

An empirical study of returns to scale of CPA firms in the post SOX era

Hsihui Chang · Hiu Lam Choy · Iny Hwang

Published online: 1 March 2015
© Springer Science+Business Media New York 2015

Abstract This paper performs empirical tests for the returns to scale pattern of certified public accountant (CPA) firms in the post Sarbanes–Oxley Act (SOX) era using data envelopment analysis. Analyzing a sample of 70 of the top 100 CPA firms from *Accounting Today* surveys over the 6 years period 2005–2010, we find that our sample CPA firms as a whole exhibit decreasing returns to scale in the post SOX era, suggesting that these CPA firms had exhausted scale economies by expansions through mergers and acquisitions and organic growth. Thus, these firms may consider stopping their expansion plans and downsizing or divestitures in order to improve their production efficiency.

Keywords CPA firms · SOX · DEA · Returns to scale · M&A · Organic growth

1 Introduction

Recent regulatory changes coupled with increasing globalization of the US economy requires certified public accountant (CPA) firms to provide clients more breadth and depth of services. Consequently, CPA firms has expanded their operations through a series of mergers and acquisitions (M&A) and organic growth. While continuing growth has been a pivotal part of CPA firms' long term viability, little is known empirically about whether CPA firms has exhausted economies of scale through expansions. Therefore, the purpose of this study is to empirically examine the returns to scale pattern of CPA firms in the post Sarbanes–Oxley Act (SOX hereafter) era. Towards this end, we use operations data on 70 non-Big 4 CPA firms to estimate their underlying production function. If these CPA firms' production function exhibits increasing returns to scale (IRS), then their recent expansions were justified and they could benefit further by continuing increasing their operation scale size. However,

H. Chang (✉) · H. L. Choy
LeBow College of Business, Drexel University, Philadelphia, PA 19104, USA
e-mail: hc336@drexel.edu

I. Hwang
College of Business Administration, Seoul National University, Seoul 151-916, South Korea

if these CPA firms' production function exhibits decreasing returns to scale (DRS), then they could have exhausted economies of scale and would need to consider decreasing their operation scale size in order to improve their productive efficiency. CPA firms exhibiting constant returns to scale (CRS) are at their optimal productive scale size (Banker 1984) and no change in their operation scale size is required.

Since its enactment in July 2002, SOX has significantly changed the playfield of CPA firms through the requirements imposed by its Sections 201 and 404. For example, Section 201 prohibits CPA firms from providing non-audit services, such as consulting services, to their audit clients and can reduce their revenues, whereas Section 404 requires CPA firms to perform additional audit work on clients' internal control system evaluations and can increase the audit fee charged for audit services (Chang et al. 2009a). Further, public companies hire other CPA firms instead of their auditors to provide consulting services as well as to assist in the establishment, documentation and assessment of internal control systems in order to comply with the requirement of Section 404. Chang et al. (2009a, 2010) demonstrate that the increased audit work load has forced leading CPA firms, such as the Big 4,¹ to drop their less profitable or high-risk audit clients. These clients have to turn to other, likely smaller, CPA firms for their audit services. In addition, some public companies switched to smaller CPA firms for better personalized services and reduced audit fees. The shift of clients from the Big 4 firms to smaller firms, together with the increased demand for additional audit services, leads to another wave of expansions through M&A and internal growth among the non-Big 4 CPA firms in the post SOX era (AICPA 2011).

The requirement for CPA firms providing audit services on public companies to register with the Public Company Accounting Oversight Board (PCAOB) and to perform its inspection has also affected the operation of CPA firms. SOX requires that CPA firms with 100 or less public clients be subjected to triennial inspection while those with more than 100 public clients be subjected to annual inspection. With the 100 public client level as the cutoff point for the different inspection requirements, small CPA firms with Less than 100 public clients can reduce their inspection costs per client by expansion (e.g. increase their public clients up to 100). Since a large portion of the inspection cost is fixed as long as the number of clients stays below 100, the more clients a CPA firm has, and the lower the inspection cost per client is. Likewise, for CPA firms with more than 100 public clients, the expansion also reduces the inspection cost per client. Consistent with this notion, Hood (2005) observes a significant increase in M&A activities within the second-tier CPA firms.² In particular, he finds a jump in M&A activities in 2004: about one-third of the second-tier CPA firms were involved in M&A activities during that year and the firms, on average, expanded their sizes by 13%.³ In addition, with increasing globalization, providing audit and accounting services to large multinational clients requires technical expertise and global resources that are not readily accessible to non-Big 4 CPA firms (Chang et al. 2009b). To build their geographic footprint and seize global opportunities, many second-tier CPA firms augment their operations by pursuing M&A of other firms in order to attract new talent, expand practice areas, and better leverage their expertise and resources (Accounting Today Institute 2012).⁴

¹ The Big 4 CPA firms consist of PricewaterhouseCoopers, Deloitte Touche Tohmatsu, Ernst and Young, and KPMG.

² We follow Hood (2005) and classify Big 4 CPA firms as the first-tier and top 100 non-Big 4 CPA firms as the second-tier.

³ The increase in operation scale size could be in part attributable to firms' internal growth.

⁴ Some firms use both M&A and organic growth to expand their business.

While the expansion enables CPA firms to pursue economies of scale, it can also lead to diseconomies of scale due to the difficulties of managing a larger workforce as these firms are getting bigger. More specifically, there are three major factors causing diseconomies of scale to occur when CPA firms over-expand their operations: (1) poor communication: As CPA firms grow their business communicating between different divisions and departments and along the chain of command becomes more difficult; (2) lack of coordination and control: As CPA firms get too big, coordination and control of activities get harder since it is harder for managers to closely monitor their subordinates and ensure they are all working together efficiently and effectively; (3) lack of motivation: it becomes more difficult for managers to build up a good team environment and sense of belonging as CPA firms expand their operations. This may result in lower employee loyalty and motivation with damaging consequences for production efficiency.

Since it is not clear in priori whether firms involved in expansions have exhausted scale economies, it is essentially an empirical research question. We address this question by examining how recent expansions among the second-tier CPA firms in the post SOX era affects their returns to scale patterns. Specifically, we adopt data envelopment analysis (DEA) to evaluate whether CPA firms' production function is characterized by CRS, DRS, or IRS. DEA is a useful analysis tool in this setting because it helps to determine whether a CPA firm has achieved their optimal productive scale size in using its inputs to generate outputs (Chandra et al. 1998; Emrouznejad et al. 2008). The optimal productive scale size is identified based on comparisons of the average productivities of CPA firms within the sample (Banker 1984). As such, it sets a realistic target of production rate improvement for the less efficient firms, based on the production frontier of the most efficient firms in the comparison group.

In a related study, Chang et al. (2009b) provides descriptive information of returns to scale patterns for 2003 and 2004, the first 2 years after the passage of SOX, on the Big 4 and non-Big 4 CPA firms using annual survey data from *Accounting Today*. They observe a CRS prevailing for the Big 4 CPA firms. However, approximately half of their second-tier CPA firms exhibit IRS while the remaining CPA firms reveal either CRS or DRS. Thus, they suggest that about half of the second-tier CPA firms can improve their production efficiency by increasing their exiting operation scale size. This current study extends Chang et al. (2009b) by examining the returns to scale pattern of non-Big 4 firms surveyed by *Accounting Today* for the period 2005–2010. The pattern of returns to scale can differ in our sample period from those documented in Chang et al. (2009b) for the following three main reasons.⁵ First, the deadline for compliance with Section 404 for accelerated filers (i.e. firms with a public float of at least \$75 million) is November 15, 2004. As the workload and cost of compliance with Section 404 is likely to be highest in the initial phase and would gradually reduce in subsequent years, the returns to scale pattern can vary with this change in workload. Therefore, it is interesting to examine how the returns to scale pattern of CPA firms evolve subsequent to the first 2 years after the passage of SOX. Second, expansions through M&A and organic growth helped the second-tier CPA firms significantly increase their operation scale size (AICPA 2011; Hood 2005). The impact of this change on operation scale size is more likely to show up in the years following the expansion rather than in the expansion year. Our 6 years sample period covering 2005–2010 allows us to capture the potential lagging effect of the expansion activities. Third, our sample period includes the period both before and after the financial crises which started in 2008. This enables us to examine whether the financial

⁵ Chang et al. (2009b) use 87 CPA firms while we use 70 CPA firms. The difference in sample composition may influence the estimation results. However, the CPA firms excluded from our sample are the Big 4 firms and those smaller ones that do not appear consistently over the 6 years sample period.

crises have a significant effect on CPA firms’ returns to scale patterns. While [Chang et al. \(2009b\)](#) observe mixed results based only on descriptive information, we provide empirical evidence from DEA-based statistical tests of returns to scale for the non-Big 4 CPA firms. Our results indicate that DRS prevails in these CPA firms in the post SOX era, suggesting that these firms had exhausted economies of scale with their recent expansions. As a result, they may consider stopping their expansion plans and contracting the operation scale size in order to get back to their optimal productive scale size.

The rest of the paper is organized as follows. In “Data envelopment analysis and its statistical tests for returns to scale” section, we discuss the DEA methodology and its statistical tests for returns to scale. In “Sample and descriptive statistics” section, we describe our sample selection process including the characteristics of our sample CPA firms, and discuss our empirical results. In “Conclusions” section, we present a summary of our findings and conclude the paper.

2 Data envelopment analysis and its statistical tests for returns to scale

Data envelopment analysis provides a nonparametric approach to estimate the production correspondence between inputs and outputs and evaluates the production inefficiency of decision making units (DMUs) from observed data. DEA has been extensively applied in various service industries including auditing, health care, and education industries. For instance, [Feroz et al. \(2005\)](#) apply DEA in their analytical procedure to determine audit scope and to assess the preliminary risk level of a client. [Ozcan et al. \(2010\)](#) use DEA to evaluate the performance of Brazilian university hospitals. [Banker et al. \(2010\)](#) employ DEA to compute the relative productivity of retail outlets of a high-end retailer. [Grigoroudis et al. \(2013\)](#) utilize DEA to evaluate the relative efficiency of a set of bank branches. [Duh et al. \(2014\)](#) measure the operating efficiency of universities using DEA. To elaborate on how this is involved, we define Y and X as the observed output and input vectors for DMUs (e.g. CPA firms). Then we can specify the production possibility set S as:

$$S = \{(Y, X) \mid Y \geq 0 \text{ can be produced from } X \geq 0\}. \tag{1}$$

The inefficiency is measured radially by the reciprocal of Shephard’s (1970) distance function. Therefore, the inefficiency of a DMU $(Y_0, X_0) \in S$ is given by:

$$\theta(Y_0, X_0) = \sup\{\theta \mid (\theta Y_0, X_0) \in S\} \tag{2}$$

[Banker \(1993\)](#) makes three main assumptions: (i) the production possibility set S is convex [i.e. If $(Y_1, X_1) \in S$ and $(Y_2, X_2) \in S$ then $(\lambda_1 Y_1 + \lambda_2 Y_2, \lambda_1 X_1 + \lambda_2 X_2) \in S$ for all $\lambda_1, \lambda_2 \geq 0$ such that $\lambda_1 + \lambda_2 = 1$], (ii) the production possibility set S exhibits monotonicity [i.e. If $(Y_1, X_1) \in S$, $Y_2 \leq Y_1$ and $X_2 \geq X_1$ then $(Y_2, X_2) \in S$], and (iii) the probability density function $f(\theta)$ is such that $f(\theta) = 0$ if $\theta < 1$ and $\int_1^{1+\delta} f(\theta)d\theta > 0$ for $\theta > 0$. By following [Banker \(1993\)](#), we can obtain a consistent estimator of the inefficiency $\hat{\theta}$ by solving the following DEA model of [Banker et al. \(1984\)](#) (BCC hereafter):

$$\hat{\theta}(Y_0, X_0) \equiv \text{Max } \theta \tag{3.0}$$

subject to

$$\sum_{j=1}^N \lambda_j Y_j \geq \theta Y_0 \tag{3.1}$$

$$\sum_{j=1}^N \lambda_j X_j - X_0 \leq 0 \tag{3.2}$$

$$\sum_{j=1}^N \lambda_j = 1 \tag{3.3}$$

$$\theta \text{ and } \lambda_j \geq 0 \tag{3.4}$$

Banker (1993) demonstrates that the DEA estimator $\hat{\theta}$ using the BCC model as in (3.0–3.4) is statistically consistent and the asymptotic empirical distribution of the DEA estimator retrieves the true distribution of $\hat{\theta}$ under the three assumptions described in the above. Since these assumptions do not impose the constraint of CRS on the production possibility set, they are consistent with variable returns to scale (VRS). We refer to such an inefficiency measure as $\hat{\theta}^B$. However, if we impose additional constraint that the production possibility set exhibits CRS, then by adding the assumption of CRS (i.e. If $(Y, X) \in S$ then $(kY, kX) \in S$ for any $k > 0$), we can obtain the Charnes et al. (1978) (CCR hereafter) inefficiency estimates $\hat{\theta}$ for a DMU by solving the same linear program as before; except that the objective function in (3.0) is maximized subject only to constraints (3.1), (3.2) and (3.4), while constraint (3.3) is eliminated.

By analogy, the CCR estimator is also statistically consistent under the four assumptions described in the above. We refer to the CCR inefficiency measure as $\hat{\theta}^C$. Given that the CCR model enables us to estimate the aggregate technical and scale inefficiency and that the BCC model helps us estimate the pure technical inefficiency of a DMU at the given scale of operation, we can calculate the scale inefficiency measure $\hat{\theta}^S$ as the aggregate inefficiency ($\hat{\theta}^C$) obtained from CCR model divided by the technical inefficiency obtained from BCC model ($\hat{\theta}^B$). That is, $\hat{\theta}^S = \hat{\theta}^C / \hat{\theta}^B$.

As described earlier, Banker (1993) proves that DEA estimators are statistically consistent. Therefore, under the null hypothesis of CRS, the asymptotic empirical distributions of DEA estimators, $\hat{\theta}^B$ and $\hat{\theta}^C$ are identical, each retrieving the true distribution of θ (Banker 1993). The asymptotic correspondence between the empirical distributions of $\hat{\theta}^B$ and $\hat{\theta}^C$ under the null hypothesis of CRS allows us to construct the following two DEA-based statistical tests for our study:

- (i) If the inefficiency θ is assumed to have an exponential distribution over $[1, \infty)$ with mean $1 + \sigma$, then $\sum_{j=1}^N (\hat{\theta}_j - 1) / \sigma$ follows the Chi-square distribution with $2N$ degrees of freedom. Thus, we can construct the following test statistic:

$$T_1 \equiv \sum_{j=1}^N (\hat{\theta}_j^C - 1) / \sum_{j=1}^N (\hat{\theta}_j^B - 1) \tag{4}$$

which can be evaluated using the F-distribution with $(2N, 2N)$ degrees of freedom.

- (ii) If the inefficiency θ is assumed to have a half-normal distribution over the range of values $[1, \infty)$ with mean $1 + \sigma$, then $\sum_{j=1}^N (\hat{\theta}_j - 1)^2 / \sigma^2$ follows the Chi-square distribution with N degrees of freedom. Thus, we can construct the following test statistic:

$$T_2 \equiv \frac{\sum_{j=1}^N (\hat{\theta}_j^C - 1)^2}{\sum_{j=1}^N (\hat{\theta}_j^B - 1)^2} \quad (5)$$

which can be evaluated using the F-distribution with (N, N) degrees of freedom.

Since [Banker \(1993\)](#) shows that the DEA estimator is biased for finite samples, the above DEA-based test statistics need not follow the F distribution. Notwithstanding this limitation, similar DEA-based statistical tests of inefficiency differences between two groups have been found to outperform conventional parametric and non-parametric tests for finite samples in extensive Monte Carlo simulation studies ([Banker and Chang 1995](#)).

Note that rejection of the null hypothesis of CRS test simply suggests that non-constant returns to scale prevails. In order to further investigate whether or not DRS holds for CPA firms, we can impose non-decreasing returns to scale (NDRS) constraint on the production possibility set to obtain the inefficiency estimator by solving the BCC model in (3) with “ \leq ” instead of “ $=$ ” in constraint (3.3). By referring to such inefficiency measure as $\hat{\theta}^D$, we can construct the test statistics for NDRS similar to those for CRS except that $\hat{\theta}^D$ is substituted into the denominator for $\hat{\theta}^B$ under the null hypothesis of NDRS. Since the CRS production frontier envelops the data less closely than NDRS, its inefficiency estimate $\hat{\theta}^C$ is greater than or equal to $\hat{\theta}^D$ calculated from the NDRS production frontier. Thus, for any DMU, equality suggests that the prevalence of NDRS and inequality implies DRS prevailing. Alternatively, we can impose non-increasing returns to scale (NIRS) constraint on the production possibility set to obtain the inefficiency estimator by solving the BCC model in (3) with “ \geq ” instead of “ $=$ ” in constraint (3.3). We refer to this inefficiency measure as $\hat{\theta}^E$ and construct the test statistics similar to those for CRS except that $\hat{\theta}^E$ is substituted into the nominator for $\hat{\theta}^C$ under the null hypothesis of NDRS. Since the NIRS production frontier envelops the observed data less closely than the VRS production frontier, the resulting inefficiency measure $\hat{\theta}^E$ is greater than or equal to $\hat{\theta}^B$ calculated from the VRS production possibility set. Thus, for any DMU, equality implies the prevalence of NDRS and inequality suggests the more restrictive property of DRS.

Conversely, we can evaluate whether or not IRS holds for CPA firms by constructing the test statistics for NIRS similar to those for CRS except that $\hat{\theta}^E$ is substituted into the denominator for $\hat{\theta}^B$ under the null hypothesis of NIRS. Since the CRS production frontier envelops the data less closely than NIRS, its inefficiency estimate $\hat{\theta}^C$ is greater than or equal to $\hat{\theta}^E$ calculated from NIRS production frontier. Thus, for any DMU, equality implies the prevalence of NIRS and inequality indicates IRS prevailing. Alternatively, we can construct the test statistics for NIRS similar to those for CRS except that $\hat{\theta}^D$ is substituted into nominator for $\hat{\theta}^C$ under the null hypothesis of NIRS. Since the NDRS production frontier envelops the observed data less closely than the VRS production frontier, the resulting inefficiency measure $\hat{\theta}^D$ is greater than or equal to $\hat{\theta}^B$ calculated from the VRS production possibility set. Thus, for any DMU, equality implies the prevalence of NIRS and inequality suggests the more restrictive property of IRS.

3 Sample and descriptive statistics

Each year *Accounting Today* surveys 100 accounting firms with the highest revenue in US. Information on these firms' revenues and number of employees is published in the issue of the following year. All data reported in the annual surveys are for domestic US operations and exclude foreign holdings. Our initial sample includes only firms in the top 100 list for each of the 6-year period 2005–2010 (*Accounting Today* issues 2006–2011). We then exclude non-CPA firms, such as Liberty Tax Services and Schenck Business Solutions, from our sample. We also exclude Big 4 CPA firms from our sample as all but one, Deloitte Touche Tohmatsu, do not report any revenues from MAS revenues from 2006 onward. Eventually, we use a total of 70 non-Big 4 top CPA firms in our analyses.

Following [Banker et al. \(2003\)](#) and [Chang et al. \(2009b\)](#), we focus on the relationship between service revenues generated and human resources employed by CPA firms. We partition the service revenues, by the types of output, into Accounting & Auditing Services (A&A), Tax Services (TAX) and Management Advisory Services (MAS). A&A includes compilations, special reports, and reviews in addition to engagements involving the attest function. TAX consists of tax research, planning and preparation work. MAS comprise such services as consulting, systems development, integrating and reselling computer equipment and software, and any other management assistance ([Banker et al. 2003](#)). The three human resource input variables considered are number of partners (PARTNERS), number of professionals (PROFESSIONALS), and number of non-professional employees (OTHERS). The designation PARTNERS includes all owners and shareholders of the accounting firms. PROFESSIONALS refer to staffs trained to perform the accounting and auditing services, and include such personnel as staff accountants, senior accountants, and managers. Those who passed CPA examinations but have not obtained CPA licenses are also included in this category. OTHERS encompasses the clerical and support personnel involved in such activities as administration, printing of reports, record keeping and the like ([Banker et al. 2003](#)).

We report in [Table 1](#) descriptive statistics on outputs and inputs of our sample CPA firms. There is generally an increase in both total revenues and number of employees over the 6 years period. For example, the mean total revenues increase by approximately 48 % from 98.8 million dollars in 2005 to 146.5 million dollars in 2010 while the mean total employees (sum of PARTNERS, PROFESSIONALS and OTHERS) increase by about 39 % from 642 in 2005 to 891 in 2010. The increase in PARTNERS is particularly acute because it increases by over 36 % from 65 in 2005 to 89 in 2010. The number of PROFESSIONALS increases by over 23 % from 431 in 2005 to 533 in 2010, and that of OTHERS increases by nearly 16 % from 145 in 2005 to 168 in 2010. The relatively larger increase in PARTNERS is consistent with the argument that CPA firms increase their size by expanding their hiring of partners mainly through mergers and acquisitions to accommodate the increasing demand for audit services in the post SOX era. The high standard deviations of the variables suggest that these non-Big 4 CPA firms vary significantly in their operation scale size and composition. Likewise, median values for all variables are much smaller than the means indicating that large disparities exist between the smallest and largest CPA firms in the sample even without the presence of Big 4 CPA firms.

Based on the data on three outputs (A&A revenues, TAX revenues and MAS revenues) and three inputs (PARTNERS, PROFESSIONALS and OTHERS) of CPA firms, we estimate their relative inefficiencies using DEA for each of the six post SOX years separately as well as for all 6 years pooled together as a whole. We test three hypotheses regarding returns to scale: (i) the null hypothesis of CRS against the alternative hypothesis of VRS, (ii) the null hypothesis

Table 1 Descriptive statistics on service revenues and human resource (N = 70)

Variables	Mean	STD	25 %	Median	75 %
<i>Year 2005</i>					
REVENUES	\$98.8M	\$182.1M	\$30.1M	\$37.1M	\$73.0M
A&A	43.2%	12.8 %	36.0 %	43.0 %	52.0 %
TAX	31.6%	9.5 %	25.0 %	32.0 %	35.0 %
MAS	25.2%	13.4 %	16.0 %	23.0 %	32.0 %
PARTNERS	65.6	108.2	20.0	29.0	59.0
PROFESSIONALS	431.1	751.3	131.0	178.5	338.0
OTHERS	145.4	306.2	33.0	44.0	110.0
<i>Year 2006</i>					
REVENUES	\$114.7M	\$205.6M	\$35.0M	\$44.4M	\$85.0M
A&A	43.5%	12.7 %	36.0 %	44.0 %	52.0 %
TAX	31.6%	9.2 %	26.0 %	30.0 %	36.0 %
MAS	24.9%	12.8 %	16.0 %	23.0 %	33.0 %
PARTNERS	71.7	115.4	24.0	31.5	59.0
PROFESSIONALS	492.5	849.8	159.0	196.0	389.0
OTHERS	166.8	378.3	38.0	50.5	113.0
<i>Year 2007</i>					
REVENUES	\$129.4M	\$221.9M	\$40.7M	\$49.8M	\$102.6M
A&A	44.4%	13.0 %	40.0 %	45.0 %	53.0 %
TAX	31.9%	8.4 %	26.0 %	32.0 %	36.0 %
MAS	23.7%	12.6 %	15.0 %	23.0 %	29.0 %
PARTNERS	76.9	114.1	24.0	37.0	68.0
PROFESSIONALS	533.9	892.0	167.0	226.0	436.0
OTHERS	171.8	365.1	44.0	61.5	106.0
<i>Year 2008</i>					
REVENUES	143.9M	\$240.9M	\$43.0M	\$58.0M	\$118.0M
A&A	44.2%	12.7 %	40.0 %	45.0 %	52.0 %
TAX	33.0%	9.2 %	27.0 %	32.5 %	37.0 %
MAS	22.8%	12.9 %	15.0 %	22.0 %	28.0 %
PARTNERS	82.1	120.4	25.0	38.0	80.0
PROFESSIONALS	555.7	855.3	185.0	249.5	453.0
OTHERS	182.9	395.0	44.0	68.0	126.0
<i>Year 2009</i>					
REVENUES	\$147.7M	\$237.1M	\$46.8M	\$59.7M	\$133.3M
A&A	44.3%	12.8 %	39.0 %	45.0 %	53.0 %
TAX	33.6%	8.9 %	28.0 %	33.0 %	38.0 %
MAS	22.1%	12.8 %	13.0 %	19.5 %	28.0 %
PARTNERS	88.8	125.1	27.0	40.0	94.0
PROFESSIONALS	548.8	819.5	176.0	263.0	517.0
OTHERS	175.8	338.7	41.0	65.5	138.0

Table 1 continued

Variables	Mean	STD	25 %	Median	75 %
<i>Year 2010</i>					
REVENUES	\$146.5M	\$224.3M	\$46.0M	\$64.6M	\$151.8M
A&A	43.9 %	12.7 %	37.0 %	45.0 %	51.0 %
TAX	34.1 %	9.1 %	30.0 %	34.0 %	38.0 %
MAS	22.0 %	12.7 %	14.0 %	18.0 %	28.0 %
PARTNERS	89.2	121.4	28.0	41.5	94.0
PROFESSIONALS	533.3	761.0	173.0	254.5	567.0
OTHERS	168.6	331.7	40.	58.5	141.0

N number of CPA firms in the sample, *REVENUES* total revenues expressed in million dollars, *A&A* proportion of A&A revenues, *TAX* proportion of TAX revenues, *MAS* proportion of MAS revenues, *PARTNERS* number of partners, *PROFESSIONALS* number of professionals, *OTHERS* number of other employees

of NDRS against the alternative hypothesis of DRS, and (iii) the null hypothesis of NIRS against the alternative hypothesis of IRS. We perform the tests for two different inefficiency distribution assumptions: the exponential distribution and the half-normal distribution.

Table 2 reports our statistical test results when the inefficiency θ is assumed to be exponentially distributed. For our sample period of 2005–2010 and for $\alpha = 10\%$, the first row of Table 2 indicates that the null hypothesis of CRS is rejected in favor of the alternate hypothesis of VRS in all 6 years exemplified. In addition, when we pool all the data point together over the 6 years, we find that the null hypothesis of CRS is rejected at 1 % significant level⁶ as indicated in the last column of Table 2. Taken together, our empirical results seem to indicate that the financial crises started in 2008 have little impact on CPA firms' returns to scale patterns.

Observing that the production function of CPA firms does not exhibit CRS, we next evaluate whether VRS stem from DRS or IRS during the post SOX era. The statistical test results are reported in the second and third rows of Table 2 under the null hypothesis of NDRS. As can be observed, the null hypothesis of NDRS is rejected at 10 % significant level for all 6 years except 2006 which has *p* values that are close to the 10 % level, suggesting that CPA firms' production function exhibits DRS.⁷ As expected from this finding, the null hypothesis of NIRS is not rejected at any conventional significant level as indicated in the fourth and fifth rows of Table 2. Similarly, estimating inefficiency over the entire sample period reaches the same conclusion that CPA firms exhibit DRS and the financial crises do not have any impact on returns to scale patterns.

We present the statistical test results in Table 3 when we maintain the assumption that the inefficiency θ is half-normally distributed. Consistent with, though statistically somewhat weaker than, those reported in Table 2, we find that the null hypothesis of CRS is rejected in favor of the alternate hypothesis of VRS in each of the 6 years and all 6 years together. Similar to those reported in Table 2, the null hypothesis of NDRS is rejected for all 6 years except 2006. Further, when we pool all the data points together over the 6-year period to

⁶ Unlike the individual year's test result, the result based on panel data pooled over 6 years needs to be interpreted with caution because a maintained assumption that the inefficiencies are independently distributed across years does not necessarily hold for the balanced panel data used in our study.

⁷ Other factors including the reduced cost of information technology may result in higher profits for these CPA firms that may explain in part why the mergers are continuing in the industry in spite of the observed DRS pattern. However, due to data limitation we are unable to further explore these issues in this study.

Table 2 Statistical test results of returns to scale when the inefficiency is assumed to be exponentially distributed (p values in parentheses)

Test statistics	Hypotheses		Years						
	Null	Alt.	2005	2006	2007	2008	2009	2010	All
$\frac{\sum_{j=1}^N (\hat{\theta}_j^C - 1)}{\sum_{j=1}^N (\hat{\theta}_j^B - 1)}$	CRS	VRS	1.377 (0.029)	1.355 (0.036)	1.438 (0.016)	1.357 (0.035)	1.372 (0.031)	1.278 (0.073)	1.377 (0.001)
$\frac{\sum_{j=1}^N (\hat{\theta}_j^C - 1)}{\sum_{j=1}^N (\hat{\theta}_j^D - 1)}$	NDRS	DRS	1.252 (0.092)	1.203 (0.137)	1.313 (0.053)	1.265 (0.082)	1.308 (0.056)	1.243 (0.098)	1.252 (0.001)
$\frac{\sum_{j=1}^N (\hat{\theta}_j^E - 1)}{\sum_{j=1}^N (\hat{\theta}_j^B - 1)}$	NDRS	DRS	1.277 (0.074)	1.229 (0.112)	1.343 (0.040)	1.285 (0.069)	1.323 (0.049)	1.244 (0.097)	1.277 (0.001)
$\frac{\sum_{j=1}^N (\hat{\theta}_j^C - 1)}{\sum_{j=1}^N (\hat{\theta}_j^E - 1)}$	NIRS	IRS	1.078 (0.328)	1.102 (0.281)	1.070 (0.342)	1.056 (0.372)	1.037 (0.415)	1.043 (0.399)	1.078 (0.137)
$\frac{\sum_{j=1}^N (\hat{\theta}_j^D - 1)}{\sum_{j=1}^N (\hat{\theta}_j^B - 1)}$	NIRS	IRS	1.099 (0.287)	1.126 (0.241)	1.095 (0.295)	1.072 (0.339)	1.049 (0.388)	1.053 (0.378)	1.089 (0.110)

$\hat{\theta}_j^B$ is obtained by solving the linear program in (3); $\hat{\theta}_j^C$ is obtained by solving the linear program in (3) except that the constraint $\sum_{j=1}^N \lambda_j = 1$ is eliminated; $\hat{\theta}_j^D$ is obtained from the linear program in (3) except that the constraint $\sum_{j=1}^N \lambda_j = 1$ is replaced by $\sum_{j=1}^N \lambda_j \leq 1$; $\hat{\theta}_j^E$ is obtained by solving the linear program in (3) except that the constraint $\sum_{j=1}^N \lambda_j = 1$ is replaced by $\sum_{j=1}^N \lambda_j \geq 1$

CRS constant returns to scale, VRS variable returns to scale, DRS (IRS) decreasing (increasing) returns to scale, NDRS (NIRS) non- decreasing (non-increasing) returns to scale

estimate the production frontier of CPA firms, the null hypothesis of NDRS is rejected at 1 % significant level. Obviously, the null hypothesis of NIRS can't be rejected at any conventional significant levels. Collectively, our findings infer that to large extent the non-Big 4 CPA firms over-expanded their human resources through M&A and organic growth in the post SOX era and may consider stopping expansion and reducing their operation scale size in order to improve their production efficiency.

4 Conclusions

Public accounting industry has experienced remarkable changes in the past two decades. The surging demand for management advisory services in 1980s and 1990s has led CPA firms to go through several waves of expansions through M&A and organic growth in pursuit of scale economies. With the new sets of regulations introduced by the Sarbanes–Oxley Act, CPA firms have been required to deal with intensified audit services but restricted non-audit services such as tax and management advisory services. These restrictions on the public accounting industry alter CPA firms' output (revenue sources) and input (allocation of professionals) structures. In addition, continuing globalization of the US economy requires CPA

Table 3 Statistical test results of returns to scale when the inefficiency is assumed to be half-normally distributed (p values in parentheses)

Test statistics	Hypotheses		Years						
	Null	Alt.	2005	2006	2007	2008	2009	2010	All
$\sum_{j=1}^N (\hat{\theta}_j^C - 1)^2 / \sum_{j=1}^N (\hat{\theta}_j^B - 1)^2$	CRS	VRS	1.454 (0.059)	1.403 (0.079)	1.568 (0.030)	1.390 (0.085)	1.539 (0.036)	1.382 (0.089)	1.454 (0.001)
$\sum_{j=1}^N (\hat{\theta}_j^C - 1)^2 / \sum_{j=1}^N (\hat{\theta}_j^D - 1)^2$	NDRS	DRS	1.339 (0.112)	1.261 (0.166)	1.448 (0.061)	1.314 (0.127)	1.490 (0.048)	1.377 (0.092)	1.338 (0.001)
$\sum_{j=1}^N (\hat{\theta}_j^E - 1)^2 / \sum_{j=1}^N (\hat{\theta}_j^B - 1)^2$	NDRS	DRS	1.368 (0.096)	1.290 (0.143)	1.485 (0.049)	1.332 (0.115)	1.506 (0.044)	1.380 (0.090)	1.367 (0.001)
$\sum_{j=1}^N (\hat{\theta}_j^C - 1)^2 / \sum_{j=1}^N (\hat{\theta}_j^E - 1)^2$	NIRS	IRS	1.063 (0.399)	1.087 (0.363)	1.055 (0.410)	1.042 (0.430)	1.022 (0.463)	1.042 (0.430)	1.063 (0.265)
$\sum_{j=1}^N (\hat{\theta}_j^D - 1)^2 / \sum_{j=1}^N (\hat{\theta}_j^B - 1)^2$	NIRS	IRS	1.086 (0.364)	1.032 (0.328)	1.082 (0.369)	1.057 (0.408)	1.033 (0.445)	1.056 (0.409)	1.086 (0.197)

$\hat{\theta}_j^B$ is obtained by solving the linear program in (3); $\hat{\theta}_j^C$ is obtained by solving the linear program in (3) except that the constraint $\sum_{j=1}^N \lambda_j = 1$ is eliminated; $\hat{\theta}_j^D$ is obtained from the linear program in (3) except that the constraint $\sum_{j=1}^N \lambda_j = 1$ is replaced by $\sum_{j=1}^N \lambda_j \leq 1$; $\hat{\theta}_j^E$ is obtained by solving the linear program in (3) except that the constraint $\sum_{j=1}^N \lambda_j = 1$ is replaced by $\sum_{j=1}^N \lambda_j \geq 1$

CRS constant returns to scale, VRS variable returns to scale, DRS (IRS) decreasing (increasing) returns to scale, NDRS (NIRS) non-decreasing (non-increasing) returns to scale

firms to build their competencies by differentiating services as well as acquiring experienced talents. To sustain growth and enhance competitive edge, many non-Big 4 CPA firms actively involved in M&A and internal growth activities after SOX. This raises an intriguing question of whether the expansion has enabled these firms to exploit scale economies or led them to overpass their optimal productive scale size.

Using data from *Accounting Today's* annual surveys of the top 100 accounting firms, we examine the returns to scale pattern of the non-Big 4 firms in the post SOX era. Specifically, we use a non-parametric DEA methodology to estimate production inefficiency and construct DEA-based statistical tests of returns to scale for these CPA firms. Our empirical test results indicate that the production function of the non-Big 4 CPA firms in our sample exhibits DRS pattern during 2005–2010. This result suggests that our sample non-Big 4 CPA firms had overgrown and exhausted economies of scale through M&A and organic growth. Therefore, in order to improve their production efficiency, these CPA firms may consider stopping their expansion plans and downsizing or divestitures. However, a caveat is in order. Our sample CPA firms may be focusing on long terms growth or improving service at the expense of short term production inefficiencies. That is, DRS may simply appear as a short term phenomena in the early years of expansions. Thus, our findings about the returns to scale pattern need to be interpreted with caution.

Given that the findings of this study indicate scale diseconomies, it would be interesting for future research to estimate the most productive scale size (MPSS) for each of these firms. Additionally, many CPA firms continue to engage in M&A activities after 2008 (AICPA 2011) in spite of the prevailing DRS pattern observed in this study. Thus, it would also be interesting for future research to look into this issue by investigating other factors, unrelated to returns to scale, affecting these firms' M&A decisions.

Acknowledgments The authors acknowledge the comments from two anonymous reviewers, William Cooper, Rajiv Banker, Ben Lev, participants of the 6th International Symposium on DEA at Temple University, INFORMS 2010 at Austin, Texas, the 4th International Conference on Management Science and Engineering Management at Chungli, Taiwan, and the Research Workshop at the National Chung Hsin University.

References

- Mergers and acquisitions. (2012). Accounting Today institute, New York.
- Accounting Today. (2006–2011). Top 100 tax and accounting firms, Faulkner & Gray Inc., New York.
- Accounting Trends & Techniques. (2011). AICPA, New York.
- Banker, R. (1984). Estimating most productive scale size using data envelopment analysis. *European Journal of Operations Research*, 17, 35–44.
- Banker, R., Charnes, A., & Cooper, W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30, 1078–1092.
- Banker, R. (1993). Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. *Management Science*, 39, 1265–1273.
- Banker, R., & Chang, H. (1995). A simulation study of hypothesis tests for differences in efficiencies. *International Journal of Production Economics*, 39, 37–54.
- Banker, R., Chang, H., & Cunningham, R. (2003). The public accounting industry production function. *Journal of Accounting and Economics*, 35, 255–281.
- Banker, R., Lee, S., Potter, G., & Srinivasan, D. (2010). The impact of supervisory monitoring on high-end retail sales productivity. *Annals of Operations Research*, 137, 25–37.
- Chandra, W., Cooper, W., & Rahman, L. (1998). Using DEA to evaluate 29 Canadian textile companies—Considering returns to scale. *International Journal of Production Economics*, 54, 129–141.
- Chang, H., Chen, J., Liao, W., Mishra, B. (2006). CEO's/CFO's swearing by the numbers: Does it impact share price of the firm?. *The Accounting Review*, 81, 1–25.
- Chang, H., Choy, H., Cooper, W., & Ruefli, T. (2009a). Using Malmquist Indexes to measure change in the productivity and efficiency of U.S. accounting firms before and after the Sarbanes–Oxley Act. *Omega—The International Journal of Management Science*, 37, 951–960.
- Chang, H., Galantine, C., & Thevaranjan, A. (2009b). Returns to scale pattern and efficient firm size in the public accounting industry: An empirical investigation. *Journal of the Operational Research Society*, 60, 1–7.
- Chang, H., Cheng, C., & Reichelt, A. (2010). Market reaction to auditor switching from big 4 to third-tier accounting firms. *Auditing: A Journal of Practice and Theory*, 29, 83–114.
- Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2, 429–444.
- Duh, R., Chen, K., Lin, R., & Kuo, L. (2014). Do internal controls improve operating efficiency of universities. *Annals of Operations Research*, 221, 173–195.
- Emrouznejad, A., Parker, B., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Sociology Economic Plan Science*, 42, 151–157.
- Feroz, E., Kim, S., & Raab, R. (2005). Analytical procedures: A data envelopment analysis approach. *Journal of Emerging Technologies in Accounting*, 2, 17–31.
- Grigoroudis, E., Tsitsiridi, E., & Zopounidis, C. (2013). Linking customer satisfaction, employee appraisal, and business performance: An evaluation methodology in the banking sector. *Annals of Operations Research*, 205, 5–27.
- Hood, D. (2005). The year of merging vigorously. *Accounting Today: Top 100 Firms*, 3.
- Ozcan, Y., Lins, M., & Lobo, M. (2010). Evaluating the performance of Brazilian university hospitals. *Annals of Operations Research*, 178, 247–261.
- Shephard, R. (1970). *Theory of Cost and Production Functions*. Princeton: Princeton University Press.