

Are Regulatory Mandate and Independence Necessary For Audit Quality?

Karim Jamal, University of Alberta
Shyam Sunder, Yale University

Revised October 13, 2008

© Copyright 2006-8, All Rights Reserved

We thank Jon Davis, Michael Gibbins, Joel Pike, Brad Pomeroy, Steve Salterio, Tom Scott and Workshop participants at the University of Alberta, Carnegie Mellon University, Chinese University of Hong Kong, City University of Hong Kong, University of Kansas, Mannheim University, Nanyang Technological University, National Taiwan University, Queens University, Sabanci University, University of Windsor, Yale School of Management, and Participants at the University of Illinois Audit Symposium and Interdisciplinary Accounting Conference in Copenhagen for helpful comments. We also thank Adam Karbani and Wayne Morgan for research assistance on this project. Financial support from the Social Sciences and Humanities Research Council (SSHRC) of Canada is gratefully acknowledged. Please send comments to karim.jamal@ualberta.ca or shyam.sunder@yale.edu.

Are Regulatory Mandate and Independence Necessary For Audit Quality?

Abstract

Mandated audit by independent accountants is a key feature of U.S. regulation of financial statements. Private incentives to demand (and supply) certification are thought to be insufficient, and independence is thought to be necessary for quality audit. This study presents archival data on certification activity in the economy, and conducts a field experiment on an unregulated market for certification of baseball cards to investigate the validity of these assumptions. Our results show that: (1) Private markets for certification services and potential for conflict of interest are ubiquitous; (2) The grading scales in certification vary in their fineness with pass/fail often used by government agencies; (3) the unregulated market for baseball card certification is dominated by firms who provide higher quality services; and (4) are not independent. The assumption that private incentives are insufficient for a well-functioning audit market may need further scrutiny; independence does not seem necessary for audit quality.

Keywords: Mandatory audit, independence, regulation, certification services, financial reporting, accounting

JEL Classifications: M49, G34, G38

Are Regulatory Mandate and Independence Necessary For Audit Quality?

The significance of independence in the work of the independent auditor is so well established that little justification is needed to establish this concept as one of the cornerstones in any structure of auditing theory.

Mautz and Sharaf, *The Philosophy of Auditing*, 1961

1. Introduction

Two key features of the market for audit of financial reports are that the audit is mandated for publicly-held companies, and that it must be conducted by independent accountants. Private demand and supply of audits are assumed to be insufficient for the proper functioning of audit markets, necessitating the mandate. This belief that audit services are undersupplied in the (unregulated) economy led the audit profession to re-organize itself during the 1990's into a broader "assurance" service provider to widen the set of audit services offered (Elliott [1998]). Independence is widely assumed to be necessary for the quality of audit, and often audit quality is equated with independence. We provide some evidence on the validity of these important foundational assumptions.

The mandate for independent audit of financial reports of public firms was created under the U.S. securities laws enacted in 1932-33. In response to a wave of corporate accounting frauds such as Enron and WorldCom, U.S. Congress enacted the Sarbanes-Oxley Act (SOX) in 2002, banned auditors from performing certain non-audit services for their audit clients, and mandated certain corporate governance reforms to enhance the independence of audits. SOX also transferred the responsibility for setting U.S. auditing standards to a new organization (Public Company Accounting Oversight Board or PCAOB) which was created by this government Act but is not formally a part of U.S. Government. Actual certification of corporate financial reports continues to be conducted by private CPA firms under the PCAOB's regulation and oversight.

Whether the SOX reforms actually improved audits remains controversial (Kinney et al. [2004])¹.

We seek to understand the consequences and necessity of the regulatory mandate for audit by independent auditors. Prior approaches to this problem involved examination of pre-1930s audits of public firms (e.g., Chow [1982], and Merino et al. [1994]), and contemporary audits of privately held firms (e.g., Chaney et al. [2004]). In this paper we add a third approach to the literature by examining unregulated markets for certification services elsewhere in the economy. Since certification services are widely available in many sectors of the economy (Power [1994]), the present study compares such services across markets. Private certification markets also provide an opportunity to observe markets with pure certifiers (who do not sell other services) competing with certifiers who cross-sell other services. For example, what might be the consequences for auditors and financial markets if audit firms could not offer non-audit services? Private certification markets provide an opportunity for exploring certification quality. In accounting context, the quality of audit, and the impact of regulation of quality has been the subject of much debate. While no two markets are exactly alike, cross-market comparisons may help us better understand the characteristics and sources of variations in certification services, especially the demand for certification in absence of regulatory mandates and lack of independence of the providers.

¹ There is some empirical evidence that the claim that consulting services impair auditor independence is incorrect. Studies have been conducted using accounting accruals, accounting re-statements, SEC enforcement actions, and litigation against auditors as dependent variables. Consulting either has no effect (Ashbaugh et al. [2003]; DeFond et al. [2002]), a positive effect (Dopuch et al. [2003]; Kinney et al. [2004]) or a small negative effect (Frankel et al. [2002]) on the quality of accounting numbers. The Frankel result is driven by a few small companies in their sample (Larcker and Richardson [2004]).

The paper is organized as follows. Section 2 outlines a framework for analysis of certification services and reports the results of our search for certification services for a wide range of goods and services sold online and offline. We also report on the fineness of reporting scales used by private certifiers and federal government agencies who conduct audits to ascertain compliance with the relevant regulations. Section 3 reports on the expertise and potential conflicts of interest of people who are available online to provide their assessments in exchange for a fee. We also report the results of a field experiment (and use the results of a prior field experiment reported in the economics literature) of graded and ungraded baseball cards traded online at EBay. Twenty-three graders compete in this online market. Some of them can be called independent certifiers in the sense that they do not provide any services other than grading, whereas others are cross-sellers and offer a host of related services, including some that create potential for conflict of interest. A summary and implications are discussed in Section 4.

Our results indicate that: (1) there is widespread demand for certification services in virtually all parts of the economy, mostly in absence of regulatory mandate. (2) These services are often provided by experts who are known to have potential conflicts of interest. (3) Private certification agencies provide reports using scales and fineness as instruments of competition. Government agencies on the other hand, generally provide standard (boilerplate) reports. (4) In the baseball card certification market, services of strict graders are more valued in the market than the services of the lenient graders. There is no race to the bottom. (5) Independence is not necessary for audit quality. Certification agencies who cross sell services provide better audit quality than independent certification agencies. These findings about the certification markets call for careful re-examination of the assumptions that the audit mandate and auditor independence are necessary for quality of financial audits. We discuss these in the concluding remarks.

2. A Framework for Analysis of Certification Services

We use “to certify” in its broad meaning of “to give assurance of quality or validity,” in the sense conveyed by terms such as corroborate, verify, validate, measure, and guarantee. Let us consider the sources of demand for and supply of certification services in absence of regulatory mandates.

Buyers like to have an assurance that they receive what they pay for. The willingness of buyers to pay for certification depends on their personal ability to assess quality. Snow shovels, with their quality more transparent, are often sold without certification; quality of toasters, cars and diamonds being less obvious, they often carry some certification. As with other credence goods, the value of certification to the buyer depends on the reputation of the certifier. Finally, buyers would want certification from a party who has the motivation to be accurate. If the buyer has reasons to doubt the incentives of the provider, especially conflict of interest, his willingness to pay for the service would be attenuated.

On the supply side of certification services, knowledge and experience add accuracy. Transactions experience in the relevant market enhances expertise, but also introduces mixed-motives (between revenues from certification and from direct transactions) for the supplier. Greater expertise should reduce the cost of discriminating among goods of varying quality

Once a good has been certified, the marginal cost of additional buyers receiving the benefit of the certification is essentially zero, and in such cases certification is a public good. The difficulties of organizing the beneficiaries to pay for the cost of public goods (since individuals have incentives to free ride) lead to well-known problems of under-production.² In case of certification services, even if there are multiple potential buyers, there is usually a single seller

² Even with well functioning competition, the market cannot supply the proper amount of a public good.... With everyone similarly motivated to understate his demand price, the good will be supplied in inadequate quantity. Hence the state must undertake the provision of the public good and finance it with a tax. George Stigler, The Citizen and the State (1975).

who has adequate incentives to pay for certification in the hope of obtaining a higher price for the good itself. Moody's ratings of corporate bonds, A.M. Best's ratings of insurance, Powers ratings of automobiles and Underwriters Laboratories ratings of appliance are examples of this phenomena. Certification of financial reports of corporations is also paid for by corporations. However, as decision makers, hired managers of corporations have mixed motives. When corporations issue new equity or debt securities, they could obtain a higher price by having the reports certified for their quality and reliability on one hand and by manipulating them to present a rosy picture on the other. Similarly, subsequent to the date of initial issue of the securities, the incentives to produce reliable financial reports are diluted by the presence of managerial mixed motives. Some managers may view the trustworthiness of their financial reports as desirable, while for others, personal benefits of opportunistically manipulating the financial reports may encroach on the desirability of a high quality audit.

Existence of written, common knowledge standards makes it easier for the buyer and the seller of certification to communicate the meaning and content of the certificate. The fineness of the classification scheme associated with the certificate would be determined by the trade-off between the errors of the certifier and the sensitivity of the buyer decisions to finer distinctions and grades. Thus the customer of a certification service could be interested in several characteristics of the certifier:

Knowledge: the customer would prefer to have certification by someone who is knowledgeable about the good or service, and look for evidence in the form of credentials—a license, degree, or experience—that gives him/her a claim to have knowledge of the subject.

Reputation. Since the customers often pay for certification services in the hope of getting better terms of trade in any subsequent transactions, both actual knowledge as well as the reputation of the certifier are valued in this market.

Motivation: the customer would prefer to obtain certification from someone who has the motive to provide good quality service, or at least, does not have a conflict of interest that may keep him/her from providing good quality service. Payment for the service provides a positive motivation for the certifier; independence is a way of ascertaining that the certifier does not have motives that may adversely affect the interests of the customer.

Standards: The value and feasibility of using written standards varies by the context of certification. Written standards are important, for example, in grading of beef, steel welds, and most engineering products. However, when it comes to art, design, historical documents, and service in a restaurant, the value and feasibility of written standards in providing better certification is less clear. Another way of framing this is that certification (for all goods) can be provided with or without the existence of formal written standards, and the two kinds of services can compete with each other.

Report: The report of certification services may vary from a single word (certified, or just a logo or seal of the certifying agency), a pass/fail report, to a grading scale (from 5-100 points) with sub-scales and qualitative comments. Lizzeri's [1999] model indicates that a monopolist certifier provides coarsely graded (pass/fail) reports and earns rents on her finer private knowledge of the underlying quality of various goods and services. Competition in the certification market leads to provision of more informative audit reports. Consumers of

certification services benefit from greater competition in certification markets in the form of more informative certification reports and possibly lower prices³.

Price. The amount and the form of pricing of the certification services vary. Variations of form include pay-for-service vs. subscription over a period of time or for larger bundles of services of which certification is one component.

The debate in the audit literature has focused largely on the fact and perception of independence, with little attention paid to the possibility of markets trading off independence, experience, competence, prices and other features. For public policy on audit reforms (e.g., audit firm rotation) understanding of such trade-offs is important.

In Section 2.1 we examine a variety of private goods to document the range of certification services available in the economy. A broad definition of certification allows us to examine four kinds of services in the economy to which some aspects of audit services may be compared: (1) expert opinions based on written standards, as in grades of beef, (2) expert opinions in absence of written standards, as in the ratings of art and wines, (3) ratings given by lay people, such as for restaurants, and (4) meters such as Nielsen's for movies, books and music. We examine certification reports in Section 2.2 and the effects of independence by conducting a field experiment of a Baseball card certification market in Section 3.

2.1 Extent of Certification Services in the Broader Economy

In order to assess the availability of certification services in the economy, we selected a sample of 817 items sold online and offline during June 12 – July 25, 2004. Goods sold online were selected from eBay.com (400 items), and goods sold offline were selected from the Producer Price Index (PPI – 358 items) and Consumer Price Index (CPI – 59 items) published by

³ It is possible that certification quality could also improve though we are unaware of a formal model of the effort exerted by the certifier.

the U.S. Bureau of Labor Statistics (www.bls.gov). The “all categories” page on eBay listed all the items available for sale in 31 main categories, and a hierarchical structure with layers of sub-categories that list thousands of items at the bottom. We selected all items at the first sub-category level (e.g., sports cards) resulting in a sample of 400 items, except catch all sub-categories such as “other.” The PPI has a similar hierarchical structure with 15 main categories. We selected a sample of all items at the second sub-category level resulting in a sample of 358 items. The CPI has a flatter hierarchical structure, and in some cases there are no sub-categories (e.g., college textbooks). We chose all unique items in the CPI which were not in the PPI samples.

A variety of certification services are available for goods and services. The most formal and traditional service is the offer of an expert opinion about compliance with formal (written) standards. It is also possible to obtain an expert opinion in the absence of a formal set of standards, ratings given by lay people (with no formal standards), and a variety of audience/popularity meters which simply record the level of activity (e.g., bestseller book, music, and film lists).

Panel A of Table 1 indicates that for the 817 goods in the sample, we were able to find an expert who would provide a certification with compliance to a formal set of standards for 743 goods (91%). For another 59 goods (7%) we were able to find an expert who would provide an opinion, but without use of a formal set of standards (e.g., to rate sculptures, antiques, paintings, and gardening services). For the remaining items, we were able to find a lay person rating for 8 items (1%), and a meter for 3 items (0.5%). We were unable to find any kind of certification service for only 4 items (0.5%, e.g., parking services, tattoos and other body art, and sale of

coupons). Overall, for 814 or 99.5% of the goods in our sample, it was possible to find some kind of certification service.

Panel B of Table 1 specifies the available combinations of certification services. For 322 items (40% of the sample) the full range of certification services (standards with expert opinion, expert opinion only with no formal standards, lay opinion, and meter) were available. For 114 items (14% of the sample) only an expert opinion using a formal set of standards was available. For all the remaining items, more than one kind of certification was available. These results support Power's (1994) characterization of our society as being an "audit society." While Power ([1994]) focuses on the rise of auditing in the public sector in the U.K., the ubiquity of auditing in the economy suggests that the phenomenon extends to the private sector in U.S. and quite likely elsewhere. While the demand for certification in the public sector may be fueled in part by political ideology, demand for certification in diverse private areas of the economy appears to be driven by broader economic forces.

Insert Table 1 about Here

2.2 Developing a Reputation for Audit Quality Via Fineness of Audit Reports

No matter how thorough and effective the auditor is, he labours in vain if he cannot clearly convey the results of his efforts in a useful form to those who need to know them.

American Accounting Association,
A Statement of Basic Auditing Concepts, (1973).

One key vehicle for building a reputation for audit quality is providing informative audit reports to users of financial statements. In essence, the standard audit reports currently issued are

pass/fail⁴. Basic economic intuition about the optimal fineness of a grading scheme being of an intermediate level can be considered alongside Lizzeri's [1999]'s prediction that a monopolist certifier will produce low precision pass/fail reports. In the course of his work, auditor develops a detailed understanding of the quality of a company's internal control system, governance, accounting policies, estimates, and disclosure. A coarse grading system (pass/fail) conveys a bare minimum of the auditor's detailed understanding about the company to the shareholders. It also induces managers of companies to disclose to the auditors only the minimum necessary to get a pass rating from them.⁵

Dubey and Geanakoplos [2005, henceforth DG) model of grading schemes suggests that a coarse (pass/fail) a grading scheme conveys less information and induces less motivation for agents to exert effort to raise quality. A fine grading scheme introduces measurement error, and discourages agents with intermediate levels of talent. An intermediate level of fineness, e.g., a 3-10 points system, balances these two considerations by providing useful information and motivating agents to work harder. The DG model also suggests that the optimal grading scheme should create a small elite grade (i.e., a hard-to-get A), and that an absolute grading scheme (e.g., score > 85 percent is an A) strictly dominates grading on a curve.

We collected data on the fineness of rating standards in the government by visiting the websites of 80 federal government departments listed as having a standard-setting function in the most recent edition of a U.S. government publication of "Standards Activities of Organizations in the United States" (Toth [1996]) and succeeded in downloading electronic copies of standards

⁴ "The report shall state whether the financial statements are presented in accordance with generally accepted accounting principles (GAAP)... The report shall contain either an expression of opinion regarding the financial statements, taken as a whole, or an assertion to the effect that an opinion cannot be expressed." (http://www.pcaobus.org/standards/interim_standards/auditing_standards/index_au.asp?series=100§ion=110). Auditor is expected to add detail when the report is "fail."

⁵ Blackwell's fineness condition suggests that, in a game against nature, a finer report would be more useful to shareholders.

for 64 agencies (80 percent of 80). We recorded a summary of the type of standards set by each agency (e.g., the Department of Agriculture sets standards for food and farm products including tobacco). Toth [1996] provided data on whether the agency audits (or certifies) entities governed by its standards. We examined the websites and/or the standards to determine whether the agency provides a minimum standard (pass/fail) or a series of grades (e.g., beef grades provided by the US Department of Agriculture).

In the sample of 64, 53 (83 percent) of the federal agencies set only minimum (i.e., pass/fail) standards. The other eleven (17 percent) agencies issue a range of grades to differentiate the quality of goods. Since many of these agencies are not able to create an effective enforcement and deterrence regime (see Law [2005]), they set quality/grading standards to induce entities to use better ingredients, production methods, and accurate labels. Given chronic complaints about under-funding of the SEC (although it had its budget doubled after the SOX legislation), one might also expect to observe graded reporting standards in financial markets. Such grades exist for ratings of bonds and insurance but not for financial reports. Out of the three predictions of the DG model, all government agencies comply with one (setting an absolute level for a passing grade), and usually do not comply with the other two (a grading system of intermediate level of fineness that includes an elite grade). The USDA, with its hierarchy of grades for foods and grains including elite grades (e.g., U.S. Prime beef) is an exception that meets all three DG prediction for optimal grading systems.

Private sector standards and seals exhibit greater variability. Pass/fail standards (e.g., certification of electrical appliances by Underwriters Laboratory), multiple seals to signal different levels of e-commerce privacy (e.g., TRUSTe provides a separate seal to signal compliance with laws relating to children, Jamal et al. [2003], [2005]), 10-point rating scales

(e.g., for grading baseball cards) 100-point rating scales (e.g., Consumers Reports' ratings of cars), and even 3,000-point scale of ELO⁶ ratings for chess players coexist. In private markets rival standard-setting organizations compete by differentiating themselves along a variety of dimensions including fineness of reports and certification processes, Jamal et al. [2003]. Some organizations let their buyers choose the fineness and detail in the report for a price (e.g., the XYZ site mentioned in Section 3.0); a consumer who wants to buy a toaster can get a pass/fail report from Underwriters Laboratories, and can also get a more detailed (finer) report from Consumers Reports. In some online markets (e.g., for grading baseball cards) there is evidence that new entrants seek to differentiate themselves from established companies by providing a finer rating system, Jin, Kato and List [2004].

We chose the unregulated online baseball card certification market for a more detailed examination because it has four features: (1) It is a competitive certification market where twenty three certification agencies compete in an online environment. (2) Some certifiers in this market are independent whereas others cross-sell services. (3) We can develop clean measures of audit quality by using both process consistency and grading strictness to assess the effect of independence on certification quality. (4) We can observe other features besides independence and audit quality that market participants might value such as price, reporting precision, or additional dimensions of service (e.g., computer grading, having two graders grade a card).

Jin, Kato and List [2004] provide details of how the baseball card certification market operates. Professional Sports Authenticator (PSA) started providing online grading service in 1991 and currently dominates the market in volume. PSA rates cards on a 10-point scale at 1 point interval (e.g., 8, 9, 10). Each point on the scale also has a qualitative label (e.g., 10 = mint).

⁶ It is not clear if ELO has an upper limit; the maximum rating ever assigned to a player has been 2851.

PSA does not use a curve to grade, and does not create a super elite grade (the top grade of 10 is given to about 10% of cards graded). Without an elite grade, PSA's practice is consistent with two of the three predictions (intermediate fineness, absolute grade cutoffs) of the DG model.

Starting in 1999, the privately held Beckett Grading Service (BGS) has gained a significant share of the market and emerged as PSA's main competitor. BGS distinguishes itself from PSA in three ways: First, it uses a 10-point scale with 0.5 point increments, and provides sub-grades for centering, corners, edges and surfaces. Second, it uses a stricter grading standard, giving the highest grade of 10 to less than 0.1 percent of the cards it grades.⁷ It builds market awareness of BGS's grading strictness by periodically posting the distribution of grades issued by BGS (a BGS distribution showing is available to its customers at <http://www.beckett.com/grading/popreport.asp?action=summary>). Third, Beckett caters to different market segments by offering three brands of service at different prices: a standard BGS service, an elite vintage service (BVG), and a value product for more price conscious customers (BCCG). The actions of Beckett are consistent with all three predictions of the DG model.⁸ The attempt by Beckett to create multiple brands for different clienteles has no parallel in auditing.

The third major online rater of baseball cards, far behind the two leaders, is Sportscard Guaranty (SGC). It entered this market in 1999 and attempted to distinguish itself by introducing a 100-point scale (Jin et al.[2004]), apparently creating some market confusion. SGC started providing a nonlinear conversion table to translate its 100-point scores into a 10-point score and the associated qualitative labels. SGC does not provide additional sub-grades. Krislov (1997)

⁷ PSA also issues a population report but a user has to pay a \$4.95 fee to access the report. The population report is also bundled with PSA's pricing guide which only provides prices of cards graded by PSA.

⁸ Publication of the empirical distribution of grades could indirectly set a "norm" of what grades should be, even if the formal grading scheme does not require use of a grading curve. Publication may also influence the quality of cards offered for grading to a service, and thus change the substantive implications of a fixed grade distribution.

provides several examples where countries (and companies) use incompatible standards to differentiate their products from their rivals'. The problems SGC faced in introducing a 100-point scale are consistent with DG's prediction that high levels of fineness are suboptimal due to measurement and interpretation errors.

Twenty other baseball card certification services vie for customers (see Table 2) using the 10-point scale in increments of 1 or 0.5, some adding extra grade categories (e.g. pristine) allowing some 10-22 possible grades. The creation of extra elite grades is consistent with DG's prediction. None of these twenty sites uses a grading curve, and all have a reporting scheme of intermediate fineness. In contrast to the government agencies, the practices of the twenty three private baseball card graders generally conform to the DG predictions.

Some government agencies have experimented with more detailed than pass/fail reports. For example, California changed its restaurant hygiene reporting system from pass/fail to a letter grade system (a score of 90-100 is A, 80-89% is B, 70-79% is C, and a score from 60-70% is reported as a number. Two consecutive ratings with a score below 60 is an F and the restaurant can be shut down for failing the inspection). All restaurants are required by law to post these grades in a prominent place where they can be easily seen by consumers. A field study by Jin and Leslie (2003) has documented a shift in consumption patterns with restaurants graded A (C or lower) reporting a significant increase (decrease) in sales. They also document a significant decrease in visits to doctors and hospitals for food related illnesses after the change in hygiene reports suggesting that the change in grading scheme improved the hygiene of California restaurants.

Press reports (e.g., Tergesen [2005]) indicate that Moody's developed a new line of business to certify the internal control systems of hedge funds with a report on a five point scale.

In September 2006, Soring Capital was the first hedge fund to be rated by Moody's (receiving a rating of 4). Moody's markets the service to hedge fund clients who will pay for the rating service just as companies pay their auditors. Morningstar may also offer a similar internal control rating service, but it plans to charge investors directly for its reports and thus hopes to be more independent of the hedge fund management, Tergesen [2005].⁹

In unregulated markets, certifiers set proprietary standards and the standard setters often certify compliance (see Jamal et al. [2003] for a description of competition in the e-commerce privacy market). This results in competition among standards-certification bundles. Certification services may also be provided by third parties licensed by the standard-setter (e.g., ISO certification). In the pre-SEC era, individual accounting firms often initiated both accounting and auditing innovations. For example, the audit of U.S. Steel by Price Waterhouse prompted the development of consolidated financial statements, as well as the long-form audit report, Vangermeersch [1986]. Certifiers of private goods such as automobiles, wines and baseball cards also develop reporting standards, often using scales of varying fineness.

For many goods and services information aggregators (e.g., MSN Auto at <http://autos.msn.com>) create composite reports from ratings provided by multiple certifiers (see Figure 1). MSN Auto report on Honda Accord, for example, provides an average of ratings from MSN's designated experts, the individual reports from each expert (on a 10-point scale), a user rating from volunteer members of the public, and an excerpt of an independent expert opinion from the *Consumer Reports* (on a 100-point scale). The *Consumer Reports* excerpt also provides

⁹ Klein (1997) models the determinants of whether the buyer or the seller should pay for certification. The competition between Moody's and Morningstar may provide an interesting opportunity to test the implications of Klein's model. In the audit literature, some preliminary evidence about having investors pay directly for audit services has been provided by Mayhew and Pike (2004).

sub-ratings for acceleration, accident avoidance, comfort and convenience, on-the-road fuel economy (all on a five-point scale) and a qualitative discussion of the car.

Similarly, an information aggregator for wines, www.vintages.com, tabulates ratings of listed wines by Robert Parker Jr. (100 point scale), James Suckling (100 point scale), Jancis Robinson (20 point scale), Tom Wilson (10 point scale), and Shari-Mogk-Edwards (5 point scale). In addition, qualitative comments from one or more of these experts are included. There is no evidence that consumers are confused by the provision of multiple assessments on varying scales.

Auditing standards currently require a report which is essentially no more nuanced than a pass/fail report. The size of the audit firm, care in client selection, and in monitoring compliance with GAAP may be the basis of auditor reputation (see the Price Waterhouse audit fee premium reported in Simunic [1980]). Given that a major objective of SOX is to convey information to shareholders about internal controls (currently provided as a pass/fail certificate), a graded audit report with sub-grades for internal controls, governance, quality of accounting methods and quality of disclosure, may be more useful to the shareholders.

In the pre-1930s unregulated era audit firms issued either a short or a long-form report, Brief [1987]. The short-form had a standardized pass/fail format of the type used currently, but could be as short as a single word (“Certified”), a line (“I certify the above statement is correct”), or a paragraph, Himmelblau [1927, pp. 12-15]. The long-form reports included comments on the propriety of accounting methods, descriptions of audit procedures carried out, and occasional representations about the market values of certain current assets as being in excess of cost. The text of long reports was client specific and varied for the same client over time, Himmelblau [1927]. The Price Waterhouse 1902 report on U.S. Steel set a new standard by providing

additional information on valuation issues such as fixed asset capitalization policy and depreciation, inventory valuation, revenue recognition, and audit procedures conducted to verify cash, Vangermeersch [1986, p. 24]. A century later, the richness and detail of these reports has been replaced by boiler plate text which conveys little of the auditor's detailed knowledge to the reader. One can argue that such knowledge could be of use to shareholder and board decisions. The standard pass/fail report is closer to practices of government agencies; and is at variance with the finer reports suggested by the DG model as well as with the dominant practice of private certifiers.¹⁰ Regulation of auditing sets a floor on the precision of audit reports, and audit firms do not find it in their interest to aim any higher.

3.0 Part II: Role of Independence in Certification Markets

Independence is an essential auditing standard because the opinion of the independent accountant is furnished for the purpose of adding justified credibility to financial statements which are primarily the representations of management. If the accountant were not independent of the management of his clients, his opinion would add nothing.

E.B. Wilcox, The CPA Handbook (1952)

There is no concept more central to the notion of audit quality than the concept of independence, Mautz and Sharaf [1961]; DeAngelo [1981]. The potential for conflict of interest is thought to completely undermine the value of an audit. To obtain evidence on willingness of people to carry out transactions under potential for conflict of interest, we approached a website which we call XYZ for information about its experts who provide opinions on Internet. XYZ is an official partner of eBay and has 128 experts who offer opinions on 1,850 items for a fee of \$9.95 (basic service) or \$29.95 (enhanced service) per assessment. These experts do not follow any written standards, nor are their opinions based on reference to them. For some of its experts,

¹⁰ Implementation issues such as legal liability are beyond the scope of this paper (see Bush et al., 2006). These issues can be dealt with by providing a safe harbour, providing these reports only to the audit committee (and not directly to investors), or by creating a grading standard like that for U.S. Prime beef and then providing a certification for companies who have high quality financial reporting.

the site provides a description, accreditation by professional bodies, educational background, and relevant business and other experience.

Since the list of XYZ experts is rotated periodically, we could not be sure that all their experts were listed at any given time. We obtained data on accreditation, education, and business interests of all listed experts who provided opinions on the site as of July 12, 2004 (see Table 3). Of the 128 experts, 50 (39%) held formal accreditation from a professional body, 32 (25%) had relevant educational credentials (e.g., opinions on art being offered by persons having an advanced degree in art history), 101 (79%) ran a business involved in the activity they were providing opinions on (e.g., a carpet shop, or a wine store). Only 12 experts (9%) did not report any formal credentials and indicated that they were hobbyists, interested in activity on which they opined.

Insert Table 3 about Here

3.1 Implications for Auditing

Almost four out of every five eBay experts are also in the business of buying/selling the goods or services on which they provide expert advice. Likewise other studies have documented the pervasiveness of potential conflict of interest in various areas of the economy (e.g., doctors, financial analysts, investment bankers – see Moore et al. [2006]). Due to funding and other social ties with the audit firms, auditing researchers may also have potential conflicts of interest in their work. The data suggest that potential conflict of interest in the market for certification is more of a norm, than an exception.

Provision of consulting services to audit clients is thought of by some as being unacceptable because of potential conflicts, Francis [2004]. This attitude in regulated domain of

auditing is in contrast to the data from the unregulated domains for private goods and professional services where consumers freely choose whether to pay cash for the opinions/assessments of self-proclaimed experts with known conflicts of interest. Is it possible that the consumers in unregulated markets are able to protect themselves against self-serving advice, and thus find it worthwhile to pay for such advice?¹¹ Could the customers of audit services do the same? If not, why? If yes, is the demand for total elimination of potential conflict of interest in auditing a case of regulatory overkill? We return to this issue in Section 4 (after reporting the results of a field study of online certifiers of baseball cards).

3.2 The Baseball Card Certification Market

In the thriving online baseball card certification market, prospective owners/sellers can hire a certifier to grade and certify the authenticity of their cards. As products, baseball cards are quite different from financial reports. Yet, the framework for analysis of certification services points to some common elements (motivations of buyers and sellers, roles of standards and reputation, and forms of reports and pricing) that may allow us to gain insights into functioning of such markets through cross-comparisons.

Issuers of securities hire auditors to add credibility to their financial statements. Empirical examination of the consequences of independence on audit quality is complicated by the difficulty of observing and controlling these characteristics. In the baseball card certification market, we have an opportunity to observe pure auditors (who only certify baseball cards) as well as cross-sellers who sell a range of additional related services. Each certifier posts a written set of proprietary standards and a schedule of prices, including premiums for faster service.

¹¹ The XYZ website has now set up a trusted marketplace where all items available for sale have been authenticated by a XYZ expert. On one hand this seems to increase the potential for conflict of interest, yet XYZ seems to think this is a useful service for their customers and that it will be profitable for XYZ as well.

Competition among rating agencies involves both setting of standards as well as provision of certification services. Providing two additional services (creating pricing guides, and acting as a dealer in the market) create potential for conflict of interest for graders. The baseball card certification market allows us to directly observe a cleaner measure of audit quality (grading strictness and test-retest process consistency). The baseball card market also provides an interesting opportunity to observe how unregulated certification providers may try to distinguish themselves by adding new service features such as value pricing (e.g., TFA), computer grading (CTA), having two experts grade each card (PRO), and even letting each customer choose their own grade (MINT) (see Table 2)¹².

The potential for a race to the bottom among certifiers who may be willing to lower their quality in competing for clients is a frequent theme in accounting (see Dye and Sunder [2001] for a discussion of such fears). Since increasing auditor independence is a major objective of SOX, it would be interesting to know how an unregulated market for certification services reacts to independence and potential for conflict of interest as well as grading strictness (leniency) and other service differentials.

3.3 Baseball Card Rating Standards and Potential Conflict of Interest

Accugrade (ASA), the first entrant in this market in 1988, pioneered the use of a 10-point rating system, and developed a tamper-proof plastic casing in which graded cards can be sealed. It no longer appears to be in the grading business, earning royalties from other rating services that use its patented package. Professional Sports Authenticator (PSA) entered the market in 1991 and currently dominates the market for single and rookie baseball cards as well as basketball and football cards. PSA uses its dominant status to collect a \$99 membership fee for

¹² Note that the limited parallel we use here is between the two certification markets—for financial reports and baseball cards—and not between the financial reports and baseball cards themselves.

access to its rating service.¹³ PSA also offers a magazine and a price guide, and runs card shows where collectors can meet experts, get their cards graded on site, and talk to PSA staff. PSA is a part of publicly-traded Collectors Universe (NASDAQ: CLTC) which offers rating services for coins, stamps, autographs and music. Beckett Grading Service (BGS), the second major grader of baseball cards, is a private company that has gained a significant market share in the card rating market. In addition to rating cards, BGS offers a magazine, price guide (which provides estimates of prices for baseball cards graded by BGS as well as by its competitors, and for ungraded cards), and card shows. BGS rates cards for racing, several sports (baseball, football, and basketball), comic books, and action figures (such as Pokemon and Digimon).

To assess the value of certification services, we collected a matched pair data set of baseball cards traded on eBay. We selected 321,045 baseball cards traded on eBay during August 19-September 3, 2004. We partitioned the cards by the decade of issue (e.g., 1930's, 1940's ... 2000's, etc.), single or rookie, and whether they had been graded. Of the 321,045 cards in our sample, 272,399 (85 percent) were singles, of which 31,778 (11 percent) had been graded. Of the 48,646 rookie cards, 12,290 (25 percent) had been graded (see Table 4). We drew a stratified (by decade of issue) random sample of 1,000 graded rookie cards for analysis.¹⁴

Insert Table 4 about Here

Market Share. Table 5A shows the estimates of market shares of six firms in the rookie card market on the basis of our stratified sample of 1,000 graded cards traded. PSA and the three

¹³ During a return visit to the PSA website in May 2006, we found that they have made the membership fee optional. Membership has certain privileges (free card grading, magazine) but is no longer required to access the PSA grading service.

¹⁴ We focus on rookie cards rather than singles because rookie cards are of higher value, they are more likely to be certified, and there are more certification agencies who provide services for rating rookie cards. The singles card market is dominated by PSA with a 78% singles market share; it has only 36 percent share in the rookie card market.

levels of services provided by Beckett have about 35 percent market share each, with GEM claiming 16 percent market share.

Insert Table 5A about Here

Fineness. The grading schemes used 9 to 22-categories on a 10-point scale with 1 or ½ point increments (with an occasional variations such as a plus sign or a “pristine” grade at the top of “mint”, Table 2). One grader (SGC) uses a 100-point scale, and provides a (non-linear) table to convert its grades onto a 10-point scale.

Strictness. We have two bases for assessing the strictness of grading. First, Jin et al. [2008] conducted a field experiment in which they gave the same 212 baseball cards to three online graders (PSA, BGS and SGC), and to three offline dealers. The average scores were 8.5 for BGS and two offline dealers, 8.7 for PSA and one offline dealer, and 8.9 for SGC. They concluded that BGS ratings used a tighter cut-off point and were more precise (see Table 5B based on Jin et al.’s Table 2). Jin et al. classified BGS as a strict grader and PSA and SGC as medium graders. There was a statistically significant difference in grades given by BGS on one hand, and by PSA and SGC on the other. The difference between PSA and SGC was not significant. Their study assessed both grading strictness (average grade) and a process measure of audit quality (consistency and test-re-test reliability). Certification agencies that were stricter were also more consistent in their ratings. We use the Jin et al. [2008] results to classify BGS as a strict grader, and the other two (PSA, SGC) as medium graders.

Insert Table 5B about Here

For graders not covered in the Jin et al. [2008] study we use the average grades observed in our samples to classify them by strictness of grading. In our sample of 1,000 rookie cards, GAI has an average rating of 9.1 which is very similar to that of PSA (8.9) so we classify GAI as a medium grader. Two graders GEM (average grade of 10) and Beckett's brand BCCG (average grade of 9.8) are classified as lenient graders (see Table 5A). The remaining graders have very few occurrences in our sample, so we do not analyze them further and do not assign them a strictness classification. Since sellers self-select which cards they send to each certification service, this measure suffers from the self-selection problem common to many archival studies. Frequency of grades provides additional corroboration to the controlled experiment run by Jin et al. [2008].

Pricing. The per-card fee varies in the range \$2-50 across the 23 graders with a higher price for faster turnaround. Similarity of posted prices suggests that the competition is driven primarily by non-price variables (see Table 2).

3.4 Data Analysis

For each graded rookie card in our sample of 1,000, we found a matched (by player, card maker, year of issue, rookie) un-graded card. Values of unrated cards were obtained from Beckett Baseball Card Monthly Guide (August 2004, Issue #234-online).¹⁵ The Beckett price guide provides a range of price estimates for each card with high, average, and low values. For each rated card sold on eBay, we recorded the player, year, card maker, grader, grade received, buyer reputation, seller reputation, number of bids, and selling price.

¹⁵ Dr James Beckett issued the first Baseball card pricing guide in 1979 by collecting information from baseball card dealers. In 1984 Beckett Publications was formed primarily as a seller of pricing guides. Over time other services were added including a magazine and sports card grading services. Then the company diversified into developing other pricing guides (e.g., football, and comic books) as well. Beckett now sells an annual pricing guide, a monthly pricing guide, and also provides an online pricing guide.

Gross and net (of the cost of certification) returns of baseball cards rated by six major certification service providers are shown in Table 6. All six certification services yield a positive gross return (selling price of a rated card exceeds the average estimate of the value for the same card in the Beckett pricing guide). However, the return net of the cost of certification is positive for the strict grader (46 percent for BGS) close to zero for the medium graders (2 percent for GAI, 4 percent for SGC, and -2 percent for PSA), and negative for the lenient graders (-14 percent for BCCG, and -23 percent for GEM). Older cards earn a higher return than the more recently issued cards, probably because older cards are more valuable relative to the fixed price of grading. Cards with a rating of 10 usually get a positive net return (except for the lenient graders BCCG and GEM), whereas ratings of 9 are likely to generate lower, and often negative, net returns. Overall, it is more profitable to get a 9 rating from a strict grader (BGS or GAI) than to get a 10 from the market leader (and medium grader) PSA.¹⁶

 Insert Table 6 about Here

To assess the relationship between the return from grading, and various characteristics of cards and grader identity, we estimated the following regression equation (Table 7):

$$R_i = \alpha_1 + \beta_1 \text{Year}_i + \beta_2 \text{Grade}_i + \beta_3 \text{GPSA}_i + \beta_4 \text{GBGS}_i + \beta_5 \text{GBeckett}_i + \beta_6 \text{GGAI}_i + \beta_7 \text{GGEM}_i + \beta_8 \text{GSGC}_i + e_i$$

Where:

R_i = net Return from getting a rookie baseball card graded. This is calculated as selling price of graded card on EBay – (book value of same card in Beckett pricing guide

¹⁶ Note that the validity of comparisons of these data across graders and vintages is limited by the selection effects. Older cards, being rarer and more valuable on average, may be more likely to be presented for certification, especially by the stricter graders. Data from Jin et al.'s field experiment yields cleaner comparisons.

+ minimum cost of certification service) divided by book value of same card in Beckett pricing guide.

Year = Actual year when a baseball card was issued

Grade = score assigned by an independent certification service on a scale from 1 to 10, increasing in increments of 0.5.

GPSA = Grader is PSA; GBGS = for BGS, GBeckett = for BECKETT (other than BGS); GGAI = for GAI, GGEM = for GEM; GSGC = for SGC. These are the top 6 graders. For each individual top grader, the test is whether return to the specific grader is greater than return from all other 17 graders.

Estimates in Table 7 indicate that the year, grade and identity of grader have a significant relationship with the selling price of a rookie baseball card. More recent, presumably lower valued, cards earn a lower net return (coefficient is negative, $p < 0.001$) because the cost of service is the same. Likewise, the grade has a positive significant coefficient ($p < 0.001$), a higher grade from a third party certification service is associated with a higher price. Certificate from a strict grader has a significant positive coefficient on price (BGS: $p < 0.001$). Certificate from a moderate grader has a lesser but still significantly positive coefficient for both the dominant grader (PSA: $p < 0.01$), and for two smaller graders (SGC: $p < 0.05$, GAI $p < 0.001$). Certificate from a lenient grader has a negative coefficient on price (GEM: $p < 0.001$). The coefficient is also negative for Beckett's value brand, though it is not statistically significant. It appears that a 10 from an easy grader does not earn financial returns; instead it signals low quality and generates net losses to the owner. Market is not fooled and there appears to be no race to the bottom.

Insert Table 7 about Here

It is not clear why Beckett has multiple brands though there is some speculation on Internet baseball discussion boards that Beckett has set up rival brands to allow grade inflation (for example, see www.seanet.com/~brucemo/card_articles/pro_grading.htm). The market appears to price Beckett's strict grading system (BGS) differently from the more lenient grading system (BCCG). There is no evidence that the multiple brands do any harm to the flagship BGS brand. Beckett seems to have been successful in implementing a market differentiation strategy and tapping different segments of customers with different brands (just as auto manufacturers sell different brands to various segments of their markets). Beckett claims to be offering a discount service for more price conscious customers, but the market suspects that Beckett is providing a more lenient grading standard in addition to lower prices in its discount brands, and prices each service accordingly.

3.5 Implications for Auditing

Care is necessary in drawing cross-market implications for auditing. Baseball cards are physical objects of substantive interest, while financial reports are symbols representing complex entities of substantive interest to buyers of certification services. In both cases, buyers are willing to pay for others' expertise, in presence of varying extent of standards as point of reference for certification. In both cases, buyers rely on the knowledge and reputation of the certifiers who compete for fee revenues. In both cases, individual certifiers choose the portfolio of certification as well as other services they may offer to their customers. Although audit of publicly-held firms is mandated by the U.S. federal securities laws, there is evidence that the demand for independent audits of public firms, originating in agency considerations pre-dated the regulatory mandate, Chow [1982].

A key regulatory instrument for improving financial reporting has been to promote the independence of auditors, in significant part by prohibiting and restricting them from providing advisory services to their audit clients. In the market for grading baseball cards seventeen of the 23 graders offer only certification services (i.e., are independent), whereas six offer additional services related either to baseball cards (e.g., pricing guides, dealers, magazines) and/or other cards, coins, stamps and other collectibles. The market is dominated by two firms which cross-sell other services (PSA and Beckett). Market prices reflect large premiums for cards certified by PSA or Beckett. The remaining firms, including all 17 independent ones, have small market shares. Even among the smaller players, the dominant certification agencies are SGC and GAI, both of whom cross-sell services. Only one independent rating agency (GEM) has any significant market share. Ironically, GEM gets its market share by providing rampant grade inflation where virtually all rookie cards graded by GEM get a grade of 10 (see Table 5A) and yield a negative market return (see Table 7). It is clear that being independent is neither necessary nor a sufficient condition for being successful in the sense of having a larger market share or being valued in this unregulated certification market. Virtually all pure (independent) certification agencies do poorly in this market, and may well go out of business as the market consolidates over time. Recall that the partners of Arthur Andersen LLP rejected Paul Volcker's offer to save the firm if they agreed to drop all its non-audit services.

If markets for certification services value not only independence but also other characteristics that may weaken certifiers' independence (such as better information derived from participation in market as dealers), their interests would be served better by finding an appropriate balance among the preferences for various characteristics. We should not assume, *a priori*, that this balance will necessarily be achieved by yielding primacy to independence over

all other considerations. As we saw in Section 3.0 (expert opinions on EBay), unregulated markets value opinions from experts who are immersed in a market (and are presumably better informed and more competent). Perhaps the potential for conflict of interest does not override the advantage of being better informed in such situations. Whether this is true in certification of financial reports deserves to be a subject of research and analysis and need not be assumed away by regulatory fiat.

There is skepticism in accounting literature about the ability of an unregulated audit (or accounting) market to function properly, with a race to the bottom being a major concern (see Dye and Sunder [2001] for arguments on the two sides). We observe the opposite in the market for baseball card certification. Beckett entered this market late, but made major inroads and gained a large market share by emphasizing grading strictness, not grade inflation. As shown in Table 7, lenient graders who engage in grade inflation earn a poor reputation and yield a negative return for buyers of their services. Recent studies of audit pricing also indicate that auditors who are industry experts, and especially city-specific industry experts, command premium audit fees and also restrain the abnormal accounting accruals of their clients (i.e., they are more strict auditors, see Francis et al.[2005; 2006]). In other markets for e-commerce privacy seals, Jamal et al. [2003] found that the higher quality certification providers dominate the market.

It would be troubling to observe grade inflation (low audit quality) in certification markets. In the baseball card certification market, the most obvious potential for grade inflation occurs in an agency called MINT where sellers of cards are asked to specify their own grades (see Table 2). The empirical distribution of grades indicates that another certification agency (GEM) also engages in grade inflation (see Table 5A). Both MINT and GEM are independent baseball card certification agencies. None of the cross-selling certification agencies engage in

grade inflation. The data suggest that the cross-selling of services might inhibit, not promote, grade inflation in certification markets. Poor profitability and market position seems to induce grade inflation.

Prima facie, it is difficult to question the value of auditor independence for the efficient functioning of the capital markets, especially in the post-Enron-Arthur Andersen world. Yet, Antle et al. (1997, p. 28) have argued:

Taking a holistic view, we have found that auditors have many incentives to protect their independence. Legal liability is significant, and any firm that would damage its independence risks an avalanche of litigation. Auditors' have substantial investments in reputations, audit technology and methodology, and directly in their financial stakes in accounting firms. We have found no evidence that the supply of non-audit services threatens auditor independence, and there is a strong intuitive case that accounting firms create value by capturing economies of scope between audit and non-audit services.

Empirical investigations of the audit market have not been supportive of the hypothesis that provision of advisory services to audit clients impairs the independence of auditors (see footnote 1 and Francis et al. [2005]; [2006], or have a bearing on audit fees, Abdel-Khalik [1990]. Experimental study of Dopuch and King [1991) concluded: "...policy makers who favor proposals to prohibit audit firms from providing both MAS and verification services to the same client should contemplate whether the prohibition will have an adverse effect on the market structure of the audit industry" (p. 89).

Our own findings about the wide-spread prevalence of conflicts of interest in the unregulated baseball certification market raise doubt about the validity of the common sense assumption about the value of independence of the certifier that under girds the current regulatory regime. We are hard-pressed to find an explanation for this apparent paradox. This evidence is consistent with Simunic's [1984] suggestion about the economies of scope arising from information links between audit and consulting. Zhang [2004] also models the provision of auditing and consulting services to the same client as a matter of trade-off between the benefits

of better-informed auditor judgment (on the basis of what the auditor may learn in the process of providing the consulting services) and the loss from conflict of interest in making those judgments. It seems plausible that while the regulatory emphasis on auditor independence may cut the profits of auditors, its consequences for the credibility and informativeness of certification are ambiguous. Antle and Demski's [1991] analysis of the role of contracting frictions in presence of information externalities between multiple services leads them to call for: "...an expansive theory of auditing, one that addresses audit procedures, organization of the audit, product bundling, contracting, and competition" (p. 20).

Audit firms appear to have engaged in poor enforcement of GAAP, especially in the 1990's and regulation does not appear to be effective in preventing poor enforcement. Perhaps the lack of profitability and the lack of a distinctive capability lead to grading leniency and poor enforcement. A more careful analysis of why auditing became a commodity in the late 1980's may help us understand the collapse of grading strictness in auditing during the 1990's. Support for blaming the cross-selling of services for the collapse of proper auditor enforcement of GAAP appears weak. The prosperity of audit firms may be the best antidote to client pressure and grade inflation.

4. Concluding Remarks

We report evidence that certification services are widely available in unregulated private good markets, including certification by experts who base their opinion on formal written standards, experts who do not use formal standards, lay opinions, and meters. Power's ([1994; 1999]) findings on the demand for audit in the public sector are corroborated; the phenomenon of "audit explosion" pervades the private sector too.

Second, we examine the fineness of certification by government agencies and private certification services. Government agencies usually issue pass/fail reports. Private certification reports use a finer scale (5, 10, or 10 points) and often include some qualitative commentary and sub-scales, breaking down the overall rating into various sub-components. Often multiple certification reports are available for the same product. The pass/fail scheme in boiler plate format of financial audit reports appears to be driven by their mandated demand, and focused on preventing litigation rather than providing information for decision making by shareholders and others. Recent appearance of un-mandated provision of graded reports on the quality of internal control of hedge funds by Moody's suggests that lifting of the mandatory requirement may give rise to demand for finer gradations in financial audit reports too.

Third, we conduct an eBay field study of certification of baseball cards where some agencies are pure auditors while others cross-sell related services. We find that the market is dominated by cross-sellers, and pure (independent) certifiers struggle to find customers in this market, resorting to grade inflation and low quality service. Strictness of grading pays, independence does not, and the two, surprisingly, do not go together. These baseball card market findings are consistent with the widespread demand for expert opinion in the economy, even in the presence of potential for conflict of interest. There is no race to the bottom; providers of low quality standards appear to fail in the marketplace. Independence is neither necessary nor sufficient for attaining grading strictness. It appears that profitability (or lack thereof) of the audit firm may be the key determinant of grading strictness.

In much of audit literature and policy deliberations, independence has been taken as *sine qua non* of high quality audit. There may well be special conditions associated with audit of financial reporting that make independence essential for an auditor to deliver quality service. If

so, it seems reasonable that such special conditions, and their logical and empirical links to independence, be identified and established. The evidence presented in this paper points to market demand for unregulated certification services provided by agents known not to fulfill the independence condition. Just as a single black swan is sufficient to falsify the proposition (held inviolable in Europe over the millennia until colonists arrived in Australia, Taleb [2007]) that all swans are white, these data seem adequate to question the general proposition that independence is a prerequisite for audit quality. Of course, it is always possible that there exist special conditions associated with financial reporting that do, in fact, make independence necessary. However, to the best of our knowledge, such argument has not yet been made and remains a challenge. Until we have such an argument, a question mark hangs on the assumption of the necessity of independence.

There are many differences between the regulated market for audit services and unregulated certification markets for other goods and services. The communication and negotiation process that precedes the issuance of the audit certificate, the presence of a diverse group of stakeholders interested in the audit report, and the opacity of audit quality, are some obvious differences. The present study is only a start at attempts to disentangle the consequences of regulation in audit markets through the use of cross-market studies. Studies that account for differences and similarities across markets in analyzing certification services and their regulation may help us better understand the existing audit regime, and find ways of making it more effective.

References

- Abdel-Khalik, A. R. 1990. "The Jointness of Audit Fees and Demand for MAS: A Self Selection Analysis," *Contemporary Accounting Research* (Spring): 295-322.
- Antle, Rick, and Joel S. Demski. 1991. "Contracting Frictions, Regulation, and the Structure of CPA Firms," *Journal of Accounting Research*U, Vol. 29 Supplement, pp. 1-24.
- Antle, Rick, Paul A. Griffin, David J. Teece, and Oliver E. Williamson. 1997. "An Economic Analysis of Auditor Independence for a Multi-Client, Multi-Service Public Accounting Firm," The Law & Economics Consulting Group, Inc.
- Ashbaugh, H., R. LaFond and B. Mayhew. 2003. Do Non-Audit Services Compromise Auditor Independence? Further Evidence. *The Accounting Review*, 78, 611-639.
- Brief, R. 1987. Corporate Financial Reporting At the Turn of The Century. *Journal of Accountancy*, Centennial Issue, May: 142-157.
- Burger, W. 1984. *United States v. Arthur Young & Co.*, Vol. 82:687, Supreme Court of the United States.
- Bush, Tim, Stella Fearnley and Shyam Sunder, "Auditor Liability Reforms in UK and US: Who Benefits?" First Draft, April 19, 2006.
- Carey, J. L. 1970. The rise of the accounting profession: To responsibility and authority, 1937–1969. New York: American Institute of Certified Public Accountants.
- Chaney, Paul K., Debra C. Jeter, and L. Shivakumar. 2004. "Self-Selection of Auditors and Audit Pricing in Private Firms," *The Accounting Review*79:1, pp. 51-72.
- Chow, Chee W. 1982. The Demand for External Auditing: Size, Debt and Ownership Influences. *The Accounting Review*, 57:2(April), 272-291.

- DeAngelo, L. 1981. Auditor Independence, 'Low Balling', and Disclosure Regulation. *Journal of Accounting and Economics*, August, pp. 113-127.
- DeFond, M., K. Raghunandan and K. Subramanyam. 2002. Do Non-Audit Service Fees Impair Auditor Independence? Evidence from Going-Concern Audit Opinions. *Journal of Accounting Research*, 40, 1247-1274.
- Dopuch, N. and R. King. 1991. "The Impact of MAS on Auditors' Independence: An Experimental Markets Study," *Journal of Accounting Research*, Vol. 29, Studies on Accounting Institutions in Markets and Organizations, pp. 60-98.
- Dopuch, N., R. King and R. Schwartz. 2003. Independence in Appearance and in Fact: An Experimental Investigation. *Contemporary Accounting Research*, 20, 79-114.
- Dubey, P., and J. Geanakoplos. 2005. Grading in Games of Status: Marking Exams and Setting Wages. Cowles Foundation Discussion Paper No. 1544R. Yale University.
- Dye, R. 1990. "MAS Spillover Effects on Auditor Independence." Working Paper.
- Dye, R., and S. Sunder. 2001. Why Not Allow FASB and IASB Standards to Compete in the U.S.? *Accounting Horizons*, September 15(3): 257-271.
- Elliott, R.K. 1998. *Assurance Services and the Audit Heritage*, *Auditing: A Journal of Practice & Theory*. 17, 1-7.
- Francis, J.R. 2004. What Do We Know About Audit Quality? *The British Accounting Review*, 36, 345-368.
- Francis, J.R., K. Reichelt, and D. Wang. 2005. The Pricing of National and City Specific Reputation for Industry Expertise in the U.S. Audit Market. *The Accounting Review*, 80 (Jan): 113-136.

- Francis, J.R., K. Reichelt, and D. Wang. 2006. National Versus Office-Specific Measures of Auditor Industry Expertise and Effects on Client Earnings Quality. Working Paper, University of Missouri.
- Frankel, R., M. Johnson and K. Nelson. 2002. The Relation Between Auditors' Fees For Non-Audit Services and Earnings Management. *The Accounting Review*, 77, 71-105.
- Gibbins, M., and K. Jamal. 2006. Management of Public Accounting Firms: The More Things Change, The More They Stay the Same? *CA Magazine*. August: 41-45.
- Himmelblau, David. 1927. *Auditors' Certificates*. New York: The Ronald Press Co.
- Jamal, K., M. Maier, and S. Sunder. 2003. *Privacy in E-Commerce: Development of Reporting Standards, Disclosure and Assurance Services in an Unregulated Market*, *Journal of Accounting Research*. Vol. 41 No. 2, 285-310.
- Jamal, K., M. Maier, and S. Sunder. 2005. *Enforced Standards Versus Evolution by General Acceptance: A Comparative Study of E-Commerce Privacy Disclosure and Practice in The U.S. and The U.K.*, *Journal of Accounting Research*. Vol 43 (1), March: 73-96.
- Jin, G., and P. Leslie. 2003. The Effects of Information on Product Quality: Evidence From Restaurant Hygiene Grade Cards. *Quarterly Journal of Economics*, 118(2), May, 409-451.
- Jin, G., A. Kato and J. List. 2004. "Evaluation of Professional Certification Markets: Evidence from the Field," Working Paper, University of Maryland.
- Jin, G., A. Kato and J. List. 2008. "That's News to Me! Information Revelation in Professional Certification Markets," *Economic Inquiry* (April): 1-19.

- Kinney, W.R., Z. Palmrose and S. Scholz. 2004. *Auditor Independence, Non-Audit Services, and Restatements: Was The U.S. Government Right?* *Journal of Accounting Research*. 42(3): 561- 588.
- Klein, D.B. 1997. Trust For Hire: Voluntary Remedies for Quality and Safety. In *Reputation: Studies in the Voluntary Elicitation of Good Conduct*. University of Michigan Press. Ann Arbor, MI.
- Krislov, S. 1997. *How Nations Choose Product Standards and Standards Change Nations*. University of Pittsburgh Press: Pittsburgh, PA.
- Larcker, D., and S. Richardson. 2004. Fees Paid To Audit Firms, Accrual Choices and Corporate Governance. *Journal of Accounting Research*, 42, 625-658.
- Law, M.T. 2005. How do Regulators Regulate? Enforcement of the Pure Food and Drugs Act, 1907-38. Working Paper, University of Vermont.
- Lizzeri, A. 1999. Information Revelation and Certification Intermediaries. *The Rand Journal of Economics*. Summer 30 (2): 214-231.
- Merino, B. D., A. G. Mayper, and R. S. Sriram. 1994. "Voluntary Audits in New York Markets in 1927: A Field Study," *Journal of Business, Finance and Accounting* 21:5, pp. 619-642.
- Mautz, R.K., and H.A. Sharaf. 1961. *The Philosophy of Auditing*. American Accounting Association Monograph No. 6. Sarasota, FL.
- Mayhew, B., and J. Pike. 2004. Does Investor Selection of Auditors Enhance Auditor Independence. *The Accounting Review*, 79: 797-823.

- Moore, D.A., P.E. Tetlock, L. Tanlu and M.H. Bazerman. 2006. Conflicts of Interest and the Case of Auditor Independence: Moral Seduction and Strategic Issue Cycling. *Academy Of Management Review* 31(1): 10-29.
- Power, M. 1994. *The Audit Explosion*. London: Demos.
- Power, M. 1999. *The Audit Society: Rituals of Verification*. New York: Oxford University Press.
- Simunic, D.A. The Pricing of Audit Services: Theory and Evidence, *Journal of Accounting Research* 18 (Spring 1980): 161–190.
- Simunic, D. A. 1984. “Auditing, Consulting and Auditing Independence,” Journal of Accounting Research Vol. ?? (Spring): 679-702.
- Taleb, Nissim Nicholas. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.
- Tergesen, A. 2005. Sizing Up A Hedge. *BusinessWeek Online*. November 28.
- Toth, R. B. 1996. *Standards Activities of Organizations in the United States*. U. S. Department of Commerce, National Institute of Standards and Technology. NIST Special Publication 806. Maryland, USA.
- Vangermeersch, Richard. 1986. *Financial Reporting Milestones in the Annual Reports of United States Steel Corporation*. New York: Gardland Publishing.
- Zhang, Yun (Clement). 2004. “Auditor’s Pre-negotiation Information, Accuracy of Financial Reports and Non-audit Services,” Ph.D. Dissertation, Yale University.

Figure 1: Example of an Aggregated Ratings Report on Internet (MSN)

2006 Honda Accord



[MSRP Price Range](#) \$18,225 - \$29,400
[Invoice Price Range](#) \$16,412 - \$26,455

MSN Ratings

[Expert](#) 8.0 [User](#) 9.3

ConsumerReports

[Overall Test Score](#)
89 [Read Snapshot](#)

[Price it with Options](#) [Save this Car](#)

[Overview](#) [Prices](#) [Features & Specs](#) [Photos](#) [Reviews](#) [Safety](#) [Reliability](#)

[At a Glance](#) [Expert Reviews](#) [Consumer Reports](#) [@User Reviews](#)

ConsumerReports.org

[Get Consumer Reports Pricing](#)

Ratings Snapshot

2006 Honda Accord EX V6

Other Trims Tested

- [EX 4-cyl](#)

Consumer Reports Overview

Highs: Acceleration, ride, handling, driving position, front-seat comfort, controls, crash-test results.

Lows: Road noise.

The Accord V6 is our top-rated family sedan. It has fairly agile handling, and the ride is steady and compliant. Inside the car the Accord is roomy, quiet and refined, although some road noise is noticeable. A telescoping steering column allows drivers to find an ideal position. The automatic shifts very smoothly and responsively. The four-cylinder engine is smoother than many V6s. Side and curtain air bags are standard. The V6 model is very quick and relatively fuel efficient. V6 models also get standard stability control for 2006. The V6 hybrid version is even quicker and gets 25 mpg overall, just one mpg better than the four-cylinder, which may not justify its \$30,000 price tag. Crash-test results are impressive.

Road Test Results

- Acceleration
- Accident avoidance
- Comfort and convenience
- Real-world fuel economy

Key

- Excellent
- Very Good
- Good
- Fair
- Poor

[For 20 more ratings and the Consumer Reports Bottom Line Price on this model, click here.](#)

Overall Rating
(Family sedans)

	P	F	G	VG	E
Highest Rated				89	
EX V6				89	
Lowest Rated			35		

Key P F G VG E
 Poor Fair Good Very Good Excellent

Table 1: Certification Services for Products Sold Online and Offline in the US

PANEL A: Four Levels of Certification Activity	Ebay (N=400)	PPI/CPI (N=417)	Total (N= 817)	% of Total (817)
Expert Opinions Based on Written Standards	344	399	743	91%
Expert Opinions Without Written Standards	45	14	59	7%
Lay People Ratings	5	3	8	1%
Meter	3	0	3	0.5%
No Certification / Rating	3	1	4	0.5%
Panel B: Breakdown of Certification Activities				
All Certification Available (Expert Opinions Based on Written Standards, Expert Opinions Without Written Standards, Lay People Ratings, and Meters)	221	101	322	40%
Expert Opinions Based on Written Standards only	4	110	114	14%
Expert Opinions Based on Written Standards, Lay People Ratings and Meter	33	57	90	11%
Expert Opinions Based on Written Standards and Expert Opinions Without Written Standards	47	34	81	10%
Expert Opinions Based on Written Standards, Expert Opinions Without Written Standards, and Lay People Ratings	17	34	51	6%
Expert Opinions Based on Written Standards, and Meter	8	28	36	4%
Expert Opinions Based on Written Standards, Expert Opinions Without Written Standards, and Meter	11	15	26	3%
Expert Opinions Based on Written Standards, and Lay People Ratings	3	20	23	3%
All other combinations	56	18	74	9%

400 goods sold on eBay and 417 goods included in the U.S. CPI and PPI (www.bls.gov) were selected. We then searched for the existence of written standards, expert opinions based on written standards (e.g., Michelin guide rating of restaurants), expert opinions that are not based on any written standards (e.g., New York Times Food Critic Rating of restaurants), lay people ratings (on various websites), and meters recording level of activity (e.g., revenue of a restaurant, or reservation time to get into a restaurant) which people can use to infer quality and/or popularity of a product. These results are summarized in panel A. The panel A results indicate that for 91% of the goods in our sample, we were able to find an expert opinion based on written standards. For the remaining 9%, we were able to find an expert opinion but without reference to any written standards (7%), lay people rating (1%), a meter (0.5%) or no certification of any kind (0.5%).

For many goods, multiple forms of certification are available; in some cases everything from expert opinions based on written standards, expert opinions which are not based on any written standards, lay people ratings, and even meters recording level of activity. In panel B, we summarize the different forms of certification available for goods in our sample. The Panel B combinations are presented in order of their frequency. The most frequent combination (occurring 40% of the time) is to have all forms of certification available (Expert Opinions Based on Written Standards, Expert Opinions Without Written Standards, Lay People Ratings, and Meters).

Table 2: Sports Card Grading Services (Summer 2004)

Name	Year Founded	Scale Categories	Grading Cost	Guarantee	Cross-seller	Feature
1. Accugrade (ASA)	1988	13	\$5-15	No	No	Invented 10-point scale + First online rating agency
2. Professional Sports Authentication (PSA)	1991	10	\$9-50	Yes	Yes	Membership Fee (\$99 – get 6 cards graded free) + Largest Market Share
3. KSA	1996	14	\$12-19	No	No	Canadian
4. American Authentication (AAI)	1996	10	\$12-22	No	No	5 card minimum per order
5. Finest Grading (FGS)	1997	14	\$5-25	No	Yes	Value Pricing Card Shows
6. Map Industries	1998	14	\$8	No	No	Free Shipping
7. Beckett Grading Service (BGS) –Vintage Grading (BVG) –Collectors Club (BCCG)	1999	19	\$8-25 \$9-26 \$5	Yes	Yes	Price guide + provide sub-grades for centering, corners, edges, and surface
8. Sportscard Guaranty (SGC)	1999	18	\$7- 50	No	Yes	Started with 100-point scale
9. The Final Authority (TFA)	1999	19	\$5-13	Yes	No	Value Pricing
10. Collection Monster (CM)	1999	19	\$10	No	Yes	Government Grading
11. Advanced Grading (AGS)	2000	11	\$9-25	No	No	Top view holders
12. Mint Grading Services	2000	18	\$6-20	Yes	No	Customer chooses grade. If MGS disagrees, pay \$1 only
13. CTA Grading	2000	11	\$10-30	No	No	Computer grading

Experts						
14. Bear Stats Grading (BSGS)	2002	14	\$10-20	No	No	Value Pricing
15. Global Authentication (GAI)	2002	19	\$6-20	No	Yes	Former PSA Experts + Dealer Focused
16. Pro Sports Grading (PRO)	---	22	\$9-50	No	No	2 experts grade each card
17. Professional Grading Service (PGS)	---	19	\$5-35	Yes	No	Help to post on eBay
18. World Class Grading (WCG)	---	20	\$5-20	Yes	No	Value Pricing
19. Champs Grading Service (CGS)	---	10	\$2-3	Yes	No	Racing Specialists
20. Grade Tech	---	21	\$14-50	No	No	Computerized grading
21. Premier Grading (PGI)	---	20	\$6-20	No	No	Value
22. Gem Trading - Gem Elite	---	9	\$8-15 \$12-19	No	No	Value
23. Sports Memorabilia Authenticator (SMA)	---	19	\$6-8	No	No	Value / Older Cards

In summer, 2004, there were 23 active sports card grading services online. We visited each website to identify the year of formation, and the rating scales. They use a 10-point scale with 1 or ½ point increments; some use +s or have discontinuities in their scale (e.g., do not award some grades such as 9.5).

We also recorded the grading fee they charge for the slowest service offered as well as the fastest service (customers also have to pay shipping and insurance costs to separate companies). Under grading cost (e.g., \$5-15), the first number is the cost of a normal grading service, and the second number is the cost of getting an expedited grading service.

Some companies offer a guarantee to return graded cards within a specified time period. Some companies have offline activities such as card shows where sports cards can be bought or sold and graded at a physical location, magazines, price guides, provide their own auction sites and a range of other services.

Table 3: Qualifications of Experts Providing Opinions on XYZ Website

	Accreditation	Education	Business	Hobby
1. Art and Antiquities (n=14)	64%	50%	71%	7%
2. Books, Maps, Manuscripts (n=5)	20%	20%	80%	20%
3. Clocks, Watches and Timepieces (n=4)	25%	25%	100%	0%
4. Clothing, Linens, Rugs and Quilts (n=6)	33%	17%	100%	0%
5. Coins, Stamps, Numismatics (n=7)	29%	14%	100%	0%
6. Electronics (n=3)	33%	0%	67%	33%
7. Famous People (n=4)	25%	0%	100%	0%
8. Furniture and Accessories (n=8)	50%	0%	75%	13%
9. Glass (n=5)	40%	20%	100%	0%
10. Guns, Knives and Swords (n=5)	60%	0%	100%	0%
11. Jewelry (n=6)	50%	33%	83%	0%
12. Knick Knacks and Collectibles (n=7)	14%	28%	86%	0%
13. Music (n=4)	25%	25%	75%	0%
14. Nature's Treasures (n=1)	0%	100%	100%	0%
15. Photography, Cameras, Projectors (n=3)	33%	67%	67%	33%
16. Porcelain, Ceramic and Pottery (n=10)	50%	20%	60%	20%
17. Silver (n=7)	43%	29%	71%	14%
18. Sports (n=6)	17%	17%	67%	33%
19. Tools, Kitchenware & Equipment (n=3)	33%	0%	100%	0%
20. Toys, Dolls, Games (n=9)	44%	11%	78%	0%
21. Transportation (n=4)	25%	25%	25%	25%
22. Wine (n=2)	50%	50%	50%	0%
23. General Appraisers (other) (n=6)	33%	67%	83%	17%
Total (n=128)	50 (39%)	32 (25%)	101(79%)	12 (9%)

A site we call XYZ is an official partner of eBay and has 128 experts who offer opinions on 1,850 separate items for a fee of \$9.95 (basic service) or \$29.95 (enhanced service). The site provides a description of some of its experts. Data in this table on accreditation, education, and business interests of experts were obtained from XYZ and reflects attributes of competence of all experts providing opinions on the site as of July 12, 2004. As far as we know, these are self-assessments of competencies with no independent verification that we are aware of. Most of the experts who provide an opinion on this site (79%) also run a related business (e.g., the wine expert owns a wine shop, and the carpet expert owns a carpet shop).

Table 4: Baseball Card Certification Services

(Baseball Cards Traded on eBay, N=321,045; 272,399 Singles and 48,646 Rookie Cards)

Issue Date	Graded	UnGraded	Total	%Graded	Graded	UnGraded	Total	%Graded
Pre1930s	1,935	2,933	4,868	40%	----	----	----	----
1930s	1,376	1,689	3,065	45%	----	----	----	----
1940s	417	1,448	1,865	22%	20	218	238	8%
1950s	6,160	22,869	29,029	21%	59	606	665	9%
1960s	7,291	28,007	35,298	21%	161	750	911	18%
1970s	6,858	14,171	21,029	33%	303	1,221	1,524	20%
1980s	3,692	8,533	12,225	30%	5,470	6,378	11,848	46%
1990s	2,693	25,163	27,856	10%	3,669	9,578	13,247	28%
2000s	1,356	135,808	137,164	1%	2,608	17,605	20,213	13%
Total	31,778	240,621	272,399	11%	12,290	36,356	48,646	25%

During August 19-Sept 3, 2004 a total of 321,045 baseball cards were traded on eBay. We counted the cards which had been graded by a 3rd party certification service. The data in the table show the number (and percent) of graded cards by decade in which they were issued. Data for single and rookie cards are shown separately to highlight the differences in the propensity to purchase grading services for different types of cards. eBay provides a breakdown by date of issue for singles cards from pre-1930s and 1930s, but not for rookie cards.

Table 5A: Frequency of Grades Given by 3rd Party Certification Services

(For a Sample of 1,000 Graded Rookie Baseball Cards)

3 rd Party Grader	Market Share %	Grade										Total	Average Score
		10	9.5	9.0	8.5	8.0	7.5	7.0	6.0	5.0	< 5		
PSA	36.8	105	0	159	0	80	0	17	4	2	1	368	8.9
BGS	22.7	1	85	73	37	13	14	1	1	1	1	227	8.9
BCCG	10.4	87	0	16	0	0	0	0	1	0	0	104	9.8
GEM	11.6	116	0	0	0	0	0	0	0	0	0	116	10.0
SGC	2.9	4	0	13	6	4	1	1	0	0	0	29	8.8
GAI	2.3	2	9	8	2	1	0	0	1	0	0	23	9.1
Other 17 Certifiers	13.3	91	2	20	10	2	1	0	1	3	3	133	9.4
Total	100%	406	96	289	55	100	16	19	8	6	5	1000	

The data in the table show the market share of twenty three baseball card certification services and the grades received by cards in our sample. The leading 3rd party certification provider (PSA) issued a grade of 10 to 4 percent of the cards (105/368). In contrast, the second certification provider (BGS) issued a grade of 10 only once out of 227 cards, but tended to offer 9.5 more frequently ($85/227 = 37$ percent). The average grade of cards certified by both PSA and BGS were both 8.9.

Table 5B: Frequency of Grades Given by 3rd Party Certification Services in Field Experiment

(Source: Jin et al. [2008], Table 2)

3 rd Party Grader	Total No. of Cards	Grade													Average Score
		10	9.5	9	8.5	8	7.5	7	6.5	6	5.5	5	4.5	4	
PSA	212	11		134		66		1		0		0		0	8.7
BGS	212	0	0	40	124	43	3	2	0	0	0	0	0	0	8.5
SGC	212	13		134	49	11	3	2		0	0	0		0	8.9
Kevin	212	0	1	40	129	37	4	1	0	0	0	0	0	0	8.5
Rick	212	0	11	57	92	45	3	2	0	1	0	0	0	1	8.5
Rodney	212	0	1	120	62	25	2	0	0	2	0	0	0	0	8.7
Total		24	13	525	456	227	15	8	0	3	0	0	0	1	8.6

Jin et al. [2008] conducted a round robin field experiment. They purchased 212 baseball cards and sent cards to three online Grading services (PSA, BGS and SGC) and three offline baseball card dealers (Kevin, Rick and Rodney) for grading. Grades assigned by these six graders are reported in the table. BGS and two offline dealers (Kevin and Rick) assigned an average grade of 8.5 and were more strict than the remaining three raters who assigned average grades of 8.7 (PSA), 8.9 (SGA) and 8.7 (Rodney).

Table 6: Average Returns to Grading of Rookie Cards (by Grader and Decade of Issue)

Rater	Number of Cards	Gross Return %	Net Return %	2000+ Gr. =10 %	1990's Gr.=10 %	1980's Gr.= 10 %	2000+ Gr.=9 %	1990's Gr.= 9 %	1980's Gr.=9 %	1970's Gr.= 9 %
BGS	227	264	46	333	-----	-----	44	86	176	-----
GAI	23	198	92	185	-----	-----	55	124	128	-----
SGC	29	125	4	303	-----	201	(27)	(65)	(1)	-----
PSA	368	238	(2)	25	47	178	(50)	(53)	(5)	683
BCCG	104	112	(14)	(15)	(11)	(3)	(68)	(72)	(25)	20
GEM	116	80.6	(23)	154	(38)	(24)	-----	-----	-----	-----

We selected a stratified-by-decade of issue sample of 1,000 from a population of 12,290 graded rookie baseball cards traded on eBay during August 19-September 3, 2004.

Gross return to getting a baseball card graded = ((selling price of a graded card – average price of the same card as per Beckett’s baseball card pricing guide for un-graded cards)/average price of the same card as per Beckett’s baseball card pricing guide for un-graded cards).

Net return to getting a baseball card graded = ((selling price of a graded card – average price of the same card as per Beckett’s baseball card pricing guide for un-graded cards – the lowest cost of grading option provided by the certification service)/average price of the same card as per Beckett’s baseball card pricing guide for un-graded cards).

The date indicates the decade when the card was issued. A card in the 2000+ category is a recently issued card. The grade 10 is the highest grade issued by the certification service (9.5 is the highest grade issued by BGS and GAI), the grade 9 is the second highest grade issued by the certification service.

Table 7: Regression Analysis

We conducted a regression where $R_i = \alpha_1 + \beta_1 \text{Year}_i + \beta_2 \text{Grade}_i + \beta_3 \text{GPSA}_i + \beta_4 \text{GBGS}_i + \beta_5 \text{GBeckett}_i + \beta_6 \text{GGAI}_i + \beta_7 \text{GGEM}_i + \beta_8 \text{GSGC}_i + e_i$

R_i = Return from getting a baseball card graded. This is calculated as selling price of graded card on EBay – (book value of same card in Beckett pricing guide + minimum cost of certification service) divided by book value of same card in Beckett pricing guide.

Year = Actual year when a baseball card was issued

Grade = score assigned by an independent certification service on a scale from 1 to 10, increasing in increments of 0.5.

GPSA = Grader is PSA; GBGS = Grader is BGS, GBeckett = Grader is BECKETT (other than BGS); GGAI = Grader is GAI, GGEM = Grader is GEM; GSGC = Grader is SGC. These are the top 5 graders. For each individual top grader, the test is whether return to the specific grader is greater than return from all other non top graders.

Regression (990 df): $R^2 = 0.1747$, $p < 0.001$. (Adjusted $R^2 = 0.168$)

	Coefficient	Std Error	T Statistic
Intercept	108.68	12.60	8.63 ***
Year	-0.058	0.006	-8.97 ***
Grade	0.72	0.06	12.65 ***
GPSA	0.37	0.16	2.22**
GBGS	1.12	0.18	6.25***
GECKETT	-0.18	0.19	-0.98
GGAI	1.21	0.31	3.87 ***
GGEM	-0.73	0.18	-3.97***
GSGC	0.48	0.28	1.68*

*** $p < 0.001$

** $p < 0.01$

* $p < 0.05$