, meeting with various elements of the self-regulatory program on a regular basis to remind them of the public interest and the public viewpoint.

The Peer Review Program

The peer review program for firms that audited SEC registrants was administered by the AICPA's SECPS, with oversight by the POB and the SEC. The primary purpose of peer review was to provide assurance to the public that firms that audited publicly-held clients had effective quality control systems. Peer reviews aimed at testing these quality control systems by testing a sample of the reviewee's accounting and auditing engagements⁴.

A description of the peer review experience of a practice of about 25 professionals is given by Sepp (1983). The reviews were primarily systems-oriented, i.e. the reviewee's quality control systems, documents, manuals and checklists were the focus of the review⁵. The reviewers first acquainted themselves with the firm's quality control policies and procedures. The reviewers

⁴ The scope of the peer review program was based on the nine elements of quality control laid down in the Statement of Auditing Standards No 25 (AICPA, 1980) and the Statement of Quality Control Standards No 1 (AICPA, 1979) : independence, assigning personnel to engagements, consultation, supervision, hiring, professional development, advancement, acceptance and continuance of clients, and inspection.

⁵ Reviews could also be engagement-oriented, in which case they focused on the application of accounting standards to specific engagements.

then selected a sample of engagements for which the workpapers, correspondence files and other documentation were walked through and checked. On the final day, an exit conference was held where the reviewers explained to the reviewee the scope of their review and any recommendations for improvement.

The results of the peer review were summarized in a report which was in boilerplate format, accompanied by a letter of comments which could take on any form. The reviewers issued a single opinion on the reviewee: clean, unmodified, modified or adverse. Typically, a peer review would result in a modified or adverse opinion only if ‘significant’ deficiencies were identified in the firm’s quality control systems. Modified or adverse opinions were rare.

B. Lack of independence of peer reviews

The self-regulatory system of peer review has come under criticism for a variety of reasons. Even though the POB was set up as an independent body, it was funded by dues paid by member firms of the SEC Practice Section, raising concerns about its ability to be independent. The scope of the peer reviews has also been criticised. Peer reviews adopted a future-looking, systems-oriented approach to audit quality. This led to what some commentators called a ‘paper chase’ (Fogarty, 1996) where reviews tended to be more about whether the firm had adequate policies, manuals and checklists in place.

The strongest criticisms of peer review, however, relate to the lack of independence inherent in a largely voluntary self-regulatory system where firms are able to choose their own reviewers. The Public Oversight Board summarized this concern as follows:

“Monitoring of firms’ accounting and auditing practices by the peer review system has come to be viewed as ineffective, as a diagnostic or remedial tool. The process has lost credibility because it has come to be perceived as ‘clubby’ and not sufficiently rigorous.”

Former SEC Chairman Williams in his Congressional testimony in February 2002 stated that

“peer review in its present form has become too incestuous. A system needs to be established which is independent of the accounting profession.”

Under the AICPA’s Peer Review Program, firms could choose their own reviewers. This left open the possibility that firms could pick ‘friendly’ reviewers (Grumet, 2005; Fogarty, 1996). If firms are permitted to choose their own reviewers, it also leaves open the possibility that they can influence the content of the resulting reports, i.e., ‘back-scratching’ (Matusiak, 1979). For example, the then-SEC Chief Accountant Lynn Turner, in a speech on the challenges of self-regulation, noted “...as an engagement partner, I was always *told in advance* which of my audits were going to be peer-reviewed. I think it would be beneficial if the staff of the Public Oversight Board could review some audits on an unannounced basis” (Turner, 2000). Similarly, Russell and Armitage (2006) in a survey of 200 accounting firms, find that about 18 % of respondents acknowledged that the firm’s peer reviewer had allowed the firm to self-select the engagements to be reviewed.

C. The new system of statutory regulation

In the wake of the Enron collapse and audit failures that came to light in 2002, SEC Chairman Harvey Pitt proposed a new regulatory structure for audit firm oversight. A month before Arthur Andersen was convicted for obstruction of justice, in May 2002, the Public Oversight Board voted to terminate its own existence.

The Public Company Accounting Oversight Board (PCAOB)

The Sarbanes-Oxley Act of 2002 established the Public Company Accounting Oversight Board (PCAOB) to oversee the audit of public companies and protect the interests of investors. The PCAOB is overseen by the SEC, and funded by fees paid by public companies determined as a function of the companies' individual market capitalization. Its duties are to register public accounting firms that prepare public company audits, establish auditing, quality control, ethics, independence and other standards related to the preparation of public company audits, conduct inspections of registered public accounting firms and to conduct investigations and disciplinary proceedings when necessary. The establishment of the PCAOB, with its responsibility to inspect public accounting firms effectively marks the end of self-regulation as the primary mechanism of regulation in the audit profession.

The PCAOB conducts *annual* inspections of audit firms with more than 100 public company clients, and inspections *once every three years* for firms with less than 100 public clients. Similar to the peer review process, these inspections involve reviewing a sample of the firm's audit engagements and assessing its quality control systems⁶. The PCAOB makes its inspection reports publicly available, with the exception that any criticisms of the firm's quality control systems are not made public if the firm has taken appropriate actions to respond to the criticisms within one year.

The PCAOB inspection reports include the name and basic details of the firm being reviewed, the number of engagements reviewed, and a list describing the audit performance deficiencies identified in the engagements reviewed. The report also includes the PCAOB's assessment of

⁶ The PCAOB examines quality control systems along the same aspects as the Peer Review Program: practices, policies and procedures concerning audit performance, training, compliance with independence standards, client acceptance and retention, and the establishment of policies and procedures.

defects identified in the firm's quality control systems, but only if the firm has not responded adequately to these criticisms within one year of inspection. However, if the inspection does not reveal any defects in the firm's quality control systems, this fact is disclosed in the inspection report. The inspection report also includes the firm's written response to the inspection report, if one exists. Unlike the previous peer review system, the PCAOB does not issue a summary opinion (clean, unmodified, modified or adverse) on each firm inspected.

III. LITERATURE REVIEW

Wallace (1991) analyzes peer review reports filed from 1980 to 1986, and investigates whether the number of deficiencies identified in firms' reviews varies systematically with characteristics of the reviewee and the reviewer, such as size and membership in a Big 8 or Big 15 category. She finds no relationship between these variables and the number of deficiencies identified in the firms' peer review. She concludes that the voluntary nature of the peer review, and firms' ability to select their own reviewers, have not significantly affected the objectivity and veracity of peer review reports. In a more recent study, Russell and Armitage (2006) perform a survey of the partners of 200 accounting firms, in which they ask questions to identify whether firms are exploiting potential loopholes in the peer review process to obtain more favourable opinions than they would otherwise merit. 17.5 % of the respondents reported that their peer reviewer had allowed them to self-select the engagements to be reviewed, of which a further 43 % acknowledged that they had then selected those engagements least likely to receive negative comments.

Hillary and Lennox (2005) provide the first large-scale evidence on the credibility of AICPA peer reviews as a signal of audit quality. They find that audit firms that received clean peer review opinions gained clients in the following year and audit firms that received modified or adverse opinions lost clients. This suggests that peer review opinions, for all their alleged weaknesses, were still a credible signal of quality for audit committees (or other decision-makers).

Casterella, Jensen and Knechel (2006) extend the work of Hillary and Lennox by testing if peer review opinions are associated with *actual* audit quality. Their findings support Hillary and Lennox - they find that peer review opinions are associated with indicators of actual quality control problems, such as overworking staff, selling tax shelters and taking on risky clients, and with malpractice claims alleging auditor negligence.

Hermanson, Houston and Rice (2007) analyze PCAOB inspection reports of small firms, and find that firms with deficiencies tend to be smaller firms that are growing rapidly and perhaps over-extending into the audit client market. Gunny and Zhang (2006) compare self-regulation (peer reviews) to statutory regulation (PCAOB inspections). In contrast to Hillary and Lennox (2005) and Casterella, Jensen and Knechel (2006), Gunny and Zhang find that peer review opinions are not associated with earnings quality of the audit firms' clients. However, they find that PCAOB opinions are associated with earnings quality and earnings informativeness of the audit firms' clients, i.e. audit firms that receive favourable PCAOB opinions have clients with higher earnings quality.

Hillary and Lennox (2005) analyze separately the determinants of peer review inspection outcomes; Gunny and Zhang (2006) and Hermanson, Houston and Rice (2007) analyze the determinants of PCAOB inspection outcomes. Hillary and Lennox (2005) find, for example, that larger firms tend to receive more favourable peer review opinions and that firms which perform reviews on other firms tend to receive more favourable peer review opinions. However, the interpretation of these results is unclear because intrinsic audit quality is a correlated omitted variable. For example, larger firms could be receiving better peer review opinions because they do provide better audit quality. I attempt to address this endogeneity by exploiting a quasi-experimental setting.

I use the change from self-regulation to statutory regulation as a natural experiment to test the existence of any systematic bias in peer review results. By comparing inspection outcomes from the peer review process with inspection outcomes under subsequent PCAOB process *for the same set of firms*, I approximate an experimental setting where the firm being reviewed is held constant, and the reviewer is changed from one setting to the other. I assume that the PCAOB, being a third-party inspector, is (i) neutral and independent of all audit firms, and (ii) reviews all firms with the same rigour. I therefore attempt to extract that portion of the peer review opinion that is attributable to ‘measurer bias’ by benchmarking the peer review outcome against the more neutral PCAOB outcome.

This study extends the emerging literature on the merits and demerits of self-regulation and statutory regulation in the audit profession. Hillary and Lennox (2005) find that peer reviews, in spite of the criticisms levelled against them, were perceived by audit clients to be credible

indicators of audit quality. The implication of these findings for the efficacy of peer review is open to interpretation. One interpretation is that peer reviews were truly objective indicators of audit quality, notwithstanding all the criticisms levelled against them, and were therefore perceived (rightly) as useful information. An alternative, and perhaps less benign, interpretation would be that criticisms of peer review are true – that peer reviews do not always provide an objective assessment of audit quality, but that audit committees (or other decision-makers in the audit client market) did not fully impound these issues while making decisions about hiring and firing auditors using peer reviews opinions. Hillary and Lennox (2005) are careful to point out in their study that “*although we find evidence that self-regulation was credible, our paper does not provide unequivocal support for the peer review program.*” I attempt to extend the interpretation of the results of Hillary and Lennox (2005), by disentangling these alternative interpretations, by providing support for the latter view.

IV. RESEARCH DESIGN

The experimental setting

I use the recent shift from self-regulation (peer reviews) to statutory regulation (PCAOB reviews) to construct a natural experiment to test for any systematic patterns of bias in peer review findings. I obtain the sample of audit firms that have *both* a peer review on record in the last AICPA SECPS public file *and* a PCAOB inspection report available from the first round of PCAOB inspections till date. For each of these firms, I compare the outcome from the peer review process (captured in PEER_REVIEW_OPINION) to the outcome from the PCAOB review process (PCAOB_OPINION). I assume that the PCAOB is an impartial, independent inspector that reviews all firms with the same rigour. I then attribute the difference between each

firm's peer review outcome and its immediately subsequent PCAOB outcome (DIFF) to lack of independence inherent in the peer review process, which I label the 'bias' in peer review. In this sense, I use a firm's PCAOB rating as a neutral, objective benchmark of its audit quality. Such a benchmark is essential to be able to comment on whether peer review ratings for a particular firm were "too high" or "too low".

In order to demonstrate the presence of any bias in peer review findings, this DIFF measure must be shown to correspond in some predictable way to characteristics of the reviewer and reviewee that could affect the independence of the review. I construct proxies that are expected capture this 'friendliness' of a peer review, drawing upon prior research and anecdotal commentaries. I then test whether these proxies have any significant ability to predict the direction of the DIFF measure.

Measuring peer review opinions

Under the AICPA's Peer Review Program, the reviewer would issue a report containing a review opinion: clean, unmodified, modified or adverse. A separate letter of comments was also typical, listing out significant weaknesses identified in the reviewee's quality control systems and in its performance of audit engagements reviewed. Hillary and Lennox (2005) develop a numerical scale on which to code these peer review opinions that incorporates both the nature of the opinion as well as the number of weaknesses identified. I use this scale to measure peer review opinions in this study. 'Clean' opinions are coded zero; 'Unmodified opinions' are coded from one to nine, depending on the number of weaknesses listed; 'Modified' opinions are coded from 10-19, depending on the number of weaknesses listed, and 'Adverse' opinions are coded from

20-29, depending on the number of weaknesses listed⁷. Lower numerical values, therefore, correspond to more favourable opinions. The peer review opinion is represented by the variable PEER_REVIEW_OPINION.

Measuring PCAOB opinions

Similar to the peer review program, the PCAOB issues an inspection report after its review of each firm⁸. The inspection report lists out the number and types of deficiencies identified in the reviewee's performance of audit engagements reviewed. The inspection report also states whether any defects were identified in the firm's quality control systems. The nature and extent of the quality control defects (if any) are, however, not disclosed in the report.

The deficiencies identified in the reports are of the following kinds:

- (a) Audit performance deficiencies, which could be:
 - (i) Potentially material misstatements in the issuer's financial statements, relating to ***departures from GAAP*** that were not identified or addressed appropriately by the audit firm being reviewed, or
 - (ii) ***Other deficiencies*** that relate to the firm's failure to obtain sufficient competent evidential matter to support its opinion on the financial statements⁹
- (b) Defects in the firm's quality control systems.

⁷ Hillary and Lennox's numerical scale, therefore, takes into account both the summary opinion and the number of weaknesses listed in the peer review report and letter of comments.

⁸ The inspection reports are made available on http://www.pcaobus.org/Inspections/Public_Reports/index.aspx

⁹ A few examples of the kinds of other deficiencies identified in the reports are: failure to perform and document substantive procedures to test revenues, failure to perform and document an evaluation of the assumptions used by management in performing goodwill impairment analysis, inappropriate reliance on the use of the work of a specialist, and inappropriate reliance on the issuer's internal controls without testing the effectiveness of those controls; among others. For a full deficiencies analysis, see AICPA (2006a) and AICPA (2006b).

The PCAOB inspection reports were collected from the PCAOB's website, analyzed and hand-coded into a summary measure of the inspection result. For each inspection report issued, I count the number of potential departures from GAAP (#DEPARTURES_GAAP), the number of other deficiencies identified (#DEFICIENCIES), and whether any quality control defects were identified (DUMMY_QC_DEFECTS). I combine these measures into a summary measure of the PCAOB opinion, defined as follows:

$$\text{PCAOB_OPINION} = \text{\#DEPARTURES_GAAP} + \text{DUMMY_DEFICIENCIES} + \text{DUMMY_QC_DEFECTS}$$

(where DUMMY_DEFICIENCIES is a dummy variable set to zero if no other audit performance deficiencies were identified, and one if some deficiencies are identified¹⁰). Similar to the measure for peer review opinions, lower numerical values correspond to more favourable opinions.

Measuring peer review 'bias'

I compare the summary measures PEER_REVIEW_OPINION and PCAOB_OPINION to assess the difference between peer review outcomes and PCAOB outcomes for each firm. Both of these measures are constructed so that zero corresponds to a 'clean' opinion (the most favourable), lower values correspond to more favourable opinions and higher values correspond to more adverse opinions. However, in order to make these measures directly comparable, I replace the

¹⁰ I use a dichotomous variable to capture audit performance deficiencies so as to minimize reliance on the actual number of weaknesses listed out in the PCAOB inspection report. I do this in order to mitigate any distortions in comparability that may arise due to differences only in the way PCAOB inspection reports are written, as compared to peer review reports. For example, it could be said that the PCAOB inspection reports mechanically list out all the audit performance weaknesses identified, whereas peer review reports focus more on 'big picture' issues and on evaluating the firm as a whole.

numerical values of these measures with their ranks, for the matched sample of firms with both peer review and PCAOB opinions available. I use only these ranks in the rest of the study¹¹. The ranks are generated by sorting the firms in descending order of the PEER_REVIEW_OPINION and PCAOB_OPINION variables. Therefore, a *higher numerical rank* corresponds to *better performance* in a given review. (For example, a firm ranked 355 has received a better peer review than a firm ranked 100).

I then compute the difference as:

$$\text{DIFF} = \text{Rank}(\text{PEER_REVIEW_OPINION}) - \text{Rank}(\text{PCAOB_OPINION})$$

Positive values of DIFF, therefore, indicate that the firm received a peer review opinion that was favourable relative to its PCAOB opinion; negative values indicate the opposite.

I also develop an alternative summary measure to compare performance across the two different reviews. DIFF_CLEAN is:

Set to 1 if the firm has received a ‘clean’ opinion from peer review but not from the PCAOB;

Set to 0 if the firm either receives a clean opinion from *both* reviews or does **not** receive a clean opinion from *both* reviews; Set to -1 if the firm has not received a clean opinion from peer review but has received a clean opinion (i.e. no weaknesses identified) from the PCAOB¹².

¹¹ I use ordinal measures, i.e. ranks to capture performance in reviews so as to avoid any cardinal interpretations of the peer review ratings and PCAOB ratings variables. These simple summary measures are only intended to indicate that a firm with a total of , say, 10 weaknesses in the PCAOB inspection has performed *worse* than a firm with no weaknesses. I make this research design choice to also mitigate the impact of any potentially influential observations, for e.g. the very large national accounting firms have a very large number of weaknesses listed out in their PCAOB inspection reports; to a large extent this is attributable to the fact that these firms have so many more audit clients for the PCAOB to inspect, when compared to a smaller firm with only one or two clients.

¹² This measure focuses purely on whether the firm receives a clean opinion or not; it abstracts away almost entirely from the actual number of weaknesses identified (if any) in either review, and should be unaffected by any mechanical differences in the way PCAOB writes its reports, compared to peer reviewers.

Characteristics of the reviewee and reviewer

I hypothesize that the ability of firms to influence peer review outcomes will vary with the characteristics of the reviewee and the reviewer chosen. I use prior literature and anecdotal commentaries to motivate the following explanatory variables.

Characteristics of the reviewee

I hypothesize that larger firms would be able to exert more influence over the peer review process. Wallace (1991) hypothesizes that the great stakes in self-regulation held by large accounting firms would lead to fewer deficiencies being discovered at these firms. Austin and Langston (1981) in a survey of approximately 130 CPAs, report that about one-third of the respondents saw self regulation as ‘a means by which the profession would be captured by dominant CPA firms and used to project the views and philosophies of these firms.’

Hillary and Lennox (2005) find that larger firms are more likely to receive favourable peer review opinions. They, however, interpret this evidence as being consistent with the notion of larger firms providing higher quality audits. I measure size as the natural log of number of clients of the firm (SIZE_REVIEWED)¹³.

Critics of peer review have characterized it as being too ‘clubby’ (POB, 2002) and ‘incestuous’ (former SEC Chairman Williams in Congress testimony). This suggests that firms are more likely to be able to find sympathetic reviewers if they themselves are involved in performing

¹³ I log-transform the number of clients as the underlying variable is highly skewed to the right due to the large national accounting firms. I replicate all the tests using the absolute number of clients, and the inferences stay the same. I also try alternative proxies for size, such as the number of personnel in the firm. The results remain unchanged.

peer reviews for other firms. Hillary and Lennox (2005) find that firms that perform reviews on other firms are likely to receive favourable peer reviews. I extend this by testing whether firms that perform reviews on other firms are more likely to be able to influence peer review outcomes in their favour, obtaining peer reviews that are favourable compared to PCAOB reviews. I define a dummy variable REVIEWED_FIRM_REVIEWED as one if the reviewed firm has performed at least one peer review on another firm in the peer review sample, and zero if otherwise.

Characteristics of the reviewer

Under the AICPA peer review program, the firm wishing to undergo a peer review can choose from three types of reviewers: (i) An AICPA-appointed review team (ii) A review team from another national or international association of CPA firms¹⁴, or (iii) another firm, which the reviewee can choose and negotiate with individually. If the firm chooses to be reviewed by an AICPA-appointed review team, the AICPA will assemble a team of reviewers, typically from firms that match the reviewee's size and areas of specialization.

It is intuitive that the 'back-scratching' result will be most pervasive when the firm negotiates directly with and chooses another firm to perform its peer review (Wallace, 1991; Grumet, 2005), as opposed to when the reviewee asks the AICPA or another CPA association to pick its reviewers. Hillary and Lennox (2005) also find that AICPA review teams are more likely to award unfavourable opinions. Therefore, I hypothesize that reviewees choosing firm-on-firm reviews have greater ability to influence peer review outcomes in their favour and will, as a result, have peer review outcomes that are favourable relative to PCAOB outcomes. I use a

¹⁴ In the peer review sample, reviews were performed by the following associations: (i) Accountants Global Network (ii) CPA Associates International (iii) CPAmerica (iv) DFK International (v) National Associated CPA firms

dummy variable `FIRM_ON_FIRM` that is set to one if the review is performed by another firm, and zero if the review is performed by the AICPA or another CPA association.

Grumet (2005) and other critics of the peer review process allege that allowing reviewees to pick and negotiate with their reviewers leaves open the possibility for firms to pick ‘friendly’ reviewers. A reviewer is more likely to be ‘friendly’, i.e. award a favourable opinion, if the reviewer and reviewee are not in competition. Hillary and Lennox (2005) find that firms reviewed by competitors tend to receive more unfavourable opinions than firms reviewed by non-competing firms. I extend this hypothesis by testing whether firms reviewed by non-competitors receive peer review opinions that are more favourable relative to PCAOB opinions. Following Hillary and Lennox (2005), I use a dummy variable `NON_COMPETITOR` set to one if the reviewing firm is not a competitor of the reviewed firm, and zero otherwise. The reviewer is considered to be non-competing if it is located more than 50 miles away from the reviewed firm¹⁵.

I expect that the more important the review business is to the reviewer, the greater is the reviewee’s ability to influence the outcome of the review. I expect firms that choose reviewers that are small relative to themselves will be better able to obtain peer review outcomes that are favourable. I use the natural log of the number of clients of the reviewer (`SIZE_REVIEWER`) to proxy for the size of the reviewer.

¹⁵ Consistent with Hillary and Lennox (2005), I use various cutoffs: 50 miles, 75 miles, 100 miles. The inferences stay robust to these definitions.

I include another variable that is designed to capture how important the peer review business is to the reviewer. Articles in the press have noted the rise of a peer review ‘cottage industry’ within the accounting profession, where some firms have specialized in the business of providing peer reviews (Rankin, 1978). These firms, typically, have more peer review clients than audit clients. I test whether the relatively greater need of peer review specialists to generate and maintain peer review clients results in them awarding reviews that are favourable when compared to PCAOB reviews. I define LN_AUDITS_TO_REVIEWS as the natural logarithm of the number of audit clients of the reviewing firm to the number of peer review clients¹⁶. Lower values of this variable indicate firms that perform a ‘large’ number of peer reviews compared to the number of issuers they audit.

I also combine the various reviewer characteristics in an additive fashion into a single summary measure FRIENDLY_SCORE, that I expect will capture the extent of independence (or lack thereof) in the peer review. FRIENDLY_SCORE is calculated as:

$$\text{FRIENDLY_SCORE} = \text{FIRM_ON_FIRM} + \text{NON_COMPETITOR} + (1 \text{ if } \text{SIZE_REVIEWER} < \text{SIZE_REVIEWED}; 0 \text{ otherwise}) + (1 \text{ if } \text{AUDITS_TO_REVIEWS_RATIO} < 1; 0 \text{ otherwise})$$

Therefore, reviews that are performed by the AICPA or other CPA associations get a FRIENDLY_SCORE of 0, reflecting that they are least likely to be ‘friendly’. At the other end of the spectrum, reviews performed by another firm which is not a competitor, which is smaller than the firm being reviewed, and which is a ‘peer review specialist’, will get a FRIENDLY_SCORE of 4, reflecting that these reviews are most likely to be ‘friendly’.

¹⁶ The prediction on this variable is, however, not unambiguously clear. It could be argued that such peer review specialists can sustain this line of business only if they develop a reputation for providing objective and rigorous reviews. This, therefore, becomes an empirical issue.

Regression specification and predictions

My main analysis consists of the following relation:

$\text{DIFF (or DIFF_CLEAN)} = f(\text{Characteristics of the reviewee, Characteristics of the reviewer})$

I test these relationships using the following specifications:

$\text{DIFF (or DIFF_CLEAN)} = \beta_0 + \beta_1 * \text{LN_SIZE_REVIEWED} + \beta_2 * \text{REVIEWED_FIRM_REVIEWS} + \beta_3 * \text{FIRM_ON_FIRM} + \beta_4 * \text{NON_COMPETITOR} + \beta_5 * \text{LN_SIZE_REVIEWER} + \beta_6 * \text{LN_AUDITS_TO_REVIEWS} + \text{Controls} + \varepsilon$

I expect that $\beta_1, \beta_2, \beta_3$ and $\beta_4 > 0$. I expect β_5 and $\beta_6 < 0$.

Alternatively,

$\text{DIFF (or DIFF_CLEAN)} = \alpha_0 + \alpha_1 * \text{LN_SIZE_REVIEWED} + \alpha_2 * \text{REVIEWED_FIRM_REVIEWS} + \alpha_3 * \text{FRIENDLY_SCORE} + \text{Controls} + \varepsilon$

I expect α_1, α_2 and $\alpha_3 > 0$.

Control variables

I include dummy variables **BIG_FIRM** (set to 1 if the firm is a large national accounting firm with more than 50 clients at the time of peer review, and 0 otherwise) and **TINY_FIRM** (set to 1 if the firm has no clients or only 1 client at the time of peer review, and 0 otherwise) to account for any differences in effects across very large firms and very small firms.

V. EMPIRICAL RESULTS

Data and sample selection

The PCAOB inspection reports were obtained from the PCAOB's website, analyzed and coded into variables by hand¹⁷. The PCAOB inspections for which public reports have been made available span a time period from May 2004 to January 2007 (the PCAOB makes reports publicly available only after a certain time lag). The peer review data and data on reviewee and reviewer characteristics were obtained from Clive Lennox's website and are the data used in Hillary and Lennox (2005)¹⁸. The peer reviews span a time period from June 1997 to September 2003.

Descriptive statistics

Table I, Panel A describes the peer review opinions received by all firms with peer review data in the Hillary and Lennox file, as well as only for those firms that are in both the peer review file *and* the PCAOB file (this intersection is the sample of primary interest for this study). Consistent with criticism of the peer review process, a very small fraction of all firms, less than 5 % of the total, receive modified or adverse opinions. The remaining 95 % of firms receive either a clean or an unmodified opinion¹⁹. Panel B describes the distribution of peer review scores. The sample of firms with both peer review data and PCAOB data (n = 355) appears to be very similar to the full peer review file (n = 1,001) in terms of the distribution of peer review opinions.

¹⁷ The inspection reports are available at: http://www.pcaobus.org/Inspections/Public_Reports/index.aspx, last accessed on May 25, 2007.

¹⁸ I am indebted to Gilles Hillary and Clive Lennox for making this data available.

¹⁹ An un-modified opinion is one where the peer review report contains some explanatory language about weaknesses identified during the review, but those weaknesses are not considered significant enough to merit a modified opinion.

Table II, Panel A describes the distribution of PCAOB inspection opinions for all firms with a PCAOB inspection report available, and for only those firms with both PCAOB opinions as well as peer review opinions. Approximately one-third of all firms have received ‘clean’ reports from the PCAOB. Panel B of Table II describes the frequency of different types of critical findings from the PCAOB inspections. The PCAOB found potentially material mis-statements pertaining to departures from GAAP in about 11 % of all firms inspected. It found some quality control deficiencies in two-thirds of all firms inspected²⁰, and identified other audit performance deficiencies in more than half of all firms inspected.

Table III shows summary statistics of the different reviewee and reviewer characteristics that are expected to be related to the reviewee’s ability to influence the outcome of a peer review. The distribution of size of the reviewee and reviewer, measured in terms of the number of clients, is highly skewed to the right due to the concentrated nature of the accounting profession. I replace these variables with their natural logarithm for use in the regression tests. The descriptive statistics for REVIEWED_FIRM_REVIEWS show that only 17.7 % of all reviewees have themselves performed a peer review, whereas the average firm that performs peer reviews performs many such reviews. The mean and median for NUMBER_OF_REVIEWS_DONE (which measures the total number of reviews performed by a given reviewer in the AICPA public file) are 3 and 5.4 respectively. This implies that there are much fewer number of unique reviewers in the sample than reviewees²¹.

²⁰ The PCAOB does not reveal publicly the exact nature and description of the quality control deficiencies identified. However, when an inspection does not reveal any such deficiencies, the PCAOB discloses this fact. Therefore, this aspect of the inspection is represented by a dichotomous variable coded to zero if no quality control defects are identified, and one if otherwise.

²¹ In the regression tests, all standard errors are corrected for clustering by reviewer.

Table IV summarizes the DIFF variables, which are developed to capture the difference between a given firm's performance in peer review as compared to its performance in PCAOB review. Positive values of DIFF or DIFF_CLEAN correspond to *better performance* on peer review compared to PCAOB review. About 60% of firms receive clean opinions from both peer review and PCAOB, or from neither. About 25% of firms receive a clean opinion from peer review but not from the PCAOB.

Table V, Panel A shows Spearman rank correlations between the different variables measuring outcomes from peer reviews and PCAOB inspections. Peer review opinions are highly positively correlated with the PCAOB opinions. The different elements of the PCAOB opinions (#DEPARTURES_GAAP, DUMMY_QC_DEFECTS, #DEFICIENCIES) are all strongly positively correlated with one another.

Panel B of Table V shows univariate Spearman ranks correlations between the primary dependent variables (DIFF and DIFF_CLEAN) and reviewer and reviewee characteristics. Positive values of DIFF or DIFF_CLEAN imply that the firm has ostensibly performed on its peer review than on its PCAOB review. The correlations between the DIFF measures and the explanatory variables are all in the direction predicted; however not all are statistically significant. The significant relationships from this table are that firms that themselves review other firms (REVIEWED_DOES_REVIEWS = 1) and firms being reviewed by other firms (FIRM_ON_FIRM = 1) are likely to have positive DIFF values, indicating that these firms receive more favourable results from peer review than from PCAOB review. FRIENDLY_SCORE is strongly positively correlated with DIFF and DIFF_CLEAN, indicating

that firms being reviewed by other firms which are non-competitors, smaller in size and peer review specialists are more likely to receive peer reviews that are favourable compared to PCAOB reviews.

Regression results

Table VI estimates a regression model to explain the DIFF variables using characteristics of the reviewee and reviewer. The regressions using DIFF as the dependent variable are estimated using OLS, and regressions using DIFF_CLEAN are estimated using ordered probit. Most of the coefficients are in the predicted direction. Panel A of Table VI clearly shows that firms that themselves review other firms receive more favourable opinions from peer reviews than from PCAOB inspections, after controlling for size. The regressions also show that firms reviewed by other firms (as opposed to being reviewed by an AICPA-appointed team or another CPA association) receive more favourable opinions from peer reviews than from government inspections.

Panel B regresses the DIFF measures on characteristics of the reviewee *and* the reviewer (only for firm-on-firm reviews, as the reviewer characteristics are difficult to define for the other reviewers). The results indicate that firms opting to be reviewed by non-competing firms receive better opinions in peer reviews than in PCAOB reviews. The negative coefficients on LN_SIZE_REVIEWED and LN_AUDITS_TO_REVIEWS indicate that firms reviewed by smaller firms and by reviewers that are likely to be ‘peer review specialists’ (as indicated by lower values of LN_AUDITS_TO_REVIEWS) are likely to receive better opinions from peer reviews than from PCAOB reviews. The regressions using FRIENDLY_SCORE as an

explanatory variable corroborate these findings: the higher the value of 'FRIENDLY_SCORE', the more likely that the firm will receive a better opinion from peer review than from PCAOB review.

I interpret these results as showing consistent support for the 'back-scratching' story. Larger firms and firms that themselves review other firms are expected to be more influential, and to be able to exert more influence over the outcome of the peer review process. These firms are the ones that experience the greatest drop in measured performance when reviews start being done by the PCAOB. Similarly, the 'back-scratching' problem can be expected to be most intense when the firm is negotiating directly with another firm to conduct its peer review, rather than when it opts to be reviewed by an AICPA-appointed team. These firms, which opt for firm-on-firm reviews where the firm picks its reviewer, experience a larger drop in measured performance in PCAOB reviews.

Reviewers that are in direct competition with the reviewee could have a greater incentive to be rigorous in their reviews. Consistent with this notion, firms reviewed by non-competitors experience a greater drop in measured performance in PCAOB reviews. Smaller firms and firms that rely significantly on peer reviews for revenues could be more susceptible to pressure to award a favourable review. Consistently, firms reviewed by smaller firms and by 'peer review specialists' experience a larger drop in measured performance once the PCAOB starts doing reviews.

Limitations and further analysis

The use of PCAOB ratings as an objective benchmark of audit quality allows us to approximate an experimental setting. However, this approach is not free of limitations. I discuss below some of the important limitations and further analysis done to consider these issues.

Analysis excluding big firms

An important assumption made in this analysis is that the PCAOB reviews all firms with the same level of rigour, and that it does not have any incentives to award particularly favourable or unfavourable reviews to any firm. This assumption is key to the use of the DIFF measures as proxies for the ‘bias’ in peer review. It is reasonable to expect that the PCAOB, having been established by an act of law (the Sarbanes-Oxley Act), is neutral and independent of the accounting profession, and inspects all firms with the same level of rigour. However, there is one notable exception to this assumption, which is the PCAOB’s treatment of the large national accounting firms vis-à-vis smaller firms. Firms with more than 100 audit issuer clients are inspected by the PCAOB every year, whereas firms with less than 100 audit issuer clients are only inspected at random once every three years. This suggests, at the very least, some differential treatment or consideration by the PCAOB of these very large firms. I therefore repeat the regression analysis in Table VI after excluding large firms²² from the sample. This analysis is produced in Table VII. Excluding the largest firms from the sample has virtually no impact on the results.

²² I try several cutoffs to determine ‘big firms’: firms with more than 100 audit issuer clients, firms with more than 50 audit issuer clients. These cutoffs do not impact the reported results.

AAERs as an alternative benchmark for audit quality

'True' experimental situations are difficult to find; and this particular setting suffers from the obvious limitation that the peer review and the PCAOB review for a particular firm do not happen at the same point in time. The peer review files extend from 1997 to 2003; the PCAOB reviews commence from 2004 onwards. For each firm in the sample, therefore, there is a time lag between the peer review and the PCAOB review²³. In addition, the Sarbanes-Oxley Act was enacted towards the end of the peer review period, and this (along with the accounting and auditing failures that prompted the Act) could have changed the environment for audit firms in various, complex ways. Another potential problem with using PCAOB reviews is that PCAOB reviews are only available for a subset (n = 355) of all the firms in the peer review file (n = 1,001). If the firms that have been selected for review till date by the PCAOB have been non-randomly selected in some particular way²⁴, this could induce some selection bias into the results.

I try to analyze the impact, if any, of these issues, by verifying the main results of the paper with an alternative proxy for audit quality. I use a variable AAER set to one if the audit firm has had clients subject to Accounting and Auditing Enforcement Releases when the peer review happens, and set to zero otherwise²⁵, as an alternative benchmark of audit quality. Not only is this variable defined *at the same point in time* as the peer review, it is also defined for the entire sample of firms (n = 1,001) in the peer review file. If Accounting and Auditing Enforcement Releases are an indication of audit failure, then if peer reviews are objective, one would expect that firms with

²³ As another check, I sort firms by their delay between the peer review and the PCAOB review, and re-run the regression tests after excluding firms that fall (i) above the median (ii) above the 75th percentile, and (iii) above the 90th percentile of the distribution of delay between reviews. The results are unaffected.

²⁴ Hermanson, Houston and Rice (2007) provide evidence that the PCAOB targeted smaller, riskier, more rapidly growing firms for its 2004 inspections.

²⁵ This variable is as defined in Hillary and Lennox (2005), and forms part of their data.

clean peer reviews are significantly less likely to have AAERs, compared to firms that do not receive clean peer reviews. Hillary and Lennox (2005) show that this relationship holds true for the entire peer review sample, i.e. firms that receive clean peer reviews are significantly less likely to have AAER clients than firms that do not receive clean peer reviews.

However, as shown in Panel B of Table VIII, the ability of clean peer reviews to distinguish between AAER firms and non-AAER firms is quite monotonically *decreasing* in the ‘friendliness’ of the reviews themselves. In other words, among the sample of firms reviewed by relatively independent reviewers ($\text{FRIENDLY_SCORE} \leq 2$), firms that have clean peer reviews are significantly less likely to have AAERs than firms that do not have clean peer reviews. As the FRIENDLY_SCORE of the reviewer increases, clean peer reviews are unable to distinguish between AAER firms and non-AAER firms, suggesting that ‘clean’ peer reviews from these reviewers may not be as objective or rigorous. This corroborates the conclusions from Table VI and Table VII.

Perceived credibility of friendly reviews

The results of these analyses all suggest that not all peer reviews were performed with the same level of rigour or objectivity; that atleast some firms might have been able to control their peer review outcomes by choosing friendly reviewers. However, Hillary and Lennox (2005) show that peer reviews were perceived to be credible indicators of audit quality in the audit client market - firms gained audit clients after receiving a positive peer review, and lost audit clients after receiving a modified or adverse review. In Table IX, I mimic the analysis in Table VIII but now to test whether the perceived credibility of peer reviews varied with the nature of the review; i.e.

were clean reviews from relatively friendly reviewers perceived to be less credible than clean reviews from more independent reviewers? Panel B of Table IX shows that there exists no clear monotonic relationship between the independence of the reviewer and the perceived credibility of the review. This suggests that audit committees (or other decision-makers) may not have fully impounded the independence (or lack thereof) of peer reviews while using these reviews in their decisions to hire and fire auditors.

V. CONCLUSION

The accounting profession in the United States has historically been self-regulated through peer reviews. The Sarbanes-Oxley Act ended the primacy of self-regulation with statutory regulation, by introducing periodic inspections of audit firms by the Public Company Accounting Oversight Board. In this study, I test whether peer reviews under the AICPA's Peer Review Program were objective, unbiased indicators of audit quality.

A major criticism of peer reviews has been that allowing firms to choose their reviewers will result in conflicts of interest, and lack of independence, as firms might try to choose 'friendly' reviewers. I use the shift from peer reviews to PCAOB reviews to construct a natural experiment. I compile a sample of firms that have *both* a peer review report in the last AICPA public file, and a PCAOB inspection report available from the first round of PCAOB inspections. For each of these firms, I compare the firm's performance under peer review with its performance under the PCAOB review. Assuming that the PCAOB is an independent, objective reviewer that reviews all firms rigorously, I use the *difference in measured performance* for a given firm between peer

review and PCAOB review, as a measure of the ‘bias’ due to lack of independence in peer reviews. I find that several reviewer and reviewee characteristics are systematically related to this difference in measured performance between the two reviews, my measure of ‘bias’. Firms that themselves review other firms tend to consistently receive more favourable opinions from their peer reviewers than from the PCAOB. Firms that are reviewed by other firms (as opposed to being reviewed by the AICPA or another CPA association) also consistently receive better opinions from peer review than from the PCAOB. Firms being reviewed by smaller firms, or by firms that tend to rely more on peer reviews for their revenues, also tend to receive better opinions from peer review than from the PCAOB. These results are consistent with the situation of *‘you scratch my back, I scratch yours’* that can arise when the reviewee has the freedom to choose and negotiate with its reviewer.

I also corroborate these results using Accounting and Auditing Enforcement Releases as an alternative benchmark for audit quality. I find that clean peer review opinions from ‘friendly’ reviewers are unable to distinguish between firms with AAERs and firms without AAERs. However, peer review opinions from the subset of firms with relatively independent reviewers are able to significantly and clearly distinguish between firms with AAERs and firms without AAERs.

Hillary and Lennox (2005) conclude that peer review opinions were perceived as credible indicators of quality in the audit client market, by showing that firms gained issuer clients after receiving a positive peer review report and lost clients after receiving a negative peer review report. However, the implications of these results are open to interpretation. One possibility is

that peer reviews were indeed objective indicators of audit quality, and were hence perceived (rightly) to be useful information. A second possibility is that the criticisms of peer review were true, and that not all peer reviews were objective or rigorous, *but* that audit committees failed to recognize these issues while using peer review reports to make hire-and-fire decisions. Hillary and Lennox acknowledge this possibility, and are careful to point out that their results should not be viewed as providing unequivocal support for the peer review system. In this context, my results help in disentangling the alternative interpretations by providing evidence for the latter view.

These results should be useful to regulators and academics in understanding one of the factors that contributed to recent accounting and governance failures in the United States. Even though audit firms with SEC issuer clients in the United States are now subject to statutory inspection after Sarbanes-Oxley, audit professions in several countries around the world (and U.S. audit firms with only privately-held clients) continue to be self-regulated in various forms. This evidence might help to interpret the peer review results observed in these spheres.

REFERENCES

AICPA. 2006a. "Large firm PCAOB inspection deficiency analysis (firms with greater than 100 SEC issuer audit clients)"

AICPA. 2006b. "Small firm PCAOB inspection deficiency analysis (firms with 100 or fewer SEC issuer audit clients)"

Austin, K.R. and D.C. Langston. 1981. "Peer review: its impact on quality control" *Journal of Accountancy* 152 (1): 78-82.

Casterella, J.R., K.L. Jensen and W.R. Knechel. 2006. "Is self-regulated peer review effective at signaling audit quality ?" Working paper.

- Coffee, J. C. 2002. "Understanding Enron: It's about the gatekeepers, stupid" *Business Lawyer* 57(4): 1403-1421.
- Edwards, F.R. 2004. "U.S. Corporate Governance: what went wrong and can it be fixed?" in *Market Discipline across Countries and Industries* ed. Borio, et.al. (MIT Press) : 237-254.
- Fogarty, T.J. 1996. "The imagery and reality of peer review in the U.S.: Insights from institutional theory" *Accounting, Organizations and Society* 21 (2/3): 243-267.
- Grumet, L. 26 September, 2005. "Rethinking the 'peer' in peer review" *Accounting Today* 19(17).
- Gunny, K. and T. Zhang. 2006. "The association between earnings quality and regulatory report opinions in the accounting industry: AICPA peer review and PCAOB inspections" Working paper.
- Hermanson, D.R., R.W. Houston and J.C. Rice. 2007. "PCAOB Inspections of smaller CPA firms: initial evidence from inspection reports" *Accounting Horizons* 21 (2): 137-152.
- Hillary, G. and C. Lennox. 2005. "The credibility of self-regulation: Evidence from the accounting profession's peer review program" *Journal of Accounting and Economics* 40: 211-229.
- Mason, E. 24 October 2005. "Quality review and its future vs. peer review and its failure" *Home Equity Wire* 19(19).
- Matusiak, L.W. 1979. "The role of the Public Oversight Board in the accounting profession's self-regulatory program" *The Ohio CPA Journal* 38(4): 144-161.
- Public Oversight Board. March 19, 2002. "The Road to Reform: A white paper from the Public Oversight Board on legislation to create a new private sector regulatory structure for the accounting profession".
- Rankin, D. June 20, 1978. "Peer review: cottage industry" *The New York Times*.
- Russell, J. and J. Armitage. 2006. "Peer review effectiveness: an analysis of potential loopholes within the USA Peer Review Program" *Managerial Auditing Journal* 21 (1): 46-62.
- Sepp, P. 1983. "The peer review experience" *Journal of Accountancy* 155 (5): 52-56.
- Turner, L.E. January 27, 2000. "Shifting paradigms in self-regulation" *Speech at the 27th Anniversary of the Securities Regulation Institute*.
- Wallace, W.A. 1991. "Peer review filings and their implications in evaluating self-regulation" *Auditing: A Journal of Practice and Theory* 10(1): 53-67.

Figure 1: A timeline of self-regulation

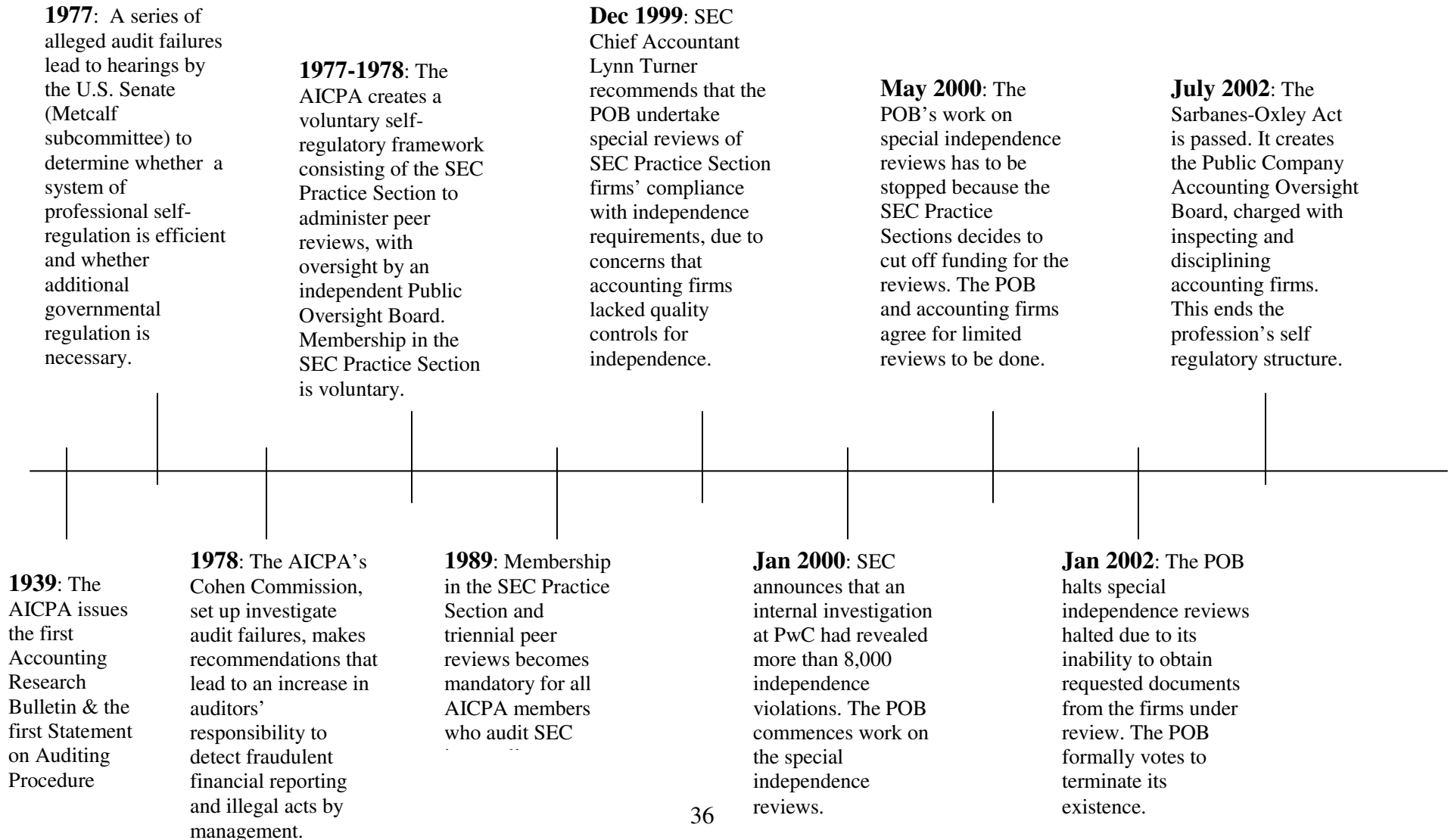


Table I – Descriptive statistics of peer review opinions

Panel A: Type of peer review opinion

All firms with peer review opinions ($n = 1,001$)

<i>Type of opinion</i>	Percentage of sample
Clean	45.25
Unmodified	50.65
Modified	3.40
Adverse	0.70
	100.00 %

Firms with peer review opinions *and* PCAOB opinions ($n = 355$)

<i>Type of opinion</i>	Percentage of sample
Clean	44.79
Unmodified	50.70
Modified	3.95
Adverse	0.28
	100.00 %

Panel B: Descriptive statistics of PEER_REVIEW_OPINION

All firms with peer review opinions ($n = 1,001$)

<i>Number of weaknesses per opinion</i>	<i>Type of opinion</i>			<i>Totals</i>	# of weaknesses
	<i>Unmodified</i>	<i>Modified</i>	<i>Adverse</i>		
0	453	0	0	453	0
1	233	8	0	241	241
2	142	11	0	153	306
3	80	7	0	87	261
4	30	6	1	37	148
5	17	2	1	20	100
6	4	0	2	6	36
7	0	0	1	1	7
8	1	0	0	1	8
9	0	0	2	2	18
Totals	960	34	7	1001	1125
Mean number of weaknesses per opinion	1.035	2.500	6.570		

Firms with peer review opinions and PCAOB opinions ($n = 355$)

<i>Number of weaknesses per opinion</i>	<i>Type of opinion</i>			<i>Totals</i>	# of weaknesses
	<i>Unmodified</i>	<i>Modified</i>	<i>Adverse</i>		
0	159	0	0	159	0
1	77	2	0	79	79
2	51	7	0	58	116
3	31	1	0	32	96
4	14	4	1	19	76
5	5	0	0	5	25
6	2	0	1	3	18
7	0	0	0	0	0
8	0	0	0	0	0
9	0	0	0	0	0
Totals	339	14	2	355	410
Mean number of weaknesses per opinion	1.077	2.5	5		

PEER_REVIEW_OPINION is defined as follows: Clean opinion = 0; Unmodified opinion with one weakness = 1, Unmodified opinion with two weaknesses = 2, Unmodified opinions with three weaknesses = 3, ...Unmodified opinions with nine weaknesses = 9; Modified opinion with one weakness = 11, Modified opinion with two weaknesses = 12, Modified opinion with three weaknesses = 13, ...Modified opinion with nine weaknesses = 19; Adverse opinion with one weakness = 21, Adverse opinion with two weaknesses = 22, ...Adverse opinion with nine weaknesses = 29.

Table II – Descriptive statistics of PCAOB opinions

Panel A: Descriptive statistics of PCAOB opinions

All firms with PCAOB opinions (*n* = 430)

<i>PCAOB_OPINION</i>	Percentage of sample
“Clean” opinion	29.30
Some defects identified	70.70
	100.00

Firms with PCAOB opinions *and* peer review opinions (*n* = 355)

<i>PCAOB_OPINION</i>	Percentage of sample
“Clean” opinion	33.24
Some defects identified	66.76
	100.00

<i>PCAOB_OPINION</i>	<i>Min</i>	<i>Q1</i>	<i>Q2</i>	<i>Q3</i>	<i>Max</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Skewness</i>	Observations
All firms with PCAOB opinions	0	0	2	2	9	1.479	1.291	1.503	430
Firms with PCAOB <i>and</i> peer review opinions	0	0	2	2	9	1.386	1.249	1.341	355

Panel B: Types of problems identified by PCAOB inspections (n = 355)

<i>Number of departures from GAAP identified in each inspection</i>	Percentage of sample
None	88.45
1	8.45
2	2.25
3	0.57
More than 3	0.28
	100.00 %

<i>Whether any quality control defects identified</i>	Percentage of sample
None identified	33.52
Some defects identified	66.48
	100.00 %

<i>Number of other audit performance deficiencies identified in each inspection</i>	Percentage of sample
None	46.76
1-5	46.48
6-10	3.38
11-15	1.97
16-20	0.28
More than 20	1.13
	100.00 %

PCAOB_OPINION for each firm undergoing inspection is defined as #DEPARTURES_GAAP + DUMMY_DEFICIENCIES + DUMMY_QC_DEFECTS where:

#DEPARTURES_GAAP is defined as the number of potentially material mis-statements arising from failures of the audit firm to identify, or to address appropriately departures from GAAP in issuers' financial statements, that were identified by the PCAOB inspection of each audit firm.

DUMMY_DEFICIENCIES is set to 1 if the PCAOB inspection identifies any deficiencies in the firm's performance of audit engagements inspected, and 0 otherwise.

DUMMY_QC_DEFECTS is set to 1 if the PCAOB inspection identifies any defects in the firm's quality control systems, and 0 otherwise.

Table III: Characteristics of reviewed and reviewing firms

<i>Characteristics of reviewed firms (n=355)</i>	<i>Minimum</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Maximum</i>	<i>Mean</i>	<i>Standard Deviation</i>	Skewness
SIZE_REVIEWED	0	1	2	5	2975	30.651	233.554	10.393
LN_SIZE_REVIEWED	0	0.693	1.099	1.792	7.998	1.401	1.144	2.436
REVIEWED_FIRM_REVIEWS	0	0	0	0	1	0.177	0.382	1.688
TINY_FIRM	0	0	0	1	1	0.383	0.487	0.481
BIG_FIRM	0	0	0	0	1	0.031	0.174	5.413

<i>Type of review</i>	<i>#Observations</i>	% of sample
Firm-on-firm	332	93.52
AICPA-appointed team	5	1.41
Other CPA association	18	5.07
Total	355	100 %

<i>Characteristics of reviewing firms (144 unique reviewing firms)</i>	<i>Minimum</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Maximum</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>Skewness</i>	# of obs
SIZE_REVIEWER	0	0	2	5	1969	39.479	240.373	6.773	332
LN_SIZE_REVIEWER	0	0	1.099	1.792	7.586	1.133	1.341	2.391	332
NUMBER_OF_MILES	0	54	237	549	2418	378.821	442.288	1.820	332
NON_COMPETITOR	0	1	1	1	1	0.762	0.426	-1.231	332
NUMBER_OF_REVIEWS_DONE	1	1	3	7	36	5.368	6.010	2.291	144
AUDITS_TO_REVIEWS_RATIO	0	0	0.195	1	420	11.128	57.969	5.925	144
LN_AUDITS_TO_REVIEWS	0	0	0.178	0.693	6.043	0.588	1.112	3.396	144
FRIENDLY_SCORE	0	2	3	4	4	2.946	1.135	-1.102	355

Reviewee characteristics

SIZE_REVIEWED is defined as the number of SEC issuer clients of the audit firm that is being peer-reviewed.

LN_SIZE_REVIEWED is defined as natural log of (1 + SIZE_REVIEWED).

REVIEWED_FIRM_REVIEWS is set to 1 if the audit firm being peer-reviewed has performed reviews on other firms in the sample, and 0 if not.

TINY_FIRM is set to 1 if the number of clients at the time of peer review is less than or equal to 1; and set to 0 otherwise.

BIG_FIRM is set to 1 if the number of clients at the time of peer review is greater than 50; and set to 0 otherwise.

Type of review

FIRM_ON_FIRM is set to 1 if the review is being performed by another audit firm, and 0 if the review is performed by an AICPA-appointed team or by another CPA association.

Reviewer characteristics (defined for firm-on-firm reviews only)

SIZE_REVIEWER is defined as the number of SEC issuer clients of the audit firm performing the peer review.

LN_SIZE_REVIEWER is defined as the natural log of (1 + SIZE_REVIEWER)

NUMBER_OF_MILES is the distance in miles between the location of the reviewed firm and the reviewing firm. The distance is set equal to 1 mile if the reviewed or reviewing firm is a Big 4 firm.

NON_COMPETITOR is set to 1 if the firm performing the review is located more than 50 miles away from the reviewed firm, and 0 if the firm performing the review is located less than 50 miles away from the reviewed firm.

NUMBER_OF_REVIEWS_DONE is defined as the number of peer reviews performed by the reviewing firm in the time frame covered by the last AICPA public file.

AUDITS_TO_REVIEWS_RATIO for the reviewing firm is defined as SIZE_REVIEWER / NUMBER_OF_REVIEWS_DONE.

LN_AUDITS_TO_REVIEWS is defined as the natural logarithm of (1+AUDITS_TO_REVIEWS_RATIO).

FRIENDLY_SCORE is defined in the following way (only for entities that perform a peer review):

FIRM_ON_FIRM + NON_COMPETITOR + (1 if SIZE_REVIEWER < SIZE_REVIEWED; 0 otherwise) + (1 if AUDITS_TO_REVIEWS_RATIO for the reviewer is < 1; 0 otherwise)

Table IV: Measuring bias in peer review opinions using PCAOB opinions as a benchmark

<i>Measure of bias</i>	<i>Min</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Max</i>	<i>Mean</i>	<i>Std Dev</i>	# of obs
DIFF	-289.5	-62	-20.5	61	268.5	0	116.378	355

<i>Measure of bias</i>	<i># of obs</i>	% of sample
DIFF_CLEAN = +1	48	13.52
DIFF_CLEAN = 0	218	61.41
DIFF_CLEAN = -1	89	25.07
		100.00

DIFF is measured as Rank(PEER_REVIEW_OPINION) – Rank(PCAOB_OPINION)

Rank(PEER_REVIEW_OPINION) is the rank of PEER_REVIEW_OPINION, sorted in reverse order. Higher ranks correspond to better performance in the peer review.

Rank(PCAOB_OPINION) is the rank of PCAOB_OPINION, sorted in reverse order. Higher ranks correspond to better performance in the PCAOB inspection.

DIFF_CLEAN is set :

Equal to 1 if the firm receives a clean opinion in peer review but not in PCAOB review.

Equal to 0 if either the firm receives a clean opinion from *both* peer review and PCAOB review; or if the firm does *not* receive a clean opinion from *both* peer review and PCAOB review.

Equal to -1 if the firm does not receive a clean opinion in peer review but receives a clean opinion in PCAOB review.

Table V: Univariate correlations

Panel A: Spearman correlations between peer review opinions and PCAOB opinions

<i>Spearman correlations</i>	<i>PEER_REVIEW_OPINION</i>	<i>PCAOB_OPINION</i>	<i>#DEPARTURES_GAAP</i>	<i>DUMMY_QC_DEFECTS</i>
PCAOB_OPINION	0.275 (0.000)***			
#DEPARTURES_GAAP	0.226 (0.000)***	0.565 (0.000)***		
DUMMY_QC_DEFECTS	0.222 (0.000)***	0.871 (0.000)***	0.256 (0.000)***	
#DEFICIENCIES	0.237 (0.000)***	0.868 (0.000)***	0.368 (0.000)***	0.684 (0.000)***

*, **, *** refers to significance at the 10 %, 5 % and 1% level respectively

PEER_REVIEW_OPINION is defined as follows: Clean opinion = 0; Unmodified opinion with one weakness = 1, Unmodified opinion with two weaknesses = 2, Unmodified opinions with three weaknesses = 3, ...Unmodified opinions with nine weaknesses = 9; Modified opinion with one weakness = 11, Modified opinion with two weaknesses = 12, Modified opinion with three weaknesses = 13, ...Modified opinion with nine weaknesses = 19; Adverse opinion with one weakness = 21, Adverse opinion with two weaknesses = 22, ...Adverse opinion with nine weaknesses = 29.

PCAOB_OPINION is defined as #DEPARTURES_GAAP + DUMMY_DEFICIENCIES + DUMMY_QC_DEFECTS

#DEPARTURES_GAAP is the number of potentially material misstatements in clients' financial statements arising from departures from GAAP, identified by the PCAOB inspection of that audit firm.

DUMMY_QC_DEFECTS is set to 1 if the PCAOB inspection of an audit firm identifies any defects in the firm's quality control systems, and 0 otherwise.

DUMMY_DEFICIENCIES is set to 1 if the PCAOB inspection identifies any deficiencies in the firm's performance of audit engagements inspected, and 0 otherwise

Panel B: Spearman correlations between DIFF and explanatory variables

<i>Spearman correlations</i>	<i>DIFF</i>	<i>DIFF_CLEA N</i>	<i>LN_SIZE_R EVIEWED</i>	<i>REVIEWED _FIRM_REV IEWS</i>	<i>FIRM_ON_ FIRM</i>	<i>NON_COM PETITOR</i>	<i>LN_SIZE_R EVIEWER</i>	<i>LN_AUDITS _TO_REVIEW S</i>
DIFF_CLEAN	0.852 (0.000)***							
LN_SIZE_RE VIEWED	0.100 (0.061)*	0.065 (0.222)						
REVIEWED_F IRM_REVIEW S	0.133 (0.012)**	0.093 (0.082)*	0.133 (0.012)**					
FIRM_ON_FI RM	0.090 (0.092)*	0.085 (0.112)	0.095 (0.074)*	0.002 (0.963)				
NON_COMPE TITOR	0.059 (0.271)	0.092 (0.085)*	-0.033 (0.535)	-0.053 (0.321)	0.141 (0.008)***			
LN_SIZE_RE VIEWER	-0.015 (0.783)	-0.079 (0.140)	0.253 (0.000)***	0.094 (0.077)*	0.287 (0.000)***	-0.151 (0.004)***		
LN_AUDITS_ TO_REVIEW S	-0.019 (0.715)	-0.070 (0.191)	0.223 (0.000)***	0.083 (0.121)	0.286 (0.000)***	-0.128 (0.016)**	0.917 (0.000)***	
FRIENDLY_S CORE	0.113 (0.033)**	0.166 (0.002)***	0.230 (0.000)***	-0.010 (0.853)	0.450 (0.000)***	0.418 (0.000)***	-0.384 (0.000)***	-0.399 (0.000)***

*, **, *** refers to significance at the 10 %, 5 % and 1% level respectively

Variable definitions are in Tables III and IV.

Table VI: Are all peer reviews objective?

Panel A: Reviewee characteristics only

	<i>Dependent variable is:</i>	
	<i>DIFF</i>	<i>DIFF_CLEAN</i>
LN_SIZE_REVIEWED	17.695 (0.051)**	0.102 (0.265)
REVIEWED_FIRM_REVIEWS	34.870 (0.032)***	0.266 (0.109)*
FIRM_ON_FIRM_REVIEW	34.237 (0.107)*	0.388 (0.133)*
<i>Other explanatory variables</i>		
TINY_FIRM	25.925 (0.148)*	0.108 (0.549)
BIG_FIRM	0.639 (0.989)	-0.447 (0.361)
Constant	-72.954 (0.006)***	
#Observations	355	355
Adjusted R ²	4.67 %	1.05 %

*, **, *** refers to significance at the 15 %, 10 % and 5% level respectively. Standard errors are estimated using the robust estimator.

Panel B: Reviewee and reviewer characteristics (for firm-on-firm reviews only)

Dependent variable is:

	<i>DIFF</i>			<i>DIFF_CLEAN</i>		
<i>Reviewee characteristics</i>						
LN_SIZE_REVIEWED	26.136 (0.006)***	25.627 (0.005)***	16.908 (0.052)**	0.214 (0.044)***	0.207 (0.048)***	0.093 (0.308)
REVIEWED_FIRM_REVIEWS	30.921 (0.057)**	30.373 (0.061)**	34.290 (0.027)***	0.197 (0.249)	0.190 (0.260)	0.262 (0.120)*
<i>Reviewer characteristics</i>						
NON_COMPETITOR	30.756 (0.021)***	31.546 (0.022)***		0.312 (0.027)***	0.321 (0.029)***	
LN_SIZE_REVIEWER	-13.020 (0.045)***			-0.171 (0.010)***		
LN_AUDITS_TO_REVIEWS		-16.103 (0.050)***			-0.212 (0.025)***	
FRIENDLY_SCORE			14.026 (0.017)***			0.174 (0.002)***
<i>Control variables</i>						
TINY_FIRM	32.771 (0.083)**	33.585 (0.071)**	31.057 (0.083)**	0.206 (0.292)	0.216 (0.259)	0.172 (0.346)

BIG_FIRM	37.738 (0.410)	35.201 (0.471)	21.208 (0.624)	0.023 (0.965)	-0.007 (0.990)	-0.193 (0.686)
Constant	-62.574 (0.007)***	-70.073 (0.003)***	-83.658 (0.001)***			
#Observations	332	332	355	332	332	355
R ²	6.21 %	5.72 %	5.89 %	2.41 %	1.99 %	2.16 %
Number of clusters	144	144	150	144	144	150

*, **, *** refers to significance at the 10 %, 5 % and 1% level respectively. Standard errors are clustered by the identity of the reviewing firm.

Dependent variables

DIFF is measured as Rank(PEER_REVIEW_OPINION) – Rank(PCAOB_OPINION)

Rank(PEER_REVIEW_OPINION) is the rank of PEER_REVIEW_OPINION, sorted in reverse order. Higher ranks correspond to better performance in the peer review.

Rank(PCAOB_OPINION) is the rank of PCAOB_OPINION, sorted in reverse order. Higher ranks correspond to better performance in the PCAOB inspection.

DIFF_CLEAN is set :

Equal to 1 if the firm receives a clean opinion in peer review but not in PCAOB review.

Equal to 0 if either the firm receives a clean opinion from *both* peer review and PCAOB review; or if the firm does *not* receive a clean opinion from *both* peer review and PCAOB review.

Equal to -1 if the firm does not receive a clean opinion in peer review but receives a clean opinion in PCAOB review.

PCAOB_OPINION is defined as #DEPARTURES_GAAP + DUMMY_QC_DEFECTS + DUMMY_DEFICIENCIES

Explanatory variables

Reviewee characteristics

LN_SIZE_REVIEWED is defined as natural log of the number of SEC issuer clients of the audit firm that is being peer-reviewed.

REVIEWED_FIRM_REVIEWS is set to 1 if the audit firm being peer-reviewed has performed reviews on other firms in the sample, and 0 if not.

Type of review

FIRM_ON_FIRM is set to 1 if the review is being performed by another audit firm, and 0 if the review is performed by an AICPA-appointed team or by another CPA association.

Reviewer characteristics (defined only for firm-on-firm reviews)

NON_COMPETITOR is set to 1 if the firm performing the review is located more than 50 miles away from the reviewed firm, and 0 if the firm performing the review is located less than 50 miles away from the reviewed firm. If the review is performed by an AICPA team or CPA association, NON_COMPETITOR is set to 1.

LN_SIZE_REVIEWER is defined as the natural log of (1+ SIZE_REVIEWER), where

LN_AUDITS_TO_REVIEWS is defined as the natural logarithm of (1+AUDITS_TO_REVIEWS_RATIO), where

FRIENDLY_SCORE is defined in the following way (only for entities that perform a peer review):

FIRM_ON_FIRM + NON_COMPETITOR + (1 if SIZE_REVIEWER < SIZE_REVIEWED; 0 otherwise) + (1 if AUDITS_TO_REVIEWS_RATIO for the reviewer is < 1; 0 otherwise)

Other control variables

TINY_FIRM is set to 1 if the number of clients at the time of peer review is less than or equal to 1; and set to 0 otherwise.

BIG_FIRM is set to 1 if the number of clients at the time of peer review is greater than 50; and set to 0 otherwise.

Table VII: Are all peer reviews objective? Analysis excluding big firms

These models are estimated after excluding all firms with more than 50 clients at the time of peer review.

<i>Dependent variable is:</i>	<i>DIFF</i>			<i>DIFF_CLEAN</i>		
<i>Reviewee characteristics</i>						
LN_SIZE_REVIEWED	22.360 (0.032)***	27.207 (0.009)***	19.961 (0.054)**	0.188 (0.075)**	0.261 (0.205)	0.161 (0.141)*
REVIEWED_FIRM_REVIEWS	38.394 (0.023)**	34.949 (0.035)***	37.961 (0.054)**	0.291 (0.089)**	0.256 (0.027)***	0.289 (0.091)**
FIRM_ON_FIRM_REVIEW	33.475 (0.113)*			0.370 (0.147)*	0.225 (0.192)	
<i>Reviewer characteristics</i>						
NON_COMPETITOR		33.578 (0.014)***			0.324 (0.020)***	
LN_SIZE_REVIEWER		-12.904 (0.086)**			-0.163 (0.030)***	
FRIENDLY_SCORE			14.139 (0.019)***			0.166 (0.003)***
<i>Control variables</i>						

TINY_FIRM	31.669 (0.098)**	34.464 (0.086)**		0.212 (0.268)		
Constant	-80.964 (0.004)***					
#Observations	344	321	344	344	321	344
R ²	3.30 %	4.82 %	4.56 %	1.35 %	2.50 %	2.35 %
Number of clusters		139	145		139	145

*, **, *** refers to significance at the 15 %, 10 % and 5% level respectively. Standard errors are clustered by the identity of the reviewing firm.

Variable definitions and notes are in Table VIII.

Table VIII: Robustness check: are peer review opinions associated with AAERs?

Panel A: Frequency of AAERs

<i>All firms with peer review opinions (n = 1,001)</i>	<i># of obs</i>	<i>% of sample</i>
Firms with AAER clients (AAER =1)	33	3.30
Firms with no AAER clients (AAER = 0)	968	96.70
	1,001	100.00

Panel B: Ability of clean peer review opinions to predict AAERs

All firms with peer review opinions (n = 1,001)

<i>Dependent variable is the % of audit firms in each category that have AAER clients.</i>	<i>Received clean peer review</i>	<i>Did not receive clean peer review</i>	<i>Difference</i>	p-value
All firms	1.3 % (n = 453)	4.9 % (n = 548)	3.6 %	0.002***
Partitioned by reviewer:				
FRIENDLY_SCORE ≤ 2 (n = 218) (more likely to be independent)	1.1 % (n = 94)	9.7 % (n = 124)	8.6 %	0.008***
FRIENDLY_SCORE = 3 (n = 414)	1.0 % (n = 203)	3.3 % (n = 211)	2.3 %	0.104*
FRIENDLY_SCORE = 4 (n = 369) (more likely to be friendly)	1.9 % (n = 156)	3.8 % (n = 213)	1.9 %	0.308

AAER is set to 1 if the audit firm has issuer clients with Accounting and Auditing Enforcement Releases at the time of peer review; and set to 0 otherwise.

FRIENDLY_SCORE is defined in the following way (only for entities that perform a peer review):

FIRM_ON_FIRM + NON_COMPETITOR + (1 if SIZE_REVIEWER < SIZE_REVIEWED; 0 otherwise) + (1 if AUDITS_TO_REVIEWS_RATIO for the reviewer is < 1; 0 otherwise)

All other definitions are in Table III.

Table IX: Do friendly peer reviews have credibility?

Panel A: Gains and losses of audit clients following peer reviews

<i>All firms with peer review opinions (n = 1,001)</i>	<i>Min</i>	<i>Q1</i>	<i>Q2</i>	<i>Q3</i>	<i>Max</i>	<i>Mean</i>	<i>Std Dev</i>	<i># of obs</i>
PCT_CHANGE_PLUS12	-200	0	0	0	300	0.674	0.230	1001

<i>All firms with peer review opinions (n = 1,001)</i>	<i># of obs</i>	<i>% of sample</i>
Firms with net gain of clients	111	11.09
Firms with no net change	757	75.62
Firms with net loss of clients	133	13.29
	1,001	100.00

Panel B: Net gains and losses of audit clients partitioned by opinion and reviewer

All firms with peer review opinions (n = 1,001)

<i>Dependent variable is the net % gain or loss of audit clients in the year following peer review(PCT_CHANGE_PLUS12)</i>	<i>Received clean peer review</i>	<i>Did not receive clean peer review</i>	<i>Difference</i>	<i>p-value</i>
All firms	3.3 % (n = 453)	-1.5 % (n = 548)	4.8 %	0.001***
Partitioned by reviewer:				
FRIENDLY_SCORE ≤ 2 (n = 218) (more likely to be independent)	3.4 % (n = 94)	-3.8 % (n = 124)	7.2 %	0.018**
FRIENDLY_SCORE = 3 (n = 414)	2.3 % (n = 203)	-0.3 % (n = 211)	2.6 %	0.305
FRIENDLY_SCORE = 4 (n = 369) (more likely to be friendly)	4.6 % (n = 156)	-1.5 % (n = 213)	6.1 %	0.004***

PCT_CHANGE_PLUS12 is the % change in the number of audit clients in the year following peer review.

All other definitions are in Table III.