The Impact of Consulting Services on Audit Quality: An Experimental Approach

Zach Kowaleski
PhD Candidate
University of Wisconsin – Madison

Brian W. Mayhew
Arthur Andersen Alumni Professor
University of Wisconsin – Madison

Amy C. Tegeler
Assistant Professor
University of Wisconsin – Milwaukee

December 21, 2017

Accepted by Philip Berger. This paper is the final Registered Report resulting from the Registration-based Editorial Process (REP) implemented by JAR for its 2017 conference; details of the process are available here https://research.chicagobooth.edu/-/media/research/arc/docs/journal/rep_policies_jar.pdf?la=en&hash=B64B8F1D368D8300BC83898C7E5CFD0557E71312 .

The accepted proposal for this report is available here https://research.chicagobooth.edu/arc/journal-of-accounting-research/online-supplements

We thank the editor and referee for insightful comments and suggestions in revising our proposal. We thank JAR conference attendees and workshop participants at the University of Wisconsin – Madison for providing helpful comments and suggestions. We thank the University of Wisconsin – Madison BRITE Lab for use of its facilities and funding for our pilot studies. We also thank the University of Wisconsin – Madison for research support. While Zach Kowaleski was an Economic Research Fellow at the PCAOB between October 2016 and October 2017, he co-wrote this paper in his personal capacity. The views expressed in this paper are the views of the authors and do not necessarily reflect the views of the Board, individual Board members or staff of the PCAOB.

Electronic copy available at: https://ssrn.com/abstract=3095694

The Impact of Consulting Services on Audit Quality:
An Experimental Approach

Abstract

We use experimental markets to examine whether providing consulting services to a nonaudit client impacts audit quality. Our paper directly addresses concerns raised by the Public Company Accounting Oversight Board that the largest public accounting firms' growth in their consulting practices threatens audit quality. We conduct an experiment proposed using a registration-based editorial process. We compare a baseline where the auditor does not provide consulting services to conditions where auditors provide consulting to audit clients or where auditors only provide consulting services to nonaudit clients. Our unique design provides evidence on whether providing consulting to nonaudit clients strengthens the salience of a client-cooperative social norm that reduces audit quality. We do not find differences in audit quality by condition in our planned analysis, however we find greater variation in audit quality in the conditions where auditors provide consulting services compared to the baseline. In unplanned analyses, our results suggest providing consulting services increases auditor cooperation with managers, increasing audit quality when managers prefer high audit quality and decreasing audit quality when managers prefer low audit quality.

Keywords: Auditor Independence, Consulting, Social Bonds, Experimental Economics

JEL Code: G41, M42

Electronic copy available at: https://ssrn.com/abstract=3095694

I. Introduction

We use experimental markets to examine whether the provision of consulting services to a nonaudit client impacts audit quality. Our paper directly addresses concerns raised by members of the Public Company Accounting Oversight Board (PCAOB) that the largest auditing firms' evolving business model to a heavy emphasis on their consulting practice threatens audit quality (Harris [2015], Doty [2015]). Over the past ten years, the largest firms have built nonaudit service practices that generate 41% of U.S. Big 4 total revenue.[1] These services now exceed the 34% of revenues generated by the audit practice. With nonaudit service revenue growing at a 10% rate, and assurance revenue at a smaller 4% rate, the gap will continue to grow (Harris [2014]).

We examine the impact of a client-cooperative social norm arising from the consulting practice on audit quality. In 2002, legislators created rules restricting the types of nonaudit services that can be performed for audit clients to address the perception that these services negatively impact audit quality (Sarbanes-Oxley [2002] (SOX)).[2] This constraint requires the Big 4 to shift consulting growth to nonaudit clients. As such, the recent nonaudit service growth does not threaten auditor independence. Nevertheless, regulators have raised new concerns around the changing culture within large accounting firms (Harris [2014]). In support of these concerns, analytical models of endogenous social norms suggest organization boundaries matter in preserving high professional norms (Fischer & Huddart [2008]). As accounting firms have become diversified,

[1] Practice revenue percentages are calculated from the Big 4 firm's annual reports to trade magazine, *Accounting Today* as presented in Figure 1.

[2] We use the terms nonaudit services and consulting interchangeably in this paper. We intend both labels to encompass consulting and advisory services that exclude tax and audit services. The SEC requires separate disclosure for tax services in recognition of the difference between tax and consulting.

complex organizations focused on goals of commercial growth and success, the norms of accounting professionalism can come in conflict with the consultants' commercial norms.

We test whether the provision of consulting services impacts audit quality by strengthening the salience of client-cooperative social norms. Consulting services rely on providers to reciprocate a client's hiring decision by cooperating with management. This strategy leads to repeat hiring that maximizes the consultant's profit. In contrast, an auditor, who is hired by a client, serves both the client management and the investing public. At times, auditors must set aside manager preferences for those of investors. We explore whether a cooperative social norm from a consulting task influences subsequent auditor decisions to cooperate with management's reporting preferences.

Our experiment was subject to a registration-based editorial process in which our experimental design and planned analysis were subject to referee and editor critique prior to acceptance. We employ a 1 X 3 design where we manipulate the provision of a consulting project and the rules governing the managers' choice of auditor. The *Baseline* condition consists of an audit market similar to previous audit market experiments (Mayhew [2001], Mayhew and Pike [2004], Hurley et al. [2016]). In this market, managers produce and sell assets to buyers, and hire auditors to verify the manager's report of the asset's value.[3] We measure the auditors' verification decisions to capture audit quality. In our nonaudit service treatments, we incorporate a second game in addition to the audit market. We use a modified trust game designed to capture the key

[3] Audit committees officially hire auditors for publicly traded companies. We simplify the hiring decision and allocate it to managers in our experimental setting. In our experiment, the manager role represents the client as a whole – both manager and board committees. Further, recent research finds management has significant influence over the auditor hiring decision (Beasley et al. [2009], Fiolleau et al. [2013]).

incentives of a typical consulting engagement. The consulting market begins with managers hiring an auditor-consultant to serve as a consultant on a project. Auditor-consultants choose how much costly effort to put into the project. The effort choice influences the project's value to the manager. In our multi-period setting, the auditor-consultant has a strong incentive to cooperate with the manager by supplying high effort to garner repeat business. We alternate the consulting market with the audit market in each period of the nonaudit service conditions. We employ two nonaudit service (*NAS*) treatments, *NAS-same* in which the auditor-consultant is hired by the same manager for both the consulting project and the audit market, and *NAS-separate* in which the auditor-consultant is hired by separate groups of managers in the consulting market and the audit market. *NAS-same* features consulting with strong economic and social bonding between the auditor and client. *NAS-separate* narrows the consulting market impact to just the social norm of cooperation as there is no economic overlap between providing the audit service and the consulting service.[4]

We predict auditors will cooperate more with managers in the *NAS* treatments than in *Baseline*. We expect this cooperation to manifest as lower audit quality. We do not find differences in audit quality by condition in our planned analysis. In an unplanned analysis, we find greater variation in audit quality in the *NAS* treatments than the *Baseline*. Half the *NAS* sessions result in high quality (HQ) markets, with managers preferring the highest quality auditor providing > 90% high audit quality. These *NAS HQ*

[4] In our original proposal available online, we planned two follow up treatments contingent upon the main experiment's results. The additional treatment would explore the effect on audit quality of either the strength of the cooperative norm or the degree of the auditor's economic independence from managers. Based on our results, we determined neither treatment was appropriate for follow up.

sessions have significantly higher audit quality than *Baseline* and the *NAS non-HQ* sessions. *NAS non-HQ* has significantly lower audit quality than *Baseline*. This pattern of differences suggests that consulting services do not lead to uniformly higher or lower audit quality, but to more extremes. The *NAS* sessions both start with higher audit quality than *Baseline* in early periods before individual sessions diverge to more extreme high and low quality levels. The results of these unplanned analyses, along with responses from the post-experiment questionnaire suggest that cooperating in the consulting market leads to cooperation in the audit market, but contrary to our prediction, cooperation does not manifest itself simply as lower audit quality. Instead, cooperation in the audit market is dictated by manager preferences. In sessions where managers prefer high audit quality to add credibility to their reports, a cooperative auditor performs high-level verification. In sessions where managers prefer lower audit quality, a cooperative auditor performs more low-level verifications. The provision of consulting services amplifies auditor responses to manager's real or perceived preferences.

Our research question differs substantially from prior research on the auditor provision of nonaudit services. Prior research focuses heavily on the economic bond between the client and auditor that results from nonaudit services. In general, archival research finds little evidence of an incremental effect of the joint provision of consulting and auditing on audit quality (Ashbaugh, et al. [2003], DeFond et al. [2002], Myers et al. [2003], Church et al. [2015]). Given the concentrated audit market, endogeneity concerns over auditor hiring decisions, and the absence of control firms, the joint provision of audit and consulting is not easily examined using archival data and econometric models (Church et al. [2015], Schneider et al. [2006]). Recent archival studies employing more

focused tests and experimental research generally have found evidence that an economic bond between clients and auditors biases auditor judgment (see Church et al. [2015]). Our study makes a unique contribution by studying consulting's impact on norms rather than economic bonding. Our treatments focus on auditor choices when that auditor provides consulting services both with and without a direct economic bond thereby enabling us to isolate the bond and norm effects. In contrast, prior research implicitly studies the combined effects of bonding incentives and social norms. We believe controlled research on this topic provides useful insights.

The remainder of this paper is organized as follows: section II reviews the relevant background and literature; section III outlines our experimental design; section IV discusses our predictions; section V presents the results of our planned analyses and unplanned analysis; section VI discusses our interpretation of these results; and section VII concludes.

II. Background and literature review

At their 1998 peak, consulting services contributed an average of 45% of Big 5 revenue (US GAO [2003]). A series of financial reporting scandals in the late 1990s and early 2000s at companies such as Enron resulted in vast changes to the regulation and oversight of the profession.[5] Among the many new regulations, SOX restricted the scope of services auditors can provide to their audit clients in an effort to enhance auditor independence. Specifically, SOX banned certain nonaudit services, such as bookkeeping,

[5] Previous audit crises arose in the 1980s with the Savings and Loan crisis (Erickson, Mayhew and Felix [2000]), and prior to that in the 1970s from events leading to the Cohen and Metcalf committees (Clikeman [2013]).

systems implementations, and internal audit outsourcing (SOX section 201).[6] As a result, four of the Big 5 divested their consulting practices in the early 2000s. However, the surviving firms have since rebuilt their consulting practices (Dey et al. [2012], Harris [2015]).

The nonaudit services restricted by SOX are now provided to nonaudit clients.[7] Building a robust nonaudit service line allows firms to attract top talent, develop in-house specialists, and benefit from potentially higher margins relative to their audit service line. In support of this growth, Elitzur, Gavious, and Minchuk (2016) analytically show that banning nonaudit services to audit clients increases both audit and nonaudit fees.

PCAOB members are concerned about the largest firms' evolving business model and the resulting threat to audit quality (Harris [2014], Harris [2015]). The PCAOB Strategic Plan [2015] identifies a threat posed by the "challenge of anticipating the implication of the expansion of consulting on auditor independence and audit quality." Within the Big 4, nonaudit services doubled from $8.5 billion in 2009 to $17.6 billion in 2014 as compared to the $1.5 billion increase in the $15 billion audit sector over the same period.[8] This disparate growth rate in the U.S. produced nonaudit services that now make up 41% of total firm revenues, while audit makes up only 34%.[9]

[6] Audit firms are permitted to provide any service that is not specifically prohibited, so long as the audit committee approves it. Common examples of permissible NAS to provide to audit clients include due diligence related to registration statements, comfort letters, tax services, employee benefit plan audits, and transaction services.

[7] Examples of NAS provided to nonaudit clients include enterprise strategy, marketing and sales, corporate finance, mergers and acquisitions, government consulting, legal services, and a wide variety of risk management services (Harris [2014]).

[8] This growth has been facilitated by acquisitions. From 2011 to 2013, the Big 4 global firms acquired 66 consulting entities (Harris [2014]).

[9] Practice revenue percentages are calculated from the Big 4 firm's annual reports to trade magazine, *Accounting Today* as presented in Figure 1.

The personnel in audit firms has shifted with the move to consulting. Audit firms use their expanding consulting practices as a selling point to recruit top students into the audit practice, as well as to increase hiring from non-accounting fields of study, such as information technology, computer science, finance, and statistics (Harris [2015]). Personnel with advisory backgrounds are also in leadership positions. Mowchan (2016) provides evidence that offices that change from a non-advisory to an advisory office managing partner experience a decrease in audit quality consistent with the concerns raised by PCAOB members. We note that a Big 4 firm named a nonaudit service partner as its top executive (Rapoport [2015]).

The nature of audit and consulting services differs substantially which could lead to challenges managing a firm that provides both services. An audit's value derives from the competence and independence of the auditor. An audit's ultimate consumer is the investing public, who rely on the auditor to objectively audit and report on the company's financial statements. In effect, the investing public demands the auditor be independent of the client by putting the public's interest ahead of the client's interests. Contrast this with nonaudit services. As a consultant, the goal is to work cooperatively with management without a direct link to the investing public. We assert that the consultant's ultimate goal is to develop a cooperative client relationship that will foster additional consulting services. While such services may benefit investors, the consultant does not have an investor focus or a fiduciary duty to investors. The resulting policy concern is that the different service lines create conflicts and inherent biases across lines, even if not provided to common clients (Harris [2014]).

In this paper, we study individual decision making in a market context. The PCAOB's concern about consulting growth within public accounting firms implicitly assumes that individual auditor decisions are impacted by the firm's participation in both service lines, even if the individuals themselves perform only audit or consulting services. Fischer and Huddart [2008] analytically model the effect of firm level norms on individual decisions consistent with the concerns raised by PCAOB members. We specifically examine individuals that provide both audit and consulting services. As a result, our experimental design provides a relatively strong test of the theory such that if such a spillover in norms exists, we are likely to find it. Correspondingly, if we do not find evidence of consulting norms spilling over onto the audit task, it would seem unlikely such an effect exists when different personnel provide each service.

While in many cases Big 4 personnel work only on audit or consulting, the firms have individual auditors who make both audit and consulting decisions. First, accounting firms allow staff to rotate or transfer between service lines. The firms use the opportunity to transfer into consulting as a selling point in order to recruit top students into the audit practice. Second, anecdotal discussions with recent graduates indicate within at least one Big 4 firm, opportunities exist for new hires to rotate nine months in auditing and three months in nonaudit services for their first two years with the firm, with a guaranteed position in nonaudit services upon promotion to senior associate. Third, professionals that specialize in consulting also participate on audit engagements as specialists. Firms emphasize this cross-participation when arguing for the benefits of their consulting practice growth. Finally, for clients receiving both audit and nonaudit services, one

partner is generally in charge of the entire client relationship, which includes both audit and consulting staff.

In addition to individuals participating in both audit and consulting projects, the interactions of professionals within an organization can create common social norms across service lines. As individuals interact, they observe each other's behavior and infer the key traits that lead to success within the organization. Social norms are established and reinforced as individuals conform to peer behavior and the influence of leadership. As such, while many professionals are involved in both audit and consulting services, even professionals that provide only one type of service can be influenced by the other service line's norms (Fischer and Huddart [2008]. While audit firm culture can positively reinforce auditor professionalism, as a firm's culture changes from emphasizing public duty to client advocacy, these commercial norms can erode the auditor's professionalism thereby reducing audit quality (Johnstone et al. [2001]).

Most nonaudit service research focuses on the impact on audit quality of auditors providing nonaudit services to audit clients. The literature examines whether providing nonaudit services improves audit quality due to a knowledge spillover effect where nonaudit services provide information that enhances audit quality (Simunic [1984]) or whether nonaudit services create an economic bond between the auditor and client that impairs independence, resulting in reduced audit quality (DeAngelo [1981a], Ashbaugh [2004]). While researchers have not found a consistent significant association between nonaudit service fees and audit quality proxies, the lack of an association does not prove auditor independence is not affected (Ashbaugh et al. [2003], DeFond et al. [2002], Myers et al. [2003], Geiger & Rama [2003], Schneider et al. [2006], Church et al.

[2015]). Prior studies, however, consistently demonstrate that investors and audit litigants perceive auditor independence as impaired when auditors provide nonaudit services (Krishnan et al. [2005], Schmidt [2012]). Further, experimental evidence consistently finds that auditors favor client-preferred positions when there is accounting ambiguity and auditor alignment with management (Church et al. [2015], Kadous et al. [2003], Mayhew et al. [2001], Mayhew & Pike [2004], Hurley et al. [2016]).

Contemporaneous research by Lisic et al. [2017] examines the relationship between consulting services and audit quality. They study the relation between the firms' self-reported proportion of annual consulting services to total firm revenue and audit quality measured by restatements. They find that higher consulting revenue is generally not associated with restatements although there is some evidence of such an association pre-SOX. Additional findings suggest investors perceive a deterioration in audit quality associated with more firm consulting, as reflected in the earnings response coefficient.

Our experimental approach allows a direct test of theory with clearly articulated constructs. Unlike an archival approach, our experiment controls the environment, directly measures the variables of interest, and controls confounding factors allowing us to test and speak to causality. Markets enable us to tease apart the effects of social norms and economic bonding to make clear inferences about the potential biases resulting from auditors providing both consulting and audit services.

Dopuch and King [1991] use an experimental market to explore the effect of management advisory services on auditor's independence. They compare two institutions – one in which the auditor can provide both consulting and auditing services and one in which the relationship is restricted. They find that prohibiting management advisory

services on an audit client does not improve market efficiencies and claim no effect on audit quality. However, they assume that the advisory service increases the likelihood the manager obtains a high value asset. As a result, the experiment design reduces the verification effort necessary to achieve a given level of audit quality when consulting services are provided. While auditors in the unrestricted condition were as accurate as the other session, they performed less verification testing consistent with nonaudit services lowering audit quality. We remove this confounding knowledge spillover effect to focus on the influence of consulting services on audit quality.

III. Experimental Design

We use an abstract market setting with student participants.[10] Our treatments directly manipulate participant norms, as students are unlikely to enter the market with strong social norms for the setting. Student subjects are particularly appropriate for experiments that focus on responses to institutions and forces expected to be learned during the experiment (Libby et al. [2002]).

Individuals often rely on social norms to inform their understanding of social situations and choose an effective response (Cialdini & Goldstein [2004]). We manipulate social norms through institutional rules of the experimental markets. Our audit market includes a cue that quality auditing (i.e. high-level verification) is the appropriate choice across all conditions. In the nonaudit service (NAS) treatments, we introduce a consulting project that strengthens the salience of a competing social norm – to cooperate with managers. We examine the effect of this competing social norm to test our theoretical predictions, as discussed in section IV.

[10] Participants are undergraduate and graduate students from a large Midwestern public university.

An experimental market allows us to analyze the influence of different institutional rules on otherwise unobservable individual actions in a controlled setting (Callahan et al. [2006]). We employ a 1 X 3 design in which participants complete multiple periods of an experimental market.[11] We manipulate the provision of nonaudit services between sessions across three levels: none (*Baseline*), *NAS-same,* and *NAS-separate*. *Baseline* consists of an audit market only, while the *NAS* conditions include a second market representing consulting services. In *NAS-same,* auditors are hired by one manager each period to provide both audit and consulting services. In *NAS-separate,* there are two groups of managers. One group hires auditors to provide audit services in the audit market only, and the other manager group hires auditors to provide consulting services in the consulting market only. The *NAS-separate* treatment contains the same social bonding as *NAS-same*, but does not include the same economic bonding. We use the generic terms provider, decision maker, and bidder in the experimental materials to reduce the risk that participants bring uncontrolled prior beliefs into the experiment (Swieringa & Weick [1982]). However, we use the terms auditor-consultant, manager, and investor throughout this paper for ease of understanding.

Each session includes ten to thirteen participants, depending on condition, in the roles of auditor-consultants, managers, and investors. *Baseline* and *NAS-same* include ten participants: three auditor-consultants, three managers, and four investors. *NAS-separate* includes a second group of three managers for a total of six managers with each manager group participating in either the audit market or the consulting market only. Each experimental session begins with an audio file of the experimenter reading the

[11] The experimental markets are programmed and conducted with z-Tree software (Fischbacher 2007).

instructions to all participants as they follow along on a printed copy they may reference throughout the session. Participants individually complete a pre-experiment quiz to facilitate understanding of the market(s). The quiz is administered electronically and provides immediate feedback regarding the correct answer and the rationale or calculation behind it. Prior studies with similar audit market designs suggest participants sufficiently understand the market given the instructions and following a few rounds of experience (Mayhew [2001]). Experimental dollars are used throughout each session and are converted to US dollars at session end.

The *NAS* conditions consist of at least 20 periods with each period including a consulting market and an audit market, with the order of the markets counterbalanced across sessions within each condition. *Baseline* consists of at least 20 periods of only the audit market. Subjects are not told the total number of periods in the session to reduce end-period effects and backward induction (Mayhew [2001]).[12] We conduct eight sessions of each condition. [13]

In the consulting market, the following actions occur:

1. Each manager chooses an auditor-consultant to hire for the consulting project. The manager pays the selected auditor-consultant E$100. In *NAS-same*, the hired auditor-consultant provides both the consulting and auditing services.

[12] We presequence the number of periods in each session consistent with Mayhew and Pike [2004] and King [1996] and hold the distribution of periods constant across conditions.

[13] We conduct a power analysis to assess the appropriateness of our sample size and likelihood of achieving our research objective. We set alpha, the upper bound for a type I error, at 0.10 consistent with a 0.05 significance level one-tailed hypothesis test. We set our sample size to 8 observations per condition, which is the number of sessions for which we reasonably expect to be able to recruit subjects. We solve for the smallest effect size we would have a 0.80 chance of detecting (i.e., power) assuming a standard deviation of 0.15, consistent with prior research (Hurley et al. [2016]) and our observed *Baseline* results. We have sufficient power to detect a difference in audit quality of 0.20 between *Baseline* and *NAS-Same* or *NAS-Separate*. Prior research (Mayhew et al. [2001]; Mayhew and Pike [2004]; and Hurley et al. [2016]) suggests this is a reasonable effect size to expect in this market. We note our actual results do not follow our directional predictions. As such, no change in sample size would provide sufficient power to detect the predicted effect.

2. Hired auditor-consultants choose whether to perform service X or service Y. Service X costs the auditor-consultant E$50 and produces E$150 for the manager. Service Y costs the auditor-consultant E$10 and produces E$100 for the manager. *The differential project service costs proxy for consulting effort. The project service cost and payoffs allow auditor-consultants to reciprocate a client's hiring decision by cooperating with management.*
3. Managers and auditor-consultants learn their consulting project earnings.

The audit market is derived from Mayhew and Pike [2004] and Hurley et al. [2016]:

1. Each manager hires an auditor. In *Baseline* and *NAS-separate*, any auditor can be hired. In *NAS-same*, the auditor-consultant hired for the consulting project remains the service provider for the audit market. Audit fees are paid by the manager to the auditor and are fixed at E$100. *These rules are a strong manipulation of institutional rules governing the provision of nonaudit services. NAS-separate is an operationalization of consulting services permitted for nonaudit clients that will never become audit clients. NAS-same is an operationalization of consulting services permitted for audit clients where we require the manager to hire the auditor for consulting work to achieve a cleaner manipulation. Auditor-consultants can be hired by more than one manager.*
2. Each manager privately selects an investment (A or B). Investment A costs E$225 and 60% of the time yields a high-valued asset (E$800, E$1000, or $1200 with equal likelihood), and 40% of the time yields a low-valued asset (E$200, E$400, or E$600 with equal likelihood). Investment B costs E$50 and always yields a low-valued asset (E$200, E$400, or E$600 with equal likelihood). *The manager's investment choice creates an information asymmetry that impedes the investor from focusing on simple expected values, thereby creating a role for the auditor to lend credibility to the manager's reports. Note that unlike Dopuch and King [1991] the consulting project has no impact on the probabilities of the asset yields from these investments.*
3. Managers privately learn the asset value generated by their investment strategy and issue an asset value report to investors. Managers can overstate their asset's value if the asset is worth less than E$1,200. *The ability of managers to misreport creates a role for auditors. There is no rational incentive to underreport.*
4. Each auditor privately learns whether s/he is hired and chooses between high- or low-level verification. If an auditor chooses high-level verification, s/he learns the asset's true value with a probability of .95, and is given the managers' reported value with a probability of .05. If an auditor-consultant chooses low-level verification, s/he learns no new information and must agree with the manager's report. High (low) verification service costs the auditor E$50 (E$10). For each asset, there is a 25 percent chance the auditor's verification choice will be checked. If checked and found to have chosen low-level verification, the auditor incurs an incremental cost of E$90. The auditor does not learn whether they are checked until the end of the period. *The verification choice represents audit quality and serves as our dependent variable. The differential verification service costs proxy for audit effort and create moral hazard for the auditor. The low verification cost represents the minimum effort level all auditors must perform.*

The probability that the auditor identifies the true asset value through high-level verification proxies for accounting and auditing uncertainty. The probability that the auditor's verification decision will be checked proxies for regulatory oversight (e.g. PCAOB inspections). The verification check also communicates high-level verification is expected.

5. Each auditor who chooses high-level verification privately learns the audit value of the asset. If the results match the manager's reported value, the auditor issues an Agree Report. If the results do not match, the auditor issues a Disagree Report. Each auditor who chooses low-level verification issues an Agree Report. *This is not a choice; the auditor's report must be consistent with verification results.*[14]
6. Investors receive the following information regarding each manager's asset: the manager's reported asset value, the auditor's report, the auditor's identification number, and the auditor's history of accurate audit market reports and failed checks. Investors engage in a first-price sealed-bid auction, with each asset's high bidder paying their bid and receiving the true asset value. All investors learn whether they are the high bidder for each asset.
7. All participants receive an earnings summary. Participants stay in the same role for the entire market. Manager identities to investors are scrambled, while auditor identities remain the same. *Manager identities are scrambled because we are not studying manager reputations.*

While information asymmetry exists during each period, results other than investment choice become common knowledge before beginning the next period to facilitate information aggregation and strategy selection. All participants receive a market history through the previous period during the experiment. This history includes the manager's reported asset value, the auditor's report, the true asset value, and the high bidder's earnings from purchasing the asset. It also includes information about the auditors' history of being hired and issuing accurate reports. The consulting market history includes the auditor-consultants' history of being hired and providing high-value project service. This approach creates strong incentives for the auditor-consultant to cooperate when providing consulting services. This information is likely more precise

[14] It would be irrational to perform a high-level verification and issue a report that contradicts the results. If an auditor intends to agree with the manager's report, choosing low-level verification is less costly, even after considering the expected cost of the verification check (expected earnings of E$67 versus E$50 for selecting high-level verification).

and timely than information provided in a natural setting thereby enhancing the role reputation plays in the market (Mayhew [2001]).

Auditor-consultants earn the sum of their consulting fees less project service costs and audit fees less verification costs. Managers earn the sum of their consulting project payoff less consulting fees and the high bid for their asset less investment costs and audit fees. For winning bids, investors earn the true asset value less their bid. Investors also earn a ten percent bonus on the cumulative value of the assets they acquire throughout the market.[15] Experimental sessions are expected to last up to 120 minutes. Conversion factors are set so that participants earn approximately $15 per hour, including a $7 show-up fee. We vary conversion rates across roles and conditions to give participants similar opportunities to earn U.S. dollars based on predicted time spent in the lab. Participants complete a post-experimental questionnaire at the end of the market.

IV. Game theoretic equilibriums and predictions

We outline the game theoretic equilibriums for our setting to establish predictions of behavior. These game theoretic predictions are not affected by our treatments. We expand the discussion with insights from theories of how changes in social norms will impact behavior to derive our hypotheses.

Consulting market

The consulting market is a modified trust game, which provides a simple way to model the principal-agent problem of an employment relationship. A consulting project is similar to an employment relationship. Management hires a consultant to perform a

[15] Participants are aware of the asset return, but are unaware of the return percentage. The purpose of the return is to provide motivation to purchase assets that, in a competitive equilibrium, provide zero return. It also leads to differences in private expectations, reducing the common value nature of the auction.

specific project. The consultant chooses how much effort to put into the project, which influences the project's value to the manager. The manager uses this information to decide whether to hire the consultant again in future periods.

In our consulting market, the manager decides which auditor-consultant to trust with a payment of E$100. If the auditor-consultant responds cooperatively by choosing high effort at a high cost, each earns a E$50 profit. If the auditor-consultant shirks by choosing low effort, the auditor-consultant benefits in the current period, earning E$90, while the manager earns nothing. The single period Nash equilibrium predicts the auditor-consultant shirks and takes the higher profit. However, in a repeated game, current behavior influences others' future actions (Watson [2008]). The multiple periods, along with an uncertain end-period, introduces a reputation mechanism supporting cooperation. In our setting, the cooperative auditor-consultant builds a reputation that increases likelihood of being hired in a future period. A manager is likely to replace a shirking auditor-consultant with a cooperative auditor-consultant in future periods. All market participants observe the outcome of the auditor-consultant's actions after the period ends. This information, along with competition, makes the reputation mechanism a strong cooperation motivator for the auditor-consultant. Prior research also suggests that reciprocity concerns provide a strong cooperation social norm (Gouldner [1960], Gachter & Falk [2002], Andreoni & Miller [1993], Dufwenberg & Kirchsteiger [2004]).

We expect managers to hire cooperative auditor-consultants for the consulting project. We expect auditor-consultants to choose a cooperation strategy on the consulting project. This strategy is consistent with concern for reciprocity as well as rational reputation concerns, as the auditor-consultants' incentive to earn repeated, moderate

profits outweighs the incentive to take a single period high profit. We are not interested in studying behavior in the consulting project, per se. We use it to establish client cooperation as a social norm.

Audit Market

The audit market is a more complicated game. After the manager decides which auditor to hire, the manager and auditor make simultaneous decisions in a manner similar to a prisoner's dilemma, with the added role of the investor player affecting payoffs. The game-theoretic equilibriums for the audit market, as described in prior research (Mayhew, Schatzberg, and Sevcik [2001], Mayhew [2001], Hurley et al. [2016], Hurley & Mayhew [2017]), are the Lemons and Reputation equilibriums.[16]

The Lemons equilibrium is the unique single period equilibrium of the game. In a single period game, the auditor shirks when hired to save investigation costs. Accordingly, investors know they cannot trust the auditor's report. Without a reliable audit report, managers cannot credibly report a high asset value, so managers cannot benefit from a high investment. As a result, the manager chooses the low investment and reports a high asset value for the resulting low value asset, because s/he knows the investors will ignore the manager and auditor reports and bid the low asset value. This Lemons equilibrium remains an equilibrium in the repeated game.

In a repeated game, auditors can build a reputation for high audit quality enabling the formation of a Reputation equilibrium where auditors choose high-level verification to build their high audit quality reputation, managers invest high, and investors

[16] The consulting and audit market are similar in that the manger hires the auditor-consultant in the first stage. The markets differ in that the auditor-consultant makes a sequential decision in the consulting market with clear knowledge of how it will benefit or harm the manager. In the audit market, the two players make simultaneous decisions that impact each other's outcomes.

incorporate the managers' report, the auditor report and the auditor's reputation for accurate reporting into their bids on the resulting assets. This equilibrium provides higher returns to the managers than the Lemons equilibrium because the auditor reduces the information asymmetry around whether the manager is choosing the high investment option, and accordingly increases the investors' bids for assets with high reported values and corresponding auditor reports.

Auditors who choose high-level verification will be accurate 95% of the time or more. Auditors who choose low-level verification will only be accurate when managers report honestly.[17] In the reputation equilibrium, managers hire reputable auditors and make the high investment. The high investment produces a high-value asset 60% of the time making it worth the manager's higher cost investment as long as investors trust the joint manager/auditor report. While managers prefer the high-quality auditor, the manager will still report low value assets as high value to investors, even though a reputable auditor will correctly disagree with the manager's report 95% of the time. To see why, consider an example. When investors observe a 200-value asset report, they will bid 200 as they know managers cannot underreport. If the manager reports that same 200-value asset as 1,200, even the highest quality auditor in our setting will be wrong 5% of the time and agree with that report. By misreporting a 200-value asset as a higher value, the manager has at least a 5% chance of receiving higher bids from investors, increasing the manager's expected return. As such, an auditor cannot rely on a manager's honest reporting to maintain their reputation for accurate auditing. The Reputation equilibrium is

[17] Verification checks provide additional information regarding the auditor's reputation. Twenty-five percent of audits are checked and results are publicly revealed. The auditor's rate of failed checks reveal information about audit effort that are otherwise inferred from the auditor's accuracy rate.

not sustainable without the auditor choosing high-level verification. The Reputation and Lemons equilibriums represent two end points in a multi-period setting, with the Reputation equilibrium generating more overall market surplus than the Lemons equilibrium, on average. Nonetheless, the multi-period nature of the game allows for numerous equilibria to evolve. In our discussion of expected behavior, we incorporate social norms into the game. We expect norms to influence which equilibrium participants play. Prior research suggests the Reputation equilibrium is difficult to achieve when there is accounting uncertainty (Mayhew et al. [2001]), and is influenced by institutional features that draw the auditor's attention toward investors or managers (Mayhew and Pike [2004], Hurley et al. [2016]).

Auditor Behavior

For each asset they are hired to verify by managers, auditors choose high- or low-level verification. While investors' interests are best protected by high level verification, we expect auditors to respond to the managers' demand for audit quality as managers provide direct auditor incentives via hiring decisions. A manager playing a reputation strategy will prefer an auditor that performs a high-level verification, while a manager playing a lemons strategy will prefer an auditor who shirks. Prior research suggests managers often prefer auditors who provide less than the highest level of quality (Hurley & Mayhew [2017]).

This audit market incorporates an auditor verification check to signal the social norm of protecting investors. Twenty-five percent of auditor verification choices are checked. If an auditor is checked and has provided a low-level verification, the auditor incurs an incremental cost of E$90, reducing their profit for that asset verification service

to E$0. The verification check does not change the game theoretic predictions of the market; it merely narrows the auditor-consultant's opportunity cost. This regulation feature captures the key construct of process-oriented regulator oversight of the audit profession (e.g. PCOAB inspections). It more importantly signals that performing a high-level verification is the appropriate expected behavior in this situation, introducing a social norm of high-quality auditing.

Individuals often rely on social norms to inform their understanding of social situations and choose an effective response (Cialdini & Goldstein [2004]). Norms provide rules or standards that guide the behavior of group members (Cialdini and Trost [1998]). Descriptive norms inform us about what is typically done by others, while injunctive norms inform us about what is typically approved and disapproved (Cialdini et al. [1991]). These injunctive norms serve as moral rules and motivate behavior through social rewards and punishments (e.g. future hiring opportunities). When making decisions, in addition to financial incentives, individuals face an emotional cost determined by behavioral norms, with non-compliance being more costly (Fischer & Huddart [2008]). Formal controls can influence behavioral norms (Tayler and Bloomfield [2011]). The institutional rules of our markets convey the injunctive social norms of appropriate behavior, while also allowing the participants to observe the descriptive norms of other participants' behavior. The verification check in the audit market is an implicit request to the auditors to perform high-level verifications, communicating the norm of high-quality auditing. In contrast, the manager's hiring decision – in both the consulting market and the audit market - is a request for reciprocation and cooperation.

Auditor cooperation in the audit market is dictated by the manager's preferences. Prior research using similar audit markets to ours show managers do not strictly prefer high quality auditors (Mayhew et al. [2001], Mayhew and Pike [2004]). Recent research suggests that managers will even forgo inserted high-quality auditors for auditors who will regularly supply low quality and just agree with manager's reports (Hurley and Mayhew [2017]). To the extent providing consulting services increases the auditor's client-cooperating focus, we will observe more client-biased auditor reports.

In *NAS-same*, the auditor-consultant earns an additional stream of revenue from the same manager by cooperating on a consulting project. This revenue stream has clear, aligned incentives for the manager and auditor-consultant. The auditor-consultant is hired to serve the interests of the manager; the manager prefers a cooperative auditor-consultant to maximize his/her project value and the auditor-consultant prefers to cooperate to get hired repeatedly. We predict that the additional revenue stream from the consulting project will create an economic bond between the manager and auditor-consultant that will influence the auditor's behavior in the audit market. The provision of both audit and consulting services to the same client obscures the auditor's perception of his role when providing audit services (Bazerman et al. [1997]). The auditor-consultant will be more likely to cooperate with the manager in the audit market and less likely to consider investor needs. When the manager prefers lower audit quality, auditors will cooperate by performing fewer verifications. Based on this discussion we predict the following hypothesis:

H1: Auditors hired for audit and consulting services by the same manager will perform lower quality audits than auditors hired only for audit services (*Baseline* > *NAS-same*).

This hypothesis reflects the motivation behind regulatory restrictions placed on permissible nonaudit services: that nonaudit services create an economic bond between the auditor and manager, impair independence, and lead to reduced audit quality. H1 assumes managers prefer lower audit quality. If a manager prefers higher audit quality, economic bonding would lead to high audit quality instead. Our market parameters allow for the economic bonding theory to manifest. We do not allow the alternative off-setting knowledge spillover used by Dopuch and King [1991].

In *NAS-separate*, the auditor-consultant also earns a separate revenue stream by cooperating with a manager on a consulting project. However, the revenue earned by cooperating with a manager on a consulting project is from a completely separate manager group than the revenue earned from providing audit services. One manager group participates in the audit market only, hiring auditor-consultants to provide audit services. A separate manager group participates in the consulting market only, hiring auditor-consultants to provide consulting services. The *NAS-separate* condition completely removes the economic bond from consulting that exists in *NAS-same* leaving only the cooperation norm. The market restriction also signals that audit and consulting services should be considered separately. The *NAS-separate* auditor-consultants face the same economic incentives as the *Baseline* auditors in the audit market. Based on economic incentives alone we would predict no difference in audit quality between the two conditions. However, cooperating with other managers on the consulting project

strengthens the salience of a cooperation norm. We test whether this norm carries over into the audit market and influences the auditor's behavior. We predict the cooperation norm from the consulting market impairs audit quality in the audit market absent economic bonding incentives.

> **H2:** Auditors participating in a separate consulting market will perform lower quality audits than auditors hired only for audit services (*Baseline* > *NAS-separate*).

Our prediction that *NAS-separate* auditors will provide lower audit quality than *Baseline* auditors is based on the expected effect of the cooperation norm from providing consulting services to any manager. Our prediction that *NAS-same* auditors will provide lower audit quality than *Baseline* auditors is based on the combined effect of the cooperation norm and the economic bond to the specific manager by whom they are hired. We predict the combination of the cooperation norm with an economic bond will have a stronger influence on audit quality than the cooperation norm alone.

> **H3:** Auditors hired for audit and consulting services by the same manager will perform lower quality audits than auditors participating in separate audit and consulting markets. (*NAS-separate* > *NAS-same*).

V. Results

Planned main analysis

We measure audit quality using the auditor's verification choice, with a high-level verification coded as high audit quality. An experimental market allows us to focus our measure of audit quality on the auditor's input rather than the joint output of manager and auditor behavior frequently used in archival studies.

Table 1 presents measured variables by condition and experimental session sorted by *AuditQuality*. On average, auditors investigate 66% of the time in *Baseline*, 62% of the time in *NAS-separate* and 71% of the time in *NAS-same*. While the differences in average audit quality across conditions are small, audit quality ranges from 38-87% in *Baseline*, 22-87% in *NAS-separate*, and 11-100% in *NAS-same*.

To test our hypotheses, we use a generalized linear mixed effects model to run a logit regression estimating our dependent variable, audit quality. We regress audit quality on planned contrast coded predictor variables that capture our hypothesized comparisons between conditions. We use by-subject and by-session random intercepts to control for non-independence.[18] Table 2 documents the results of our hypothesis tests. The overall analysis of variance suggests there are no meaningful differences across the three sessions. Our planned contrast shows H1 is not supported. There is not a significant difference in audit quality between *Baseline* and *NAS-same*. Furthermore, the predicted pattern of results where *Baseline* > *NAS-separate* > *NAS-same* is not supported. In fact, we find the rather surprising result that *NAS-same* has the highest level of audit quality on average. Additional planned contrasts comparing *NAS-separate* to *Baseline* and to *NAS-same* are both insignificant suggesting a lack of support for either H2 or H3.[19] There are no differences in audit quality between any of the conditions, contrary to our hypotheses.

INSERT TABLES 1, 2 AND FIGURE 2 HERE

[18] We limit analyses to periods 6-20 from each session. We drop the first five periods to remove learning effects. We drop all periods beyond the first 20 to reduce end-period effects.
[19] A significant difference between *Baseline* and *NAS-separate* would have provided evidence that a cooperation norm alone is sufficient to impair audit quality even when no economic bond exists (H2), and a difference between *NAS-separate* and *NAS-same* would have suggested an economic bond incrementally reduces quality (H3).

Our design assumes auditor-consultants will cooperate in the consulting portion of the market. Actual consulting cooperation levels vary across sessions. Widespread defection from consulting market cooperation in a session would make testing the effects of cooperation difficult. The realized levels of cooperation – 90% on average - are generally consistent with our expectations, suggesting we successfully introduced a cooperation norm. However, we note two of 16 *NAS* sessions with lower than expected cooperation. To test our results sensitivity, we drop sessions with low cooperation and run our planned analyses with only those sessions where the consulting market established cooperation. We continue to find no difference in audit quality between conditions.[20]

Unplanned analysis

Although our planned analysis finds no statistical differences in audit quality by condition, we use unplanned analyses to explore the deviation from our predictions. Visual inspection of the results in figure 2 shows greater dispersion in audit quality in the *NAS* conditions than *Baseline*; a statistical comparison of the standard deviations in Table 2 confirms this difference is marginally significant. We conduct a variance ratio test and find a marginally significant difference between the standard deviation of the *NAS* sessions versus *Baseline* ($p = 0.09$). This difference is primarily driven by the contrast between *NAS-same* and *Baseline* ($p = 0.07$). This evidence suggests that the provision of consulting services leads to more variation in audit quality.

[20] We consider sessions that cooperate less than half of the time (50%) as non-cooperative. One session in *NAS-separate* fell below this threshold with 33% cooperation. We also test a cooperation cut-off of 75%, the midpoint between full cooperation and non-cooperation. One session in *NAS-same* fell below this threshold with 69% cooperation. Average cooperation after excluding these non-cooperative sessions is 96%.

Our hypothesis development argues that providing consulting services increases the auditor's client-cooperating focus, leading to more client-biased auditor reports. We assert that auditor cooperation in the audit market is dictated by the managers' preferences. While we expected cooperation to manifest as lower audit quality, we acknowledge it could result in higher audit quality if managers prefer high quality.[21] In addition to actual manager preferences, audit quality can be influenced by auditor perception of manager preferences or norms. The *NAS* conditions include sessions with very high and low audit quality.

To further understand manager preferences, we inspect audit quality and market share for each auditor by session in table 3.[22] In half of the *NAS* sessions, a high-quality auditor dominated the market based on the auditor choosing high quality > 90% of the time and garnering the greatest market share. Four out of the eight sessions each of *NAS-same* and *NAS-separate* meet this 'high quality' definition (HQ), while no *Baseline* sessions meet it. This evidence suggests some managers demanded high audit quality and auditors responded by providing it.

INSERT TABLE 3

In an unplanned analysis documented in table 4, we find a significant difference in audit quality between *NAS* sessions meeting the HQ definition (*NAS HQ)* and the remaining *NAS* sessions (*NAS non-HQ).* Further, audit quality in *Baseline* is significantly

[21] Based on prior studies, we expected managers to prefer less than high quality auditing, however, the game-theoretic equilibrium includes two endpoints: Lemons and Reputation. While the Reputation equilibrium is difficult to achieve, the parameters of our study gave it a good opportunity to arise: accounting and auditing uncertainty was low (high-verifications were 95% accurate) and verification checks signaled a norm of high audit quality.

[22] Market share equals the number of times the auditor is hired divided by the total number of hiring opportunities in the session (i.e. 3 manager hires per period x total number of periods in the session).

lower than *NAS HQ* and higher than *NAS non-HQ*.[23] This pattern of differences suggests that consulting services do not lead to uniformly higher or lower audit quality, but to more extremes. In sessions where managers prefer high audit quality to add credibility to their reports, a cooperative auditor performs high-level verification. In sessions where managers prefer lower audit quality, a cooperative auditor performs more low-level verifications. Consulting services appears to increase auditor cooperation by magnifying both high and low quality in the *NAS* treatments relative to *Baseline*.

INSERT TABLE 4 AND FIGURE 3

Supplemental analysis

While we are primarily interested in studying auditor behavior, the dynamic nature of the interaction between the market participants provides an opportunity to analyze manager and investor behavior. We expect manager and investor behavior generally to follow prior research (Hurley et al. [2016]), and as such, make no formal predictions about the influence of the treatment conditions.

We look at manager investment and reporting choices in a planned analysis in order to confirm our game theoretic expectations about manager strategies. Additionally, we explore the data in an unplanned analysis to further understand apparent differences in manager preferences between the *NAS HQ* and *NAS non-HQ* sessions.

Manager preferences are reflected in both their auditor selections and their investing behavior. The manager can choose whether to play a reputation or lemons strategy (Mayhew [2001]). Prior research suggests auditors respond to manager's

[23] We perform sensitivity analyses excluding the non-cooperative consulting sessions and the HQ session with 100% high-level verifications, as managers did not have a choice in audit quality. Our results are consistent: *Baseline vs NAS HQ* ($p = 0.02$), *Baseline vs. NAS non-HQ* ($p = 0.05$).

preferences for audit quality (Mayhew and Pike [2004], Hurley and Mayhew [2017]). We expect that in general, managers will demand a moderate level of audit quality, but that the more auditors acquiesce to the managers' preference for less than high audit quality, the further managers' will deviate from a reputation strategy. In contrast, we expect that managers' investment levels will increase and misreporting will decrease as audit quality increases.[24] In a planned analysis in table 5, we regress managers' investing decisions (i.e. propensity to invest high) and misreporting decisions on our measure of audit quality. We find that managers who hire more accurate auditors are more likely to invest and less likely to misreport, consistent with the predicted reputation or lemons strategy.

INSERT TABLE 5

We conduct an unplanned analysis of manager investing behavior by HQ designation in table 6. We expect high investment in the *NAS HQ* sessions given a high-quality auditor dominates the market, and lower investment in *NAS non-HQ* sessions. Our results are directionally consistent with this expectation as there is more investment in the NAS HQ sessions, but this higher investment is not statistically different from *Baseline* and *NAS non-HQ* levels. Together these analyses provide evidence that manager preferences varied, while remaining consistent with game theoretic predictions that managers will follow either a lemons or reputation equilibrium strategy.

INSERT TABLE 6

We conduct an unplanned analysis of auditor choices in the first five periods to further understand auditors' perception of manager preferences or norms. Behavior in the

[24] As discussed in section IV, a reputation strategy does not require managers to be honest. Managers always have a weak preference to lie when they receive a low-value asset. The decrease in misreporting follows the increase in investment.

early periods can reveal insights about auditors' initial perceptions. Audit quality provided in early rounds can influence managers' eventual preferences as they realize the consequences of playing different strategies. Figure 4 shows *NAS* auditors started with higher audit quality than *Baseline* auditors. *NAS* auditors delivered 75% high audit quality in the first five periods, significantly higher than *Baseline* audit quality of 58% ($p = 0.04$). There was no difference between *NAS-same* and *NAS-separate* or between *NAS HQ* and *NAS non-HQ*. *Baseline* auditors provide a relatively constant level of audit quality throughout the session, while *NAS* auditors tend to sharply increase or decrease audit quality after the first five periods. The higher audit quality in early periods of the *NAS* markets suggests auditors providing consulting services perceive a norm of high effort and quality before they learn manager preferences.

INSERT FIGURE 4

We look at investors' bids on each asset to assess the functioning of the audit markets. The auditor's credibility drives the reliability of reported asset values greater than 200 accompanied by an agree report.[25] The investor does not know if the auditor performed a high-level verification or merely agreed with the manager's reported value. The auditor's historical accuracy serves as a signal of how much assurance investors can expect on the reported asset value. As such, we expect investors to discount agree reports from less reputable (i.e. less accurate) auditors and submit lower bids. In analyzing investor behavior, our tests focus on how they react to auditors' agree reports. We calculate investors' confidence in the report as the difference between the high bid and

[25] A disagree report from the auditor is a clear indication that the asset is worth less than its reported value. A reported asset value of 200 is always accurate because managers cannot understate asset values.

the asset's reported value. We regress this difference on auditor accuracy and the manager's report.[26] Table 7 documents the full model we employ along with the results. We find a significant main effect of auditor accuracy suggesting investors are more confident in asset values when auditors are more accurate. We also find a negative coefficient on manager reports consistent with investors not believing higher manager reports in general. The interaction between the manager's reported asset value and auditor accuracy shows auditor accuracy matters more as managers' reported asset values increase. That is, the increase in investor confidence from auditor accuracy is greater at higher reported asset values. We also include our treatment variables to assess whether investors perceive consulting as an incremental threat to reporting quality and find no such effect. In general, investor behavior is consistent with our expectations and past market studies, suggesting the overall market dynamics function as intended.

INSERT TABLE 7

Post-experiment questionnaire

Our main analyses stand on their own as tests of our theory. Generally, deviations from game-theoretic predictions of self-interested behavior serve as evidence of the social norms we are testing (Camerer and Fehr [2004]). However, we also asked participants about their strategies ex post. Their responses provide additional nuance to understanding the impact of consulting on audit markets. Our post-experiment questionnaire provides supplemental information concerning the participants' strategies and perceptions of social norms in the markets. Appendix A lists the questions, which include both scale-based, and free response questions. To measure the auditor's

[26] We run the same analysis with audit quality as measured in our main analysis and find similar results although somewhat weaker.

perception of the social norms in the audit market, we ask questions about what they believe is an appropriate action for the situation, what they believe is expected of them, what they feel obligated to do, and what they expect other auditors to do (Bicchieri and Chavez [2010]). Participant self-awareness in these types of experiments is limited. Accordingly, we consider their self-reported perceptions and interpretation as supplementary to our main tests. Appendix A presents mean responses to each questionnaire item by condition (planned analysis) as well as by HQ designation (unplanned analysis). We inspect responses to the post-experiment questionnaire to corroborate our interpretation of the pattern of results. No clear pattern emerges by condition, consistent with the lack of results in our planned analysis. Next, we look at responses by HQ designation.

NAS auditor-consultants perceive a norm of cooperation in the consulting market, as indicated by their responses to questions P1-P4 being significantly above the midpoint of the scale ($p < 0.01$). While *NAS HQ* auditor ratings are directionally higher than *NAS non-HQ* auditors, the difference is not statistically significant for P1-P3 and only marginally significant for P4 ($p = 0.05$), suggesting auditors across the *NAS* conditions similarly perceived the norm of cooperation in the consulting market.

In the audit market, the perceived norm around the provision of verification services differed. *NAS HQ* auditors perceived the highest norm of verification, followed by *Baseline* auditors, then by *NAS non-HQ* auditors. Contrast analysis finds a significant difference between *NAS HQ* and *NAS non-HQ* mean responses to questions P5-P8, coupled with a lack of difference between *Baseline* and the average of *NAS HQ* and *non-*

HQ, consistent with the pattern of actual audit quality. P7 directionally follows the same pattern, but is not significant.

Auditors described their strategy and the factors they considered when making decisions in the market in open-ended responses. The most common factor cited in all conditions was to get hired.[27] Some auditors described their strategy as performing high verification to please managers and get hired more; others indicated they performed low verification services to please managers. Some anticipated managers would prefer one type of service and switched their strategy once they saw which auditors got rehired. This theme supports our interpretation that in general, auditors responded to their perception of managers' demanded level of audit quality. Other responses expressed a desire to maximize profit, smooth earnings, or follow others' behavior.

Manager's post-experiment responses confirm variation in audit quality preferences. We ask to what extent managers prefer an auditor with high accuracy (question M7). Contrast analysis finds a marginally significant difference between *NAS HQ* and *NAS non-HQ* coupled with a lack of difference between *Baseline* and the average of *NAS HQ* and *NAS non-HQ*. The pattern of manager preference for audit quality follows the pattern of realized audit quality. In their open-ended responses, some managers expressed a clear preference for high audit quality, while others expressed a preference for something other than high quality (i.e. either low or a mix). *NAS* managers appear to have been more successful demanding audit quality than *Baseline* managers as

[27] One of the authors reviewed auditors' open-ended responses blind to HQ designation, but not to condition, and coded each auditor response into one of the emergent themes: getting hired, maximizing earnings, smoothing earnings by avoiding verification checks, or other. Sixteen *Baseline*, 13 *NAS-separate*, and 17 *NAS-same* auditors (out of 24 in each condition) described their strategy with a focus on getting hired in the audit market.

their expressed preferences more closely aligned with the audit quality realized in their markets.[28] The similar variation in manager preferences by condition, coupled with the more extreme audit quality responses suggest that the cooperation in the consulting market amplified the auditor's tendency to respond to manager preferences.

V. Discussion

The audit quality pattern results together with manager preference indicators and the post-experiment questionnaire responses are consistent with our prediction that cooperating in a consulting market leads to cooperation in the audit market. Contrary to our prediction, cooperation in the audit market does not manifest itself simply as lower audit quality. Instead, it appears to amplify auditor responses to manager preferences, real or perceived. When managers prefer low audit quality, an auditor faces conflicting pressures. Verification checks signal that they should provide high audit quality, while the manager's hiring power incentivizes low quality. Manager pressure to provide low quality appears more salient when the auditor cooperates with managers in the consulting market compared to auditors who only experience the audit market. When managers prefer high quality, the norm implied by the verification check is aligned with manager preferences. The only remaining conflicting pressure is the incentive to increase profits through shirking. While the *Baseline* auditors generally respond by providing higher

[28] One of the authors reviewed managers' open-ended responses blind to HQ designation but not to condition and coded each manager response as having clearly expressed a preference for high audit quality, low audit quality, mixed quality, or unclear. Four managers expressed a clear preference for high audit quality in *NAS-same*, and all four were in HQ sessions. Four of the five managers preferring high quality in *NAS-separate* were in HQ sessions. In *Baseline*, four of five managers expressing a preference for high quality were in sessions with audit quality in the top half, but none were high enough to meet the definition of a "high quality" market. Of the managers expressing a preference for low or mixed audit quality, seven of nine in NAS-same were in non-HQ sessions, eight of 12 in NAS-separate were in non-HQ sessions, while the 12 in Baseline were evenly split between the higher and lower halves of the Baseline sessions.

audit quality when demanded, they also appear to sacrifice some quality to increase profits. A cooperative auditor-consultant in the *NAS* treatments appears to provide the demanded high audit quality consistent with a greater focus on manager preferences relative to profits.

A recent study of the effect of social ties between CEOs and audit engagement partners in South Korea finds a pattern of results consistent with our interpretation of our experimental market results (Kwon and Yi [2017]). Social ties could theoretically impair audit quality through collusion or improve audit quality through more efficient information sharing and trust. Kwon and Yi [2017] find that CEOs and engagement partners with social ties (i.e., school ties) are associated with high quality financial statements (i.e., absolute magnitude of performance-adjusted discretionary accruals). The effects are more pronounced among firms more likely to demand high audit quality (i.e., large, cross-listed firms) and disappear among those with weak governance.

VII. Conclusion

We examine whether providing nonaudit services to nonaudit clients impairs audit quality. Recent growth in large accounting firms' consulting practices has raised concerns that this evolving business model will lead to impairments in audit quality. We conduct an audit market experiment subject to a registration-based editorial process in which we vary the institutional rules governing nonaudit service relationships between auditors and clients in order to study the effect on auditor behavior. We introduce a cooperative consulting services task between manager and auditor participants. We predict that auditor engagement in this task will strengthen the salience of a cooperation norm with management that will carry over into the subsequent audit market and impair audit

quality. Our experimental approach allows us to study this effect without independence concerns or the potential benefits of knowledge spillover.

In our planned analysis we do not find that the provision of consulting services impairs audit quality uniformly; rather we find that it leads to greater variation between markets with higher high audit quality and lower low audit quality. We predicted cooperating in a consulting task would lead to greater cooperation on the audit task, and predicted cooperation to manifest itself as lower audit quality. However, in principle it could result in higher audit quality if that is what managers prefer. We find variation in manager preferences for audit quality and auditor perceptions of manager preferences. In unplanned analysis, we find half the NAS sessions result in high quality (HQ) markets, with managers most frequently hiring the highest quality auditor providing > 90% audit quality. These *NAS HQ* sessions have significantly higher audit quality than *Baseline* and the *NAS non-HQ* sessions. *NAS non-HQ* has significantly lower audit quality than *Baseline*. The results of our unplanned analyses, along with responses from the post-experiment questionnaire suggest that cooperating in the consulting market leads to cooperation in the audit market, with cooperation being dictated by manager preferences. In sessions where managers prefer high audit quality to add credibility to their reports, a cooperative auditor performs high-level verification. In sessions where managers prefer lower audit quality, a cooperative auditor performs more low-level verifications. Our unplanned analyses suggest cooperation on the consulting task amplifies auditor responses to managers' preferences.

Our study is subject to certain limitations. First, our experiment demonstrates the potential of institutional rules to influence behavior through norms, but does not provide

evidence that such norms have changed. It is difficult to measure the norm even in our experimental markets and it would be very difficult to examine the strength of existing social norms within the Big 4 accounting firms. Second, our experimental design does not capture all factors that are present in a natural audit market. We design it to capture key auditor incentives, but acknowledge it does not include all forces that influence an individual auditor's decisions. We focus instead on the theory that providing consulting will influence auditor behavior even when not provided to audit clients. Third, our experimental design permits auditor-consultants to compete on audit quality reputation alone. Price competition through bidding offers an additional mechanism through which to compete and specialize. We choose not to include auditor bidding, as it would incorporate additional complexity to the experiment without impacting our predictions.

Finally, the way we operationalize audit quality does not necessarily extend to other settings. Our experiment design operationalizes audit quality as an effort choice, provides unambiguous but imperfect audit evidence, and allows no reporting discretion. This approach is consistent with the nature of many audit tasks where an exception is indicative of a problem and the accounting is relatively objective. It is unclear whether our approach generalizes to more subjective accounting settings, such as highly uncertain estimates, where some contradictory evidence can be expected.[29] Similarly, settings where auditors have reporting discretion could also generate different results.

[29] Kachelmeier and Van Landuyt [2017] find social bonds impair audit quality when there is measurement uncertainty. Their operationalization of audit quality is an effort decision made after seeing audit evidence that can contain false positives, creating a scenario where auditors can give management the benefit of the doubt. Our results comparing *Baseline* and *NAS non-HQ* are consistent with their findings. Our results comparing *Baseline* and *NAS HQ* demonstrate a different effect when managers prefer high audit quality.

We believe controlled research on the influence of nonaudit services on audit quality provides useful insights to researchers, practitioners, and regulators on the potential threats to audit quality from the Big 4 firms' evolving business models. Future research can subject our interpretation of our unplanned analysis results to further scrutiny by replicating the findings in a controlled setting to establish a causal link between manager preferences and auditor responses. More importantly, because our results suggest auditors respond to manager preferences, it is important to understand when and why managers demand high versus low audit quality. Our study parameters create an environment, much like the real world, where high audit quality has the potential to add value to manager reports through rational investor responses, yet some managers continue to prefer a Pareto-inferior strategy of low quality audits and skeptical investors. Future research can expand our understanding of the demand for audit quality beyond economic incentives. A growing body of literature in behavioral economics provides evidence that human decision-making processes are not driven solely by economic cost-benefit consideration, but are influenced by psychological factors and biases (Kahneman and Tversky [1979]; Kahneman [2003]). Identifying the conditions that lead managers to prefer high versus low quality auditors will expand the theory of the demand for audit quality and provide direction to regulators and practitioners striving to improve audit quality.

REFERENCES

ANDREONI, J., and J.H. MILLER. "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence." *The Economic Journal* 103 (1993): 570-585.

ASHBAUGH, H. "Ethical Issues Related to the Provision of Audit and Non-Audit Services: Evidence from Academic Research." *Journal of Business Ethics* 52 (2004): 143–148.

ASHBAUGH, H., R. LAFOND, and B.W. MAYHEW. "Do Nonaudit Services Compromise Auditor Independence? Further Evidence." *The Accounting Review* 78 (2003): 611–639.

BAZERMAN, M.H., K.P. MORGAN, and G.F. LOEWENSTEIN. "The Impossibility of Auditor Independence." *Sloan Management Review* 38 (1997): 89-94.

BEASLEY, M.S., J.V. CARCELLO, D.R. HERMANSON, and T.L. NEAL. "The Audit Committee Oversight Process." *Contemporary Accounting Research* 26 (2009): 65-122.

BICCHIERI, C., and A. CHAVEZ. "Behaving as Expected: Public Information and Fairness Norms." *Journal of Behavioral Decision Making* 23(2010): 161-178.

CALLAHAN, C.M., E.A GABRIEL, and B.J. SAINTY. "A Review and Classification of Experimental Economics Research in Accounting." *Journal of Accounting Literature* 25 (2006): 59–126.

CAMERER, C.F, and E. FEHR. "Measuring Social Norms and Preferences Using Experimental games: A Guide for Social Scientists." In *Foundations of Human Sociality: Experimental and Ethnographic Evidence from Fifteen Small-Scale Societies*. Editors J. Henrich et al. Oxford: Oxford University Press, 2004. Pgs 55-95.

CHURCH, B.K., J.G. JENKINS, S.A. MCCRACKEN, P.B. ROUSH, and J.D. STANLEY. "Auditor Independence in Fact: Research, Regulatory, and Practice Implications Drawn from Experimental and Archival Research." *Accounting Horizons* 29 (2015): 217–238.

CIALDINI, R.B., and N.J. GOLDSTEIN. "Social Influence: Compliance and Conformity." *Annual Review of Psychology* 55 (2004): 591–621.

CIALDINI, R.B., C.A. KALLGREN, and R.R. RENO. "A Focus Theory of Normative Conduct - a Theoretical Refinement and Reevaluation of the Role of Norms in Human-Behavior." *Advances in Experimental Social Psychology* 24 (1991): 201–234.

CIALDINI, R.B., and M.R. TROST. "Social Influence: Social Norms, Conformity and Compliance." In *The Handbook of Social Psychology, Vol. 2*. 4th ed. Editors D.T. Gilbert et al. Boston: McGraw-Hill, 1998. Pgs. 151–192.

CLIKEMAN, P.M. Called to Account: Financial Frauds that Shaped the Accounting Profession. 2nd ed. New York: Routledge, 2013.

DEANGELO, L.E. "Auditor Independence, 'Low Balling', and Disclosure Regulation." *Journal of Accounting and Economics* 3 (1981a): 113–127.

DEFOND, M.L., K. RAGHUNANDAN, and K.R. SUBRAMANYAM. "Do Non-Audit Service Fees Impair Auditor Independence? Evidence from Going Concern Audit Opinions." *Journal of Accounting Research* 40 (2002): 1247–1274.

DEY, R.M., A. ROBIN, and D. TESSONI. "Advisory Services Rise Again at Large Audit Firms." *CPA Journal* 82 (2012): 58–67.

DOPUCH, N., and R.R. KING. "The Impact of MAS on Auditors' Independence: An Experimental Markets Study." *Journal of Accounting Research* 29 (1991): 60–98.

DOTY, J. "Auditing Plenary Session" Public Company Accounting Oversight Board at the American Accounting Association Audit Mid-Year Conference. 16 January 2015.

DUFWENBERG, M., and G. KIRCHSTEIGER. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2004): 268–298.

ELITZUR, R., A. GAVIOUS, and Y. MINCHUK. “Audit Fees Post-SOX: An Auction Model.” Working paper, University of Toronto, Ben Gurion University, and Shamoon College of Engineering. Available at http://ssrn.com/abstract=2801202.
ERICKSON, M., B.W. MAYHEW, and W.L. FELIX. “Why Do Audits Fail? Evidence from Lincoln Savings and Loan.” *Journal of Accounting Research* 38 (2000): 165–194.
FIOLLEAU, K., K. HOANG, K. JAMAL, and S. SUNDER. “How Do Regulatory Reforms to Enhance Auditor Independence Work in Practice?” *Contemporary Accounting Research* 30 (2013): 864–890.
FISCHBACHER, U. “Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments.” *Experimental Economics* 10 (2007): 171–178.
FISCHER, P., and S. HUDDART. “Optimal Contracting with Endogenous Social Norms.” *The American Economic Review* 98 (2008): 1459–1475.
GACHTER, S., and A. FALK. “Reputation and Reciprocity: Consequences for the Labour Relation.” *The Scandinavian Journal of Economics* 104 (2002): 1–26.
GEIGER, M.A., and D.V. RAMA. “Audit Fees, Nonaudit Fees, and Auditor Reporting on Stressed Companies.” *Auditing: A Journal of Practice & Theory* 22 (2003): 53–69.
GOULDNER, A.W. “The Norm of Reciprocity: A Preliminary Statement.” *American Sociological Review* 25 (1960): 161–178.
HARRIS, S. “The Rise of Advisory Services in Audit Firms.” Public Company Accounting Oversight Board at the Practicing Law Institute 12th Annual Directors' Institute on Corporate Governance. 24 November 2014.
HARRIS, S. “Issues for the Academic Community to Consider.” Public Company Accounting Oversight Board at the PCAOB/AAA Annual Meeting. 16 April 2015.
HURLEY, P.J., and B.W. MAYHEW. “Market Reactions to a High-Quality Auditor and Managerial Preference for Audit Quality.” Working paper, Northeastern University and the University of Wisconsin-Madison, 2017. Available at http://ssrn.com/abstract=2258865.
HURLEY, P.J., B.W. MAYHEW, and K. OBERMIRE. “Realigning Auditors’ Incentives: Experimental Evidence of a Third-Party Hiring System.” Working paper, Northeastern University and the University of Wisconsin-Madison, 2016.
JOHNSTONE, K.M., M.H. SUTTON, and T.D. WARFIELD. “Antecedents and Consequences of Independence Risk: Framework for Analysis.” *Accounting Horizons* 15 (2001): 1–18.
KACHELMEIER, S.J., and B.W. VAN LANDUYT. “Prompting the Benefit of the Doubt: The Joint Effect of Auditor-Client Social Bonds and Measurement Uncertainty on Audit Adjustments.” *Journal of Accounting Research* 55 (2017): 963–994.
KADOUS, K., S.J. KENNEDY, and M.E. PEECHER. “The Effect of Quality Assessment and Directional Goal Commitment on Auditors’ Acceptance of Client-Preferred Accounting Methods.” *The Accounting Review* 78 (2003): 759–778.
KAHNEMAN, D. 2003. “Maps of bounded rationality: Psychology for behavioral economics.” *The American Economic Review* 93 (5): 1449-1475.
KAHNEMAN, D., and A. TVERSKY. 1979. “Prospect theory: An analysis of decision under risk.” *Econometrica* 47 (2): 263-291.
KALLGREN, C.A., R.R. RENO, and R.B. CIALDINI. “A Focus Theory of Normative Conduct: When Norms Do and Do Not Affect Behavior.” *Personality and Social Psychology Bulletin* 26 (2000): 1002–1012.

KING, R.R. “Reputation Formation for Reliable Reporting: An Experimental Investigation.” *The Accounting Review* 71 (1996): 375–396.
KRISHNAN, J., H. SAMI, and Y. ZHANG. “Does the Provision of Nonaudit Services Affect Investor Perceptions of Auditor Independence?” *Auditing: A Journal of Practice & Theory* 24 (2005): 111–135.
KWON, S.Y. and Y. HAN. “Do Social Ties between CEOs and Engagement Audit Partners Affect Audit Quality and Audit Fees?” *Auditing: A Journal of Practice & Theory* (2017): Forthcoming.
LIBBY, R., R. BLOOMFIELD, and M.W. NELSON. “Experimental Research in Financial Accounting.” *Accounting, Organizations and Society* 27 (2002): 775–810.
LISIC, L.L., L.A. MYERS, R.J. PAWLEWICZ, and T.A. SEIDEL. “Do Accounting Firm Consulting Revenues Affect Audit Quality? Evidence from the Pre- and Post-SOX Eras.” *Contemporary Accounting Research* (2017): Forthcoming.
MAYHEW, B.W. “Auditor Reputation Building.” *Journal of Accounting Research* 39 (2001): 599–617.
MAYHEW, B.W., and J.E. PIKE. “Does Investor Selection of Auditors Enhance Auditor Independence?” *The Accounting Review* 79 (2004): 797–822.
MAYHEW, B.W., J.W. SCHATZBERG, and G.R. SEVCIK. “The Effect of Accounting Uncertainty and Auditor Reputation on Auditor Objectivity.” *Auditing: A Journal of Practice & Theory* 20 (2001): 49–70.
MOWCHAN, M. “Do Office Managing Partners Influence Audit Quality?” Working paper, Arizona State University. Available at http://ssrn.com/abstract=2784285.
MYERS, J.N., L.A. MYERS, and T.C. OMER. “Exploring the Term of the Auditor-Client Relationship and the Quality of Earnings: A Case for Mandatory Auditor Rotation?” *The Accounting Review* 78 (2003): 779–799.
PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). PCAOB Strategic Plan, 2015-2019”. Washington, D.C.: PCAOB, 2015. Available at http://pcaobus.org/sites/digitalpublications/Pages/PCAOB-strategic-plan-2015-2019.aspx
RAPOPORT, M. “KPMG Taps Veteran as First Female CEO.” *Wall Street Journal* (2015). http://www.wsj.com/articles/kpmg-taps-veteran-as-first-female-ceo-1429589268.
SARBANES-OXLEY ACT OF 2002. Public Law 107-2040. H.R. 3763.
SCHMIDT, J.J. “Perceived Auditor Independence and Audit Litigation: The Role of Nonaudit Services Fees.” *The Accounting Review* 87 (2012): 1033–1065.
SCHNEIDER, A., B.K. CHURCH, and K.M. ELY. “Non-Audit Services and Auditor Independence: A Review of the Literature.” *Journal of Accounting Literature* 25 (2006): 169–211.
SIMUNIC, D.A. “Auditing, Consulting, and Auditor Independence.” *Journal of Accounting Research* 22 (1984): 679–702.
SWIERINGA, R.J., and K.E. WEICK. “An Assessment of Laboratory Experiments in Accounting.” *Journal of Accounting Research* 20 (1982): 56–101.
TAYLER, W.B., and R.J. BLOOMFIELD. “Norms, Conformity, and Controls.” *Journal of Accounting Research* 49 (2011): 753–790.
UNITED STATES GENERAL ACCOUNTING OFFICE. “Public Accounting Firms: Mandated Study on Consolidation and Competition.” Report to the Senate Committee on Banking, Housing, and Urban Affairs and the House Committee on Financial Services (2003). Available at http://www.gao.gov/new.items/d03864.pdf.
WATSON, J. Strategy: An Introduction to Game Theory. 2nd ed. New York: W.W. Norton & Company, 2008.

Figure 1: Revenues from Audit and Assurance (AUDIT), Tax, and Nonaudit Services (NAS) as a percentage of total U.S. revenue for the Big 4 public accounting firms

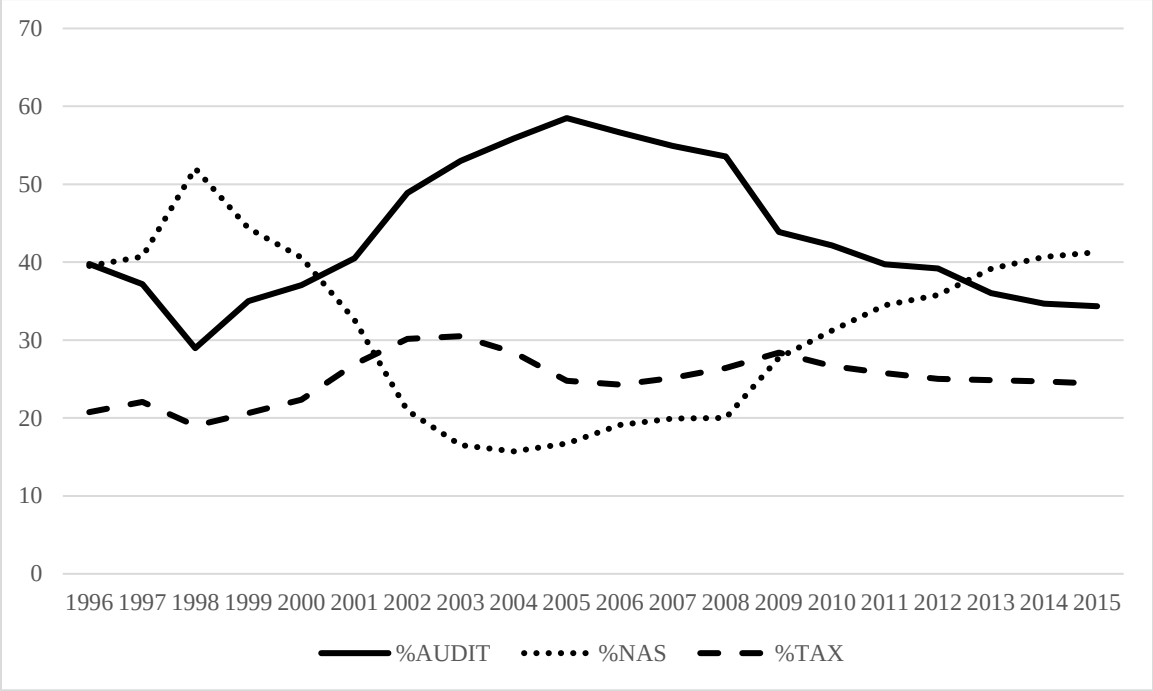

Fig. 1 - Source: "Top 100 Firms" annual report published by the Accounting Today Trade Magazine, consistent with Dey et al. [2012]. For 2003, 2004, and 2005, KPMG included NAS revenue in reported assurance revenue. We do not include KPMG data for these years.

Figure 2: PLANNED - Audit Quality by Condition

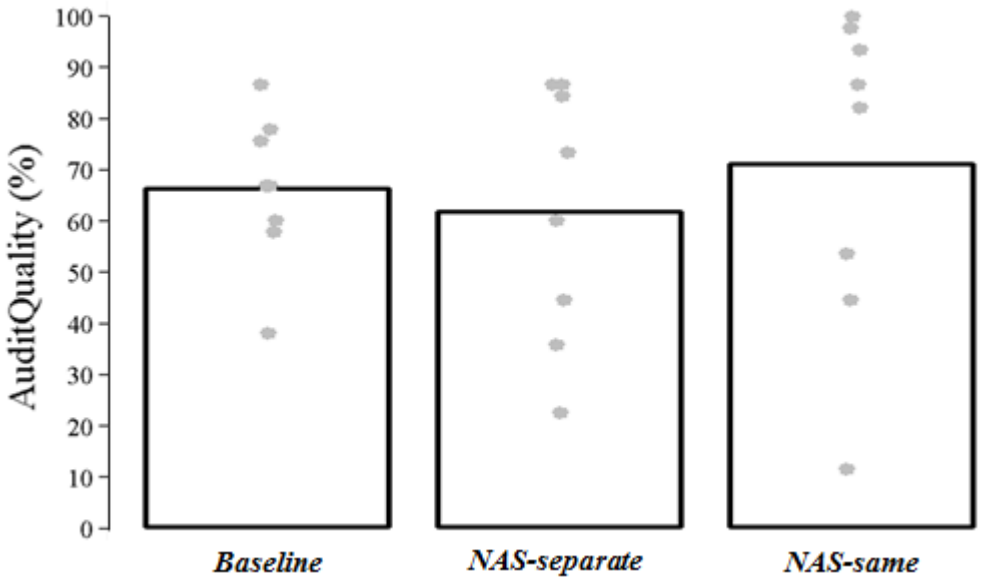

Fig. 2. - Bars show sample means on *AuditQuality*, which equals one when auditor chooses high-level verification, zero otherwise. Gray points plot average *AuditQuality* by session. Refer to Table 1 for description of Conditions.

Figure 3: UNPLANNED - Audit Quality by High Quality session designation

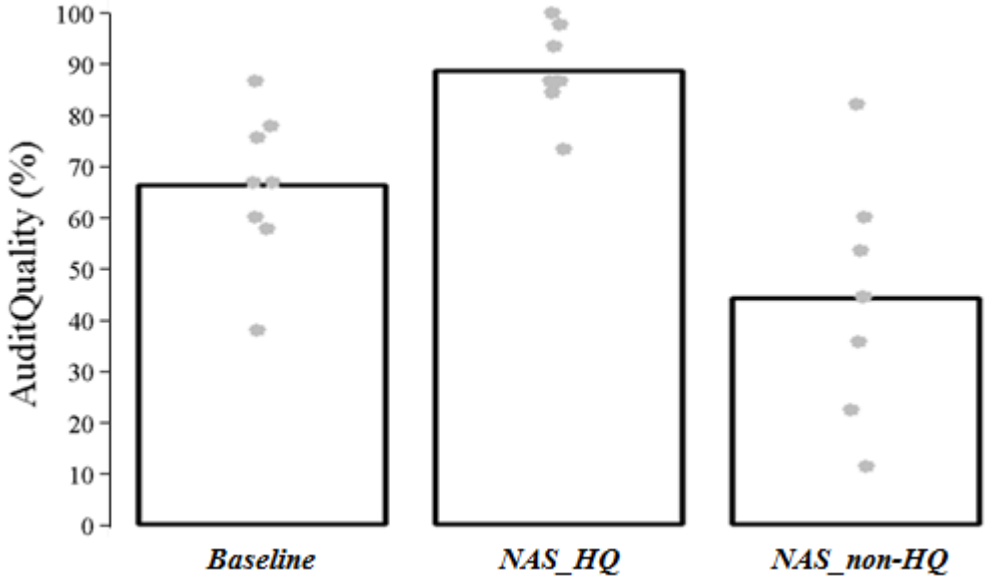

Fig. 3. - Bars show sample means on *AuditQuality*, which equals one when auditor chooses high-level verification, zero otherwise. Gray points plot average *AuditQuality* by session. *Baseline* solely consists of an auditor asset market. *NAS_HQ* includes all NAS sessions where a high-quality auditor (>90% *AuditQuality*) was hired most frequently. *NAS_non-HQ* includes all other NAS sessions. This analysis was not planned.

Figure 4: UNPLANNED - Audit Quality by NAS High Quality designation and Period

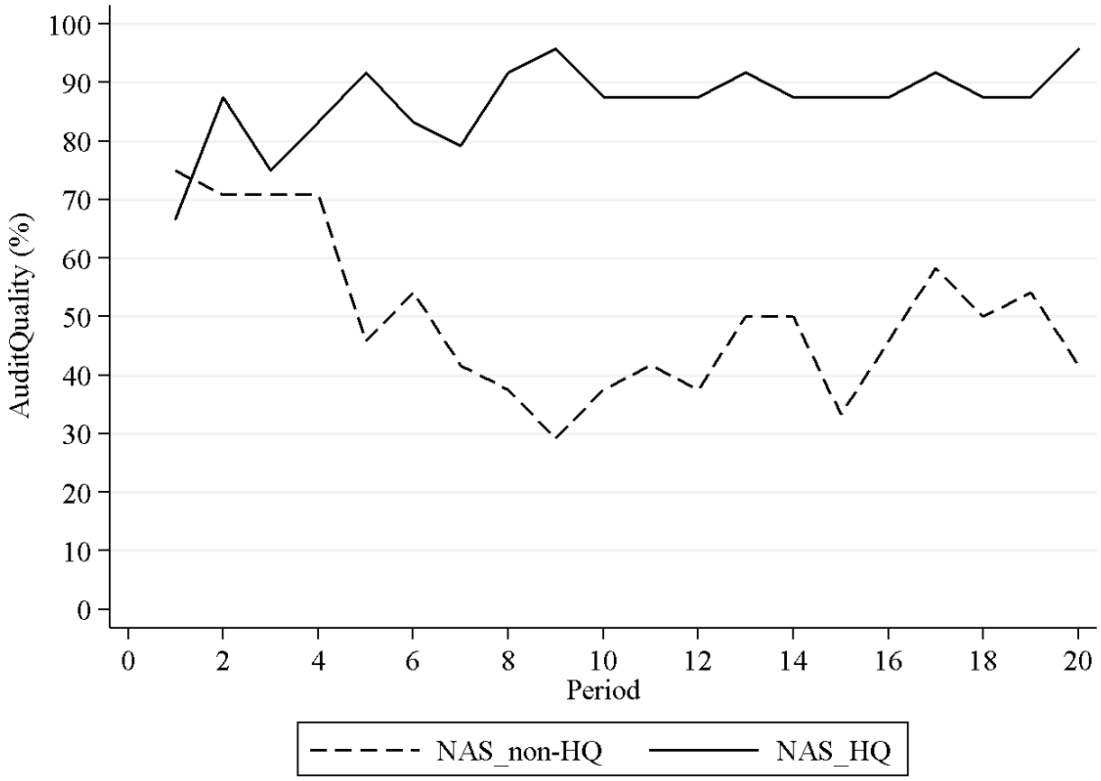

Fig. 4 – Plots *AuditQuality* by period to compare trends in *NAS_HQ* with *NAS_non-HQ*. *AuditQuality* equals one when auditor chooses high-level verification, zero otherwise. *NAS_HQ* includes all sessions where a high-quality auditor (>90% *AuditQuality*) was hired most frequently. *NAS_non-HQ* includes all other NAS sessions. This analysis was not planned.

Table 1: PLANNED - Descriptive Statistics by Condition

Baseline (n=80)									
Variable	1	2	3	4	5	6	7	8	Avg
AuditQuality	0.87	0.78	0.76	0.67	0.67	0.60	0.58	0.38	0.66
AuditorAccurate	0.80	0.78	0.82	0.71	0.71	0.76	0.76	0.58	0.74
MgrInvestment	0.42	0.33	0.24	0.38	0.40	0.42	0.67	0.11	0.37
MgrMisreporting	0.80	0.73	0.82	0.89	0.76	0.71	0.60	0.69	0.75
HighBid	550.24	614.82	593.44	536.91	568.98	563.60	727.62	524.76	585.05
HighBidderEarned	-8.02	-94.82	-77.89	-70.24	-97.87	27.51	-105.40	-58.09	-60.60
HQ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NAS-Separate (n=104)									
Variable	1	2	3	4	5	6	7	8	Avg
AuditQuality	0.87	0.87	0.84	0.73	0.60	0.44	0.36	0.22	0.62
AuditorAccurate	0.82	0.89	0.87	0.80	0.80	0.44	0.47	0.31	0.68
MgrInvestment	0.31	0.53	0.27	0.62	0.62	0.47	0.24	0.27	0.42
MgrMisreporting	0.89	0.51	0.76	0.33	0.44	0.73	0.82	0.89	0.67
HighBid	578.24	561.69	566.51	624.62	728.60	602.98	604.18	628.40	611.90
HighBidderEarned	-49.36	24.98	-42.07	-6.84	-168.60	-34.09	-88.62	-143.96	-63.57
ConsultantEffort	0.93	0.91	1.00	1.00	1.00	0.96	1.00	0.33	0.89
HQ	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.50
NAS-Same (n=80)									
Variable	1	2	3	4	5	6	7	8	Avg
AuditQuality	1.00	0.98	0.93	0.87	0.82	0.53	0.44	0.11	0.71
AuditorAccurate	0.93	0.93	0.89	0.98	0.89	0.64	0.60	0.18	0.76
MgrInvestment	0.93	0.51	0.44	0.49	0.69	0.33	0.56	0.20	0.52
MgrMisreporting	0.69	0.67	0.78	0.36	0.49	0.80	0.71	0.93	0.68
HighBid	731.58	669.69	746.51	623.49	687.31	599.69	686.53	564.18	663.62
HighBidderEarned	-33.80	-29.69	-195.40	-32.38	6.02	-137.47	-122.09	-155.29	-87.51
ConsultantEffort	0.96	1.00	0.93	0.96	0.96	0.69	0.93	0.91	0.92
HQ	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.50

Table 1 describes the mean by variable from each session. We drop the first five periods to remove learning effects. We drop all periods beyond the first 20 to reduce end-period effects. *Baseline* solely consists of an auditor asset market. *NAS-same* indicates that the manager jointly hires for both the consulting project and asset market. *NAS-separate* indicates that separate groups of managers participate in the consulting and asset markets. Noted *n* indicates the number of participants in each condition.

Variable Definitions: *AuditQuality* = 0 auditor chooses low-level verification, 1 high-level verification; *AuditorAccurate* = 0 audit report inaccurate, 1 accurate; *MgrInvestment* = 0 manager chooses low-investment, 1 high-investment; *MgrMisreporting* = 0 manager reports true value, 1 lies; *HighBid* = Highest bid that was made by an investor for the manager's asset; *HighBidderEarned* = Amount received (lost) by the HighBid investor; *ConsultantEffort* = 0 auditor-consultant selects low-effort, 1 high-effort; *HQ* = 1 if, in unplanned analysis, the high-quality auditor (>90% *AuditQuality*) is hired most frequently.

Table 2: PLANNED - Audit Quality Hypotheses (Baseline > NAS-sep > NAS-same)

Panel A: Summary statistics: mean and standard deviation	Audit Quality	
Condition	Mean	Std Dev
Baseline	66%	0.15
NAS-separate	62%	0.25
NAS-same	71%	0.32

Panel B: Analysis of Variance			
Source	*df*	F-stat	p-value
Group	2	0.44	0.64
Residual	21		

Panel C: Planned Contrasts			
Source	*df*	z-stat	p-value
Baseline > *NAS-same* (H1)	1	-0.58	0.56
NAS-separate vs. *Baseline* and *NAS-same*	1	-0.75	0.46
Baseline > *NAS-separate* (H2)	1	0.35	0.72
NAS-separate > *NAS-same* (H3)	1	-0.93	0.35
By-subject random intercept	Yes		
By-session random intercept	Yes		

Table 2 presents the results of our audit quality hypothesis tests. Each panel includes 360 observations, 72 subject groups, and 24 session groups. The dependent variable, *AuditQuality*, equals one when auditor chooses high-level verification, zero otherwise. Panel A presents the mean and standard deviation by condition. Panels B and C present a generalized linear mixed effects model used to run a logit regression. Panel B *Group* indicates group assignment is specified in the model. Panel C presents the planned contrasts used for hypothesis testing.

See table 1 for condition descriptions

p-values are two-tailed as results are inconsistent with directional predictions.

Table 3: UNPLANNED - Descriptive Statistics by Auditor

NAS-Same						*NAS-Separate*						*Baseline*			
SessionAQ	*MktShr*	*Rank*	*AQ*	*CRank*	*C_Effort*	*SessionAQ*	*MktShr*	*Rank*	*AQ*	*Mkt Share*	*CRank*	*C_Effort*	*SessionAQ*	*MktShr*	*Rank*	*AQ*
100.00%	**51%**	**1**	**100.00%**	1	90.63%	88.33%	**65%**	**1**	**100.00%**	45%	1	88.89%	81.94%	38%	3	59.26%
	27%	2	100.00%	2	88.24%		20%	3	41.67%	8%	3	0.00%		36%	1	100.00%
	22%	3	100.00%	3	71.43%		15%	2	100.00%	47%	2	82.14%		26%	2	89.47%
91.67%	**51%**	**1**	**100.00%**	1	97.30%	85.00%	**50%**	**1**	**93.33%**	70%	1	100.00%	80.00%	40%	2	75.00%
	43%	2	83.87%	2	90.32%		33%	3	70.00%	5%	3	33.33%		33%	1	100.00%
	6%	3	75.00%	3	25.00%		17%	2	90.00%	25%	2	86.67%		27%	3	62.50%
86.36%	**55%**	**1**	**100.00%**	1	94.44%	84.72%	**50%**	**1**	**100.00%**	61%	1	100.00%	71.67%	43%	2	69.23%
	23%	2	93.33%	2	80.00%		40%	3	68.97%	21%	3	73.33%		28%	3	64.71%
	23%	3	46.67%	3	80.00%		10%	2	71.43%	18%	2	84.62%		28%	1	82.35%
85.00%	**50%**	**1**	**100.00%**	1	100.00%	71.43%	**52%**	**1**	**93.94%**	35%	1	100.00%	63.64%	42%	1	85.71%
	33%	2	70.00%	3	75.00%		24%	2	93.33%	32%	2	100.00%		30%	2	55.00%
	17%	3	70.00%	2	80.00%		24%	3	0.00%	33%	3	100.00%		27%	3	38.89%
78.33%	48%	2	72.41%	3	89.66%	68.33%	63%	3	55.26%	2%	2	0.00%	63.49%	41%	3	57.69%
	42%	1	96.00%	1	100.00%		20%	1	100.00%	98%	1	100.00%		32%	1	70.00%
	10%	3	33.33%	2	100.00%		17%	2	80.00%	0%	N/A	N/A		27%	2	64.71%
57.14%	40%	2	64.00%	1	92.00%	46.67%	48%	3	0.00%	8%	3	0.00%	60.00%	37%	3	9.09%
	33%	3	23.81%	2	57.14%		28%	2	88.24%	72%	1	95.35%		35%	1	95.24%
	27%	1	88.24%	3	35.29%		23%	1	92.86%	20%	2	83.33%		28%	2	82.35%
53.33%	48%	1	65.52%	1	100.00%	42.86%	40%	1	68.00%	87%	1	96.36%	58.73%	41%	2	53.85%
	28%	3	23.53%	2	94.12%		37%	2	43.48%	10%	3	33.33%		30%	1	84.21%
	23%	2	64.29%	3	64.29%		24%	3	0.00%	3%	2	50.00%		29%	3	38.89%
16.67%	50%	3	0.00%	2	90.00%	31.82%	47%	2	22.58%	11%	3	14.29%	40.00%	43%	1	50.00%
	38%	1	34.78%	1	91.30%		29%	1	63.16%	21%	2	28.57%		37%	2	36.36%
	12%	2	28.57%	3	71.43%		24%	3	12.50%	68%	1	55.56%		20%	3	25.00%

Table 3 presents descriptive statistics of each auditor's behavior by session and market designation. *SessionAQ* presents % times auditor selects high-level verification by session; this variable does not reconcile with Table 1 as all periods are included. *MktShr* is the number of times hired divided by all possible hires. We *Rank* auditor by *AQ* (% times auditor selects high-level verification); we *CRank* consultants by *C_Effort* (% times auditor-consultant selects high-effort). We use this table to identify NAS_HQ sessions (identified as lightly shaded). *NAS_HQ* includes all sessions where a high-quality auditor (>90%) was hired most frequently as noted with bold lettering. *NAS_non-HQ* includes all other NAS sessions not meeting the HQ definition.

Table 4: UNPLANNED - High Quality sessions

Panel A: Descriptive statistics: mean and standard deviation	Audit Quality	
Condition	Mean	Std Dev
NAS HQ	89%	8%
Baseline	66%	15%
NAS Non-HQ	44%	22%

Panel B: Ex Post Tests

Source	*df*	z-stat	p-value
Baseline vs NAS HQ	1	-2.69	0.007
Baseline vs NAS Non-HQ	1	2.25	0.025
By-subject random intercept	Yes		
By-session random intercept	Yes		

Table 4 presents the results of our unplanned analysis over High Quality (HQ) sessions. Eight NAS sessions were dominated by a high quality auditor (HQ), where the auditor with the highest market share provided >90% audit quality. The remaining eight NAS sessions did not meet this criterion (NAS non-HQ) nor did any Baseline sessions. Each panel includes 1,080 observations, 72 subject groups, and 24 session groups. The dependent variable, *AuditQuality*, equals one when auditor chooses high-level verification, zero otherwise. Panel A presents the mean and standard deviation by HQ designation and condition. Panel B presents the results from our generalized linear mixed effects logit regression. p-values are two-tailed.

Table 5: PLANNED - Manager Analysis

	MgrInvestment	*MgrMisreporting*
AuditorAccuracyPer	3.68	-3.41
	(<0.01)	(< 0.01)
By-subject random intercept	Yes	Yes
By-session random intercept	Yes	Yes

Table 5 presents two generalized linear mixed effects models with one of two dependent variables. *MgrInvestment* is coded as 1 if high investment, 0 if low. *MgrMisreporting* is coded as 1 if the manager misreports the asset value, 0 otherwise. The independent variable, *AuditorAccuracyPer* measures times auditor accurate divided by times hired at beginning of period. The *MgrInvestment* model uses 1,080 observations. As the *MgrMisreporting* model drops observations with maximum value assets with no potential to misreport, it only uses 999 observations. Both models use 72 subject groups, and 24 session groups. Coefficients are presented above the p-value for each test.

Table 6: UNPLANNED - Manager Analysis in High Quality sessions

Panel A: Descriptive statistics: mean and standard deviation	Mgr Investment	
Condition	Mean	Std Dev
NAS HQ	51%	21%
Baseline	37%	16%
NAS Non-HQ	42%	19%

Panel B: Ex Post Tests

Source	*df*	z-stat	p-value
NAS HQ vs Baseline and NAS Non-HQ	1	1.37	0.169
By-subject random intercept	Yes		
By-session random intercept	Yes		

Table 6 presents the results of our unplanned analysis over High Quality (HQ) sessions. Eight NAS sessions were dominated by a high quality auditor (HQ), where the auditor with the highest market share provided >90% audit quality. The remaining eight NAS sessions did not meet this criteria (NAS non-HQ) nor did any Baseline sessions. Each panel includes 1080 observations, 72 subject groups, and 24 session groups. The dependent variable, *MgrInvestment*, equals one if high investment, 0 if low. Panel A presents the mean and standard deviation by HQ designation and condition. Panel B presents the results from our generalized linear mixed effects logit regression. p-values are two-tailed.

Table 7: PLANNED - Investor Analysis

	(1)	(2)
AuditorAccuracyPer_mc	288.43	287.55
	(< 0.01)	(< 0.01)
MgrReport_c	-0.35	-0.35
	(< 0.01)	(< 0.01)
*AuditorAccuracyPer_mc * MgrReport_c*	0.87	0.87
	(< 0.01)	(< 0.01)
Period_c	-3.85	-3.85
	(< 0.01)	(< 0.01)
NAS-separate		-7.96
		(0.83)
NAS-same		28.31
		(0.44)
By-subject random intercept	Yes	Yes
By-session random intercept	Yes	Yes

Table 7 presents two generalized linear mixed effect models with a dependent variable of *Confidence*. Model one studies the effects of audit accuracy. Model two studies whether investors perceive consulting as an incremental threat to reporting quality, above actual accuracy. *Confidence* is measured as the high bid for an asset minus the reported value. *AuditorAccuracyPer_mc* measures times auditor accurate divided by times hired at beginning of period, centered at the mean 0.746. *MgrReport_c* measures the manager's reported asset value, midpoint centered at 800. *Period_c* reports the period in the market, midpoint centered at 13. See table 1 for condition descriptions. The sample is limited to instances when the managers report more than 200, an agree report was issued, and the auditor has been hired in at least one previous period. Each model uses 611 observations, 84 subject groups, and 24 session groups. Coefficients are presented above the p-values for each test.

Appendix A Table 1: Post-experiment questionnaire responses

Panel A: Provider (Auditor) responses (n=72)		*Planned analysis* Mean (SD)		*Unplanned analysis* Mean (SD)		p-value	
Question	*Baseline*	*NAS-separate*	*NAS-same*	*NAS HQ*	*NAS non-HQ*	*HQ vs. non-HQ*	*HQ & non-HQ vs. Baseline*
When you were hired in the project market…							
P1 … what did you believe was the appropriate service choice?	N/A	5.35 (2.23)	6.04 (1.23)	6.13 (1.55)	5.29 (1.97)	0.11	N/A
P2 … what service did you believe the Decision Maker expected you to choose?	N/A	5.65 (1.85)	6.75 (0.85)	6.57 (0.79)	5.88 (1.94)	0.12	N/A
P3 … what service did you feel obligated to choose?	N/A	5.00 (2.26)	5.83 (1.71)	5.87 (1.84)	5.00 (2.13)	0.14	N/A
P4 What service did you expect other Providers to choose when hired?	N/A	4.48 (2.27)	5.88 (1.57)	5.78 (1.59)	4.62 (2.30)	0.05	N/A
When you were hired in the asset market…							
P5 … what did you believe was the appropriate verification choice?	4.79 (1.67)	4.96 (2.22)	4.92 (1.91)	5.58 (1.67)	4.29 (2.22)	0.02	0.76
P6 … what verification level did you believe the Decision Maker expected you to choose?	3.58 (2.19)	4.04 (2.07)	4.46 (2.23)	5.08 (1.82)	3.42 (2.15)	0.01	0.20
P7 … what verification level did you believe the Bidder expected you to choose?	5.00 (2.11)	5.21 (1.77)	5.25 (2.29)	5.58 (1.64)	4.88 (2.33)	0.23	0.66
P8 … what verification level did you feel obligated to choose?	4.46 (2.19)	4.92 (1.89)	5.04 (2.10)	5.75 (1.59)	4.21 (2.04)	0.01	0.29
P9 What verification level did you expect other Providers to choose when hired?	3.88 (1.75)	4.67 (1.83)	5.00 (1.74)	5.50 (1.38)	4.17 (1.90)	0.01	0.03

Panel A presents auditor responses to post-experiment questions about factors in the decision-making. Questions P1-P4 regarding project market choices are measures on a scale of 1 (always service Y) to 7 (always service X). Questions P5-P9 regarding the asset market are measured on a scale from 1 (always low) to 7 (always high). The planned presentation by condition is supplemented with unplanned presentation by HQ designation and contrast tests. A significant difference between HQ and non-HQ coupled with a lack of difference between Baseline and the average of HQ and non-HQ supports a HQ > Baseline > non-HQ pattern, consistent with our interpretation of the results.

Open-ended provider (auditor) questions – Responses not presented

What was your overall strategy as a Provider in this experiment?
Was your strategy consistent? Did you change strategies as you progressed through the periods?
If you changed your strategy: 1) Why did you change it? 2) When did you change it? And 3) How did you change it?
What factors did you consider when deciding whether to perform High- or Low-level effort in the project market?
What factors did you consider when deciding whether to perform a High- or Low-level verification in the asset market?

Appendix A Table 1 (continued): Post-experiment questionnaire responses

Panel B: Decision Maker (Manager) responses	*Planned analysis* Mean (SD)			*Unplanned analysis* Mean (SD)		p-value	
(n=72) Question	*Baseline*	*NAS-separate**	*NAS-same*	*NAS HQ*	*NAS non-HQ*	*HQ vs. non-HQ*	*HQ & non-HQ vs. Baseline*
In the project market, when selecting a Provider…							
M1 … how often did you expect them to choose Service X?	N/A	5.54 (1.67)	5.75 (1.33)	5.88 (1.33)	5.42 (1.64)	0.29	N/A
M2 … did you prefer a Provider with a past history of selecting Service X?	N/A	6.50 (1.32)	5.63 (1.97)	6.04 (1.63)	6.08 (1.84)	0.93	N/A
M3 … how much did you consider the Provider's accuracy in the asset market?	N/A	3.00 (2.11)	5.38 (1.53)	4.21 (2.21)	4.17 (2.20)	0.95	N/A
In the asset market,							
M4 When choosing your investment, how much did you consider the risk of not getting a high valued asset?	4.46 (1.93)	5.17 (1.37)	4.29 (2.05)	4.88 (1.75)	4.58 (1.84)	0.59	0.56
M5 When selecting a Provider to hire, how often did you expect them to choose high-level verification?	4.50 (1.50)	4.54 (2.13)	4.71 (1.60)	4.88 (1.94)	4.38 (1.79)	0.33	0.78
M6 When selecting a Provider to hire, how much did you consider your planned investment choice (i.e. investment A or B)?	4.79 (2.00)	4.54 (2.28)	4.42 (2.22)	4.50 (2.43)	4.46 (2.06)	0.95	0.57
M7 When selecting a Provider to hire how much did you prefer a Provider with a history of high accuracy?	4.50 (2.15)	4.63 (1.95)	4.17 (1.77)	4.92 (1.86)	3.87 (1.74)	0.07	0.83
M8 When selecting a Provider to hire, how much did you prefer a Provider with a history of failed checks?	4.33 (1.95)	3.92 (1.89)	2.75 (1.51)	3.21 (1.96)	3.46 (1.64)	0.64	0.03
M9 When selecting a Provider to hire, how much did you consider which Provider you hired in the previous periods?	5.21 (1.67)	5.50 (1.62)	5.83 (1.44)	6.21 (1.18)	5.09 (1.65)	0.01	0.25
M10 When selecting a Provider to hire, how much did you consider the Provider's service history in the project market?	N/A	2.96 (2.03)	4.92 (1.95)	4.65 (2.14)	3.29 (2.10)	0.03	N/A
M11 When selecting what asset amount to report to Bidders, how much did you consider which Provider would be investigating your reported asset value?	5.92 (1.28)	6.04 (1.46)	4.71 (1.73)	5.43 (1.80)	5.29 (1.68)	0.76	0.17

Panel B presents manager responses to post-experiment questions about factors in decision-making. Responses are measured on a scale from 1 (not at all) to 7 (a lot).

Open-ended decision-maker (manager) questions – Responses not presented

What was your overall strategy as a Decision Maker in this experiment?
Was your strategy consistent? Did you change strategies as you progressed through the periods?
If you changed your strategy: 1) Why did you change it? 2) When did you change it? And 3) How did you change it?
What factors did you consider when deciding which Provider to hire in the project market? [b, c]
What factors did you consider when deciding which Provider to hire in the asset market? [b, c]
What factors did you consider when deciding which Provider to hire? [a]
What factors did you consider when deciding whether select Investment A or Investment B?
What factors did you consider when deciding what amount to report to Bidders?

Appendix A Table 1 (continued): Post-experiment questionnaire responses

Panel C: Bidder (Investor) responses		Mean (SD)			*Unplanned analysis* Mean (SD)		p-value	
(n=91) Question		*Baseline*	*NAS-separate*	*NAS-same*	*NAS HQ*	*NAS non-HQ*	*HQ vs. non-HQ*	*HQ & non-HQ vs. Baseline*
	When deciding the amount to BID for each asset, how much did you consider…							
B1	… the reported asset value that was provided by the Decision Maker?	4.97 (1.62)	4.97 (1.80)	5.32 (1.54)	5.17 (1.71)	5.13 (1.65)	0.92	0.62
B2	… the Provider's Agree or Disagree report?	5.26 (1.79)	5.97 (1.24)	5.52 (1.81)	6.34 (0.94)	5.16 (1.81)	< 0.01	0.16
B3	… the Provider's history of accuracy?	5.83 (1.39)	6.31 (1.17)	5.80 (1.79)	6.28 (1.33)	5.83 (1.68)	0.25	0.51
B4	… the bids made by other Bidders in previous periods?	4.19 (1.96)	4.28 (2.03)	4.57 (2.06)	5.14 (1.86)	3.77 (2.00)	0.01	0.54
B5	… the true value of the assets in previous periods?	4.58 (1.52)	4.79 (1.88)	5.10 (1.90)	5.45 (1.92)	4.48 (1.75)	0.03	0.32
B6	… that you would receive a return that is based on the total true value of all assets that you purchased during the market?	4.45 (1.86)	5.00 (1.65)	4.94 (1.50)	5.24 (1.70)	4.71 (1.40)	0.22	0.16
B7	… the Provider's history of failed checks	3.84 (2.13)	4.45 (1.99)	3.97 (2.21)	4.17 (2.17)	4.23 (2.08)	0.92	0.45

Panel C presents investor responses to post-experiment questions about factors in decision-making. Responses are measured on a scale from 1 (not at all) to 7 (a lot).

Open-ended bidder (investor) questions – Responses not presented

What was your overall strategy as a Bidder in this experiment?
Was your strategy consistent of did you change your strategy as you progressed through the periods?
If you changed your strategy: 1) Why did you change it? 2) When did you change it? And 3) How did you change it?

[All participants]

Demographic Questions – Responses not presented

What is your gender?
How old are you today?
What is your first language?
What is your standing at the University? [Freshman, Sophomore, Junior, Senior, Graduate, Other]
What is your expected major?

[a] Question is included in questionnaires for NAS-same, but excluded from some other treatments.

[b] Question is included in questionnaires for NAS-separate, but excluded from some other treatments.

[c] Question is included in questionnaires for NAS-different, but excluded from some other treatments.