Attribution Testing
Exposing the influence of reverse causation in the attribution of better performance to ISO 9000 Quality Management Systems

Gavin P M Dick – Kent Business School, University of Kent
Inaki Heras – The Basque Country University, San Sebastian, Spain
Marti Casadesus – University of Girona, Girona, Spain

Working Paper No. 115
March 2006
Attribution testing
Exposing the influence of reverse causation in the attribution of better performance to ISO 9000 Quality Management Systems.

Gavin P.M. Dick
Kent Business School, University of Kent,
Iñaki Heras
The Basque Country University, San Sebastian, Spain
Martí Casadesús
University of Girona, Girona, Spain

Abstract
ISO 9000 Management Systems adoption (QCert) has proven to be a persistent and growing phenomenon, yet to date little research has been done that can safely attribute improved business performance benefits to it. The paper examines the evidence for the causal links between QCert and improved performance in the empirical literature. Tests for attribution of performance improvement are proposed that analyze, effect, cause and effect magnitude and these are illustrated to show how they influence the interpretation of results. The attribution testing method is then used to interpret the results of two USA and two European longitudinal studies and the role of reverse causation is shown to be a major mechanism that explains the superior performance of the certified firms. The analyses cast doubt on any inference causality being drawn from the broad literature that finds an association of ISO 9000 accreditation with better business performance since it indicates that the strongest direction of causality is that firms with superior performance are more likely to have certification, not that certified firms are more likely to have superior performance. The findings have profound implications for the interpretation of causation in the substantial literature that shows QCert is associated with improved business performance. For researchers the paper provides logic for testing the influence of reverse causation on results and it demonstrates the potential confusion of attribution that can lead to the development of flawed or incomplete theory.

Keywords:
Causality, ISO 9000, Performance, Quality.
Introduction

Although most “new” ideas in management have short life spans and are discarded when eclipsed by the next fad (Carson, Lanier, Carson and Guidry, 2000), ISO 9000 Management Systems adoption (QCert) has proven to be a persistent and growing phenomenon. Its persistence suggests that it is not simply another management fad but will remain an influential global management meta-standard (Uzumeri, 1997).

Despite the high cost of achieving and maintaining registration to the ISO 9000 Management System Standards, more than half a million organizations in 149 countries have made the investment (ISO, 2004). ISO 9000 Registrars make bold claims for the business benefits of QCert for instance in the USA ANAB (2005) claim 16 benefits from QCert including increased operational efficiency, cost savings from less rework, customer satisfaction, competitive edge, perceived higher quality and increased market share. In Europe similar claims are made by BSI the lead registrar in the UK (Breeze, 2004), but are these claims for attribution of improved performance to QCert valid? We suggest that the narrow scope of ISO 9000 compared to the ‘excellence’ models of quality gives us reason to doubt that QCert can be sufficient to be the cause. For instance Ferguson’s (1996) suggests that the requirement of ISO 9000 meets only 40% of the requirements of the Baldrige Award and an even more pessimistic view is take by Reimann and Hertz (1994) who suggest an overlap of only 10%.

This doubt motivated us to explore the evidence for benefits and in particular what evidence there is to show whether QCert can be attributed as the cause.

In this paper we propose a methodology for attributing causation of performance and contrast its results against those for normal cross-sectional methods. We demonstrate that the proposed attribution testing logic can lead us to very different results and conclusions than those obtained from cross-sectional methods. Our analysis leads us to conclude that reverse causation may well be a major mechanism that explains the superior performance often found in certified firms.

The paper starts with looking at the causal links between QCert and improved performance before looking at evidence for them in the empirical literature. We then discuss methods for testing causality and demonstrate their use on previously published data. We conclude with a discussion of the implications of our analysis for the attribution of performance in studies that examine QCert and the wider implications that reverse causation has for theory development.

---

1 Most reports are broadly consistent with the indicators of costs that come from a study by the Singapore Productivity and Standards Boards (1999), that suggests that companies need to plan for implementation costs of around $445 per employee, while maintenance costs will run at approximately $120 per employee. Indications from a recent studies (Casadesús and Karapetrovic, 2005) suggest that costs of implementation and maintenance are both falling so these estimates my be overstated.
Literature

A Quality and Business Causal Model

Although there is generally agreement in the literature on the association between quality and performance, we need to note that there is little commonality in how they measure business performance or define quality (Sousa and Voss, 2002).

The literature is in broad agreement on the potential causal chain between improved quality systems and better performance. Both Garvin’s (1984) and Deming’s (1986) Quality Model reason that as quality improves, waste is eliminated, costs are reduced, and financial performance improves. In the context of ISO 9000 Quality Management Systems the causal links can be extended as follows - a certified quality management system can achieve an increased emphasis on quality (Dick, Gallimore and Brown, 2000) leading to less waste and duplication of effort, and improvement in product quality. This means there are lower costs and fewer customer defections which lead to increased sales volume, while lowering the average cost of acquiring new business. These in turn lead to improved profitability from a combination of lower cost of production, lower sales expenses and scale economies from greater sales volume. Indeed, even if not all the quality benefits materialize, the possession of the ‘Quality Badge’ alone should lead to increased sales opportunities and so, improve profitability from increased sales volume. This causal model of improvements flowing from QCert to improved business performance is summarized in Figure 1. We next examine the evidence in the literature for the performance benefits and causal attribution to QCert.

Figure 1
The expected links between management system certification to ISO 9000 standards and business performance

ISO Certification \(\rightarrow\) Quality Emphasis \(\rightarrow\) Quality Improvement \(\rightarrow\) Business Benefits \(\rightarrow\) Business Performance

Certified by a third party Registrar as meeting an ISO 9000 Quality Management Standard

Increased emphasis on internal quality dimensions

Less waste and duplication of effort

Reduced costs improve competitiveness

Cost of sales reduces leading to increased profits

Increased emphasis on external quality dimensions

Service and product quality received by customers improves

Fewer customer defections

Profitability benefits from scale economies, and lower sales acquisition costs

Badge of quality opens sales opportunities

Sales volume increases

ISO 9000 Quality Management and Performance Literature

Here we review the empirical work in peer-reviewed journals from 1990-2005 that include reference to both ISO 9000, certification, and performance or benefits. The search used the BIDS, Emerald Management Reviews (formerly Anbar) and EBSCO databases to identify source materials. A four-stage approach to selection of articles was used. Initial screening of the 2000 or so search listing excluded
materials that were not in peer reviewed journals, followed by a relevance screening to exclude articles that did not explicitly measure business benefits or performance variables. At this stage it was found that the great majority of articles were focused on implementation issues, motives, and expectations rather than post-registration performance benefits which left approximately 100 articles. Next, the research methods of each article were assessed, and only survey research that had sample sizes with sufficient power\(^2\) (Cohen, 1988) and reported the statistical significance of their results were chosen. An exception was made for the three major surveys done by, or on behalf of, commercial firms prior to 1997; since peer reviewed journal articles prior to 1997 that met all the inclusion criteria were scarce. Finally, papers that included firms registered after 2000 were excluded so that the findings could be viewed as being uninfluenced by the major quality standards revision (ISO 9001:2000) that applied from 2001 onwards. Clearly, this process cannot claim to have captured every item of relevant research but it can be viewed as a substantial sample of the literature, which is unlikely to have any systematic bias in its selection.

This methodology resulted in a set of 30 research papers\(^3\) to be analyzed that can be broken down into three groups. The first of these are descriptive studies that provide no information on their statistical validity. These were reported in the early 1990’s the early days of ISO growth in Europe. The second group consists of ‘snapshot’ cross sectional studies that provide evidence of statistical validity. Academic papers of this type started to appear in 1997 and continue with many of these using intervening variables to explain how performance gains were achieved. The third group are longitudinal studies that provide information on business performance prior to and post certification that have the potential for indicating direction of causality. We start by reviewing the findings of the first two groups before summarizing their findings against the links in the ISO 9000 and Business Performance Model (Figure 1). We then review in more depth the third group that can provide evidence of causation.

Although there are many studies reporting expectations of increased market share and improved product quality from ISO 9000 implementation (for example, Ebrahimpour, Withers, and Hikmet 1997), there are much fewer empirical studies on the business performance benefits actually achieved.

One of the earliest studies by the Institute of Quality Assurance (IQA, 1993) found the most common gains related to better quality. Support for this is found in the Singapore research of Quazi and Padijabo (1998) and Chou-Chua, Goh and Wan (2003). Similarly, an analysis of 363 Norwegian firms by Sun (2000), found that QCert was associated with reducing customer complaints, product defects, and costs associated with rework and warranties. However, showing more limited

\(^{2}\) We discarded research that used a sample of less than 76. Cohen (1988 p31) calculates that samples larger than 76 are needed to ensure that a relationship > 0.4 that exists will be detected at a significance level of 0.05 and a power of 0.8 . Thus samples of less than 76 are unlikely to detect statistically significant effects that are medium in magnitude and very unlikely to detect small effect.

\(^{3}\) This may seem a very small number compared to the apparently vast research output relating to ISO 9000 but is comparable to findings of Ahire, Landeros and Goulhar (1995) who found only 29 empirical articles from the 226 on Total Quality Management that they reviewed.
benefits are Prabhu et al (2000), where fewer defects and greater productivity were found but customer satisfaction and reliability showed no gains.

However, most surveys find marketing gains are the most common advantage claimed by registered firms. Lloyd’s Register of Quality Assurance (1993) telephone survey of 400 quality managers claimed that internal benefits and market gains were achieved, and that these increased over time. Confirmation of these internal benefits is found in Buttle’s (1997) survey of 1220 certified UK companies which found that, as well as improving operations, marketing gains, were achieved by most of the firms following quality certification. Additional support for Buttle’s findings is found in the study by Casadesús, Heras, and Ochoa (2000) and Santos and Escanciano (2002) of firms in Spain. Further afield in Singapore Quazi and Padüjabo (1998) and Chou-Chua, Goh and Wan (2003) found marketing gains in addition to the earlier mentioned improvements in product quality.

Unfortunately, there are few studies that have used objective measures of profitability. The most quoted of these, despite its poor statistical methodology, is Lloyd’s Register of Quality Assurance survey (1996), that found certified companies’ sales growth, profit margins, and return on capital were much better than the industry average. However, research in the USA by Simmons and White (1999) and by Naveh and Marcus (2005) found no sales growth or margin improvements associated with QCert but they did find that the return on assets was better for firms who were registered compared to those who were not.

In contrast to the research we have reviewed so far indicating better business performance and improved profitability, Tsekouras, Dimara and Skuras (2002) and Dimara et al (2004) found no statistically significant benefits associated with QCert in their studies of Greek firms. Likewise, a study of 1000 firms in Australia and New Zealand found that quality certification had no significant, positive relationship with business performance (Terziovski, Samson, and Dow, 1997). They also noted that the principal motivation for pursuing quality certification was the ability of the certificate to open customers’ doors that were previously closed, or would close, if quality certification were not achieved. So, could motives for pursuing accreditation have a bearing on whether benefits are achieved?

Where studies do not report the full range of benefits suggested by the model presented in Figure 1, could this be due to organizations reacting to external pressure to be certified? Some studies (for instance, Gore 1994) have suggested when firms are reacting to external pressure for certification, they may see ISO 9000 registration as the prime objective, and adopt a minimalist approach to achieve it. These firms may possess quality certification but they do not value the quality management system that quality certification requires, so will achieve limited benefits. Support for this proposition is found in a study of 272 Australian firms by Jones, Arndt and Kustin (1997). It found evidence that firms that sought quality certification because of externally imposed perceptions of the necessity to ‘obtain a certificate’ were found to experience fewer beneficial outcomes of

---

4 The study compared self-reported sales and profitability of firms with QCert against Government industrial statistics that did not utilise directly comparable measures or methods.
certification than firms who had a ‘developmental’ view of quality improvement. These developmental firms’ motives included a desire to use quality certification to improve the company’s internal processes, and help lower quality costs, and/or increase customer focus. Using an identical typology of motives, a study in Malaysia of 405 firms (Yahya and Goh, 2001) also found that developmental motives made a difference, but only to the internal benefits achieved such as lower waste. Further support for benefits being contingent on firms having internal reasons for pursuing accreditation are found in three other studies (Brown, Van de Wiele and Loughton, 1998; Singels, Ruel and Van de Water, 2001; Yeung, Lee and Chan, 2003).

The importance of going beyond the minimum needed to obtain registration is shown by Huang, et al., (1999) that found that firms that had a strong motivation to thoroughly implement ISO 9000 beyond the mere purpose of obtaining the certificate obtained the greater benefits. In these firms, certification was linked strongly to increased international business, and to a lesser extent, better product quality and lower costs. Similarly Naveh and Marcus (2005) and Briscoe, Fawcett and Todd (2005) found that operating performance benefits were associated with high usage of QCert. Further insights into the motivation theme are provided by the research of Abraham, et al. (2000), who found that certification provided little guarantee of high performance outcomes unless accompanied by substantial changes in leadership, structure, and communications. However, the claim that QCert is primarily sought in response to customer or regulatory requirements is probably overstated as has been shown in the research of Anderson, Daly and Johnson (1999) who disconfirmed it as a primary reason for adoption. Instead they found evidence to support that QCert is a credible public signal of effective quality management practices.

However, in contrast to authors finding that motivation has a bearing on results, Terziovski, Samson, and Dow (1997) found that their variable ‘TQM environment,’ (indicative of a developmental view of quality) had no significant influence on the relationship between quality certification and business performance. While Leung et al (1999) observed that whether motives were customer driven or not made very little difference to whether benefits of accreditation outweighed the costs of achieving it. Their survey of 405 firms in Hong Kong found that the majority of firms reported that the benefits exceeded the cost.

Overall, the evaluation of the links between QCert and improved performance reveals that there is evidence in the field’s empirical research to suggest that the broad range of benefits shown in Table 1 are possible but uncertain unless motivation for pursuing QCert is for internal or developmental reasons. Particularly weak is the research evidence for claims that quality systems registration is associated with better quality. However, the evidence is stronger for reduced costs and increased sales or market share, both of which are consistent with the evidence for improved profitability. However, the reasonably strong evidence for increased profitability is not consistent with the idea that that gains
are expected to be attenuated as we move forward through the model. This suggests that not all the increased profitability can be safely attributed to QCert.

However, caution is needed in implying that certification is the cause of any benefits, since the methodologies that are used in twenty seven out of the thirty studies we have examined here, cannot test that certification is the cause as the methods used can only indicate association. Could the model’s proposition of forward causality between ISO 9000 certification and improved business performance be erroneous? Could it be the reverse causation also exists i.e. that better business performance precedes QCert and is being mistakenly attributed to QCert. In other words, could it be that organizations with above average business performance tend to pursue QCert more than less profitable firms and this explains or inflates the better performance found in the presence of QCert?

To examine this causation question we now examine in detail the three research articles that used research designs that could provide evidence of causality. Each of these longitudinal studies starts at the point when registration to ISO 9000 standards began to expand in the country or sector examined.

The first was Häversjö’s (2000) analysis of the returns on capital employed of 800 Danish companies between 1989 and 1995. Häversjö’s longitudinal results (Häversjö’s Table 1) show that the average financial performance of the certified organizations was superior to the non-certified organizations both before and after QCert but no post-registration performance gains of significance could be detected.

The second article that used a research design that could provide evidence of causality is Wayhan, Kirche and Khumawalas’ (2002) analysis of the performance of 96 organizations in the USA between 1990 and 1998. Their table of results (Wayhan, et al.’s Table 1) also show that the 48 registered organizations had a consistently better return on assets employed, both before and after their registration, compared to a control group of 48 non-registered organizations. As with Häversjö no significant post-registration performance gains were found.

The third examined the performance of 544 US firms between 1987 and 1997 (Corbett, Montes-Sancho and Kirsch, 2005) and like the other two longitudinal studies they also found that the firms with, or intending to pursue QCert had better than average performances than other firms in their industry. To overcome the problem of most firms being more profitable than average prior to accreditation they used a labor-intensive research design that matched each firm with one with a similar pre-registration performance. Their findings indicate achievement of small post QCert gains in ROA and sales that cumulatively become significant.

So it seems that of the three longitudinal studies that could indicate attribution only Corbett et al (2005) provides any evidence of benefits being gained from QCert. However, taken together these longitudinal studies do provide consistent evidence that adopters of QCert tend to be firms with above average performance. This is a cause for concern since it indicates that there is a systematic bias in

---

5 Attenuation is the concept that as we move through each step in the model (Figure 1) the effect is less than the cause that precedes it. Cumulatively this means that end performance gains such as increased profitability are likely to have less evidence of achievement than quality improvements such as less waste or better service quality that are nearer the QCert effect.
patterns of adoption that inflates the performance benefits found in snapshot cross-sectional research when QCert firms are compared to firms without QCert. The systematic bias could also lead to incorrect inferences being drawn about the effect of moderating variables since these could also be attributed to being the characteristics of firms with higher performance rather than influences which lead to QCert bringing benefits.

Clearly our review of the empirical literature suggests that there is a paucity of research designs that can show that benefits found can be safely attributed to QCert. With this in mind we discuss next a system for testing for attribution of performance and then go on to demonstrate its use.

Methodology

Attribution testing

Ideas on causation have exercised philosophers since Aristotle but perhaps the most appropriate modern regularity theory for use in the management field of enquiry is that of a cause being a sufficient condition for the occurrence of some effect with the rider that the cause must precede the effect and other possible explanations are eliminated (White, 1990). In practice in the social sciences causality is usually accepted in empirical research as requiring three conditions; variables that logically might influence one another must be associated, the causal variable must produce its influence before the outcome occurs and other possible explanations must be eliminated such as a third variable that influences both variables (Blaikie, 2003).

So in our context QCert has been shown in our literature review to have a chain of influences that might be a sufficient condition for the occurrence of better financial and sales performance. In other words we have plausible sequence of casual relationships (Figure 1) that we can view as mechanisms that can explain why QCert could cause improved financial performance and we have found associations between them that indicate that a cause and effect relationship exists. However, for causation to be attributed we also need to satisfy the other two conditions. We need to show that QCert preceded better performance and we need to find ways of separating performance differences to identify what part QCert influences.

So how can we these ideas be operationalised? Firstly, we need research designs that go beyond the dichotomous idea of comparing certified firms with those that are not-certified by splitting from the non-certified firms those firms that will be certified in the future (not-yet-certified). Thus in our design we have three cross-sections. Not-yet-certified, certified and a control group of non-certified firms. Figure 2 shows these three groups and their relationships.

To test for attribution of performance to QCert we would need to show that three conditions are satisfied. Firstly, we need to show that an effect exists in the presence of QCert that is absent when it is not present. Secondly, we would need to show that the cause can be attributed to QCert and that it preceded better performance, and finally we would need to demonstrate that the magnitude of the QCert influence on differences in performance found was significant. So we would need to test for effect by showing that Certified firms [x] had better performance than Non-certified firms [y]. We would need to show causality by testing that Not-yet-certified firms [z] had similar or worse performance than Non-certified firms and we would need to test for the magnitude of the influence by showing that
Certified firms had better performance compared to Not-yet-certified ones. Thus attribution to QCert requires: $x > y$, $z = y$, $x > z$.

However, if we find that Not-yet-certified firms and Certified firms have better performance than Non-certified firms and there is little or no difference in performance between Not-yet-certified firms and Certified firms then reverse causation is found. This indicates other causes are responsible for the better performance plus the presence of some mechanism(s) that results in better performing firms showing a greater propensity to adopt QCert than Non Certified firms. Thus attribution of better performance to other causes requires: $x > y$, $z > y$, $x = z$.

If better performance is found in both Not-yet-certified and Certified than Non-certified firms, and Certified firms have better performance than Not-yet-certified firms, then co-attribution of better performance is found. In other words, some of the performance differences found can be attributed to QCert with the rest being attributed to other causes and better performing firms being more likely to adopt QCert. Thus attributable QCert performance = $(x - y) - (z - y)$ and reverse causation = $(x - y) - (x - z)$.

Therefore, we propose to use three two tailed t-tests to determine performance attribution to QCert:

1) Test 1 (T1), Certified firms having significantly better performance than Non-certified firms [effect test = $x > y$].

2) Test 2 (T2), Not-yet-certified firms having no significant difference to Non-certified firms [cause test = $z = y$].

3) Test 3 (T2) test, Certified firms having significantly different performance to Not-yet-certified firms [magnitude test = $x > z$].
Figure 2 graphically summarizes these tests. Using the causality logic, we can see that T1 tests for a significant effect associated with our cause. While T2 is a causal test which must be non-significant for us to view the effect as being the predominant influence of the cause. Finally, T3 is a test of the magnitude of the influence of the cause on the effect.

At this point we need to make it clear that we are not suggesting that these tests are the ultimate solution in attribution testing, for that more complex methods (i.e. Pearl, 2005) or costly research designs are required (i.e. Corbett et al 2005). Rather the tests proposed represent a minimum needed to test for the potential influence of reverse causation in the interpretation of performance attribution.

Test Data

The research data that we are going to use to demonstrate the attribution testing methods comes from the Basque Autonomous Community, which is with Madrid and Cataluña considered to be one of the regions in Spain where ISO 9000 registrations are concentrated. The data is identical to that used by (Heras, Casadesús and Dick, 2002) but in that paper was only analyzed using a T1 type test that could only indicate association.

The data for that study was gathered from the Ardán database, an Entrepreneurial Information Service of the Consortium of the Exempt Zone of Vigo. The database is one of the most complete at domestic level in Spain for both economic and financial information, since it includes data for more than 100,000 companies, and more than 500 items of annual data for each company and year. The data is recorded from, among other sources, the outcome and balance sheets that companies submit to the Mercantile Register.

For the analysis, two samples were drawn from the database, which were a sample of 400 that were ISO 9000 certified companies the first of whom were registered in 1995, and another sample matched by industrial sector of 400 non-certified companies. Data was available for the years 1994, 1995, 1996, 1997 and 1998, and included the sales revenue for each accounting year, as well as the profitability ratio (ROA, the ratio of net profit before interest and tax on total assets). In addition, for the certified companies, the data set included information on their last quality certification registration date. This information on registration dates was checked with the registration bodies and where necessary with the companies to ensure that the date we used was the true date of the firm's initial registration to ISO 9000.

Possible sources of bias in the two samples were checked. Firstly, we noted that the two samples were not homogenous. Certified firms had on average larger sales turnover than non-certified firms did. To test that any difference in profitability of the certified companies is not a direct result of their larger average sales, we used the z-test of proportions, with a level of significance set at $\alpha=0.05$, as well as a t-test for differences in means. Both these calculations indicate that there is no significant effect.

Likewise, the distribution by industrial sector of both the certified and non-certified companies was analyzed and we found that the average profitability for certified firms was higher for each sector (manufacturing, construction, trade and services) compared to the non-certified firms. In order to evaluate if there were statistically significant differences in the profitability ratio among industrial sectors the average profitability ratio for all the sectors and years was calculated to verify
if sector differences were creating a bias in the results. Using t-tests for
differences in means, no statistically significant differences were identified (level
of significance set at $\alpha = 0.05$). Therefore, we feel confident that any differences
between ISO certified and non-certified companies that we may find are not
related to the sector distribution of the two samples.

In summary, the research design consists of three samples of firms: Certified, Not-
yet-certified and Non-certified for each of the five years, and two variables, sales
growth, and return on total assets employed (ROA).

Findings

We start by briefly presenting the original published findings (Heras, et al, 2002) of
our longitudinal study that used this data so that we can contrast the findings with
those using the attribution testing method we propose. In the original
methodology a dichotomous split was made with Not-yet certified firms being
excluded from the analysis. The results for sales growth are presented in Table 1
and the findings indicate that Certified firms achieve substantially greater
cumulative sales growth (56%) than Non-certified firms (40%) during the five years
with two out of the four years being statistically significant at or above the 0.05
level.

<table>
<thead>
<tr>
<th>Year</th>
<th>Non-certified</th>
<th>Certified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994-95</td>
<td>13.88%</td>
<td>*25.69%</td>
</tr>
<tr>
<td>1995-96</td>
<td>5.30%</td>
<td>*10.40%</td>
</tr>
<tr>
<td>1996-97</td>
<td>11.77%</td>
<td>10.84%</td>
</tr>
<tr>
<td>1997-98</td>
<td>8.70%</td>
<td>9.31%</td>
</tr>
<tr>
<td>Cumulative</td>
<td>*40%</td>
<td>*56%</td>
</tr>
</tbody>
</table>

*indicates that the t-test of means is significant at level of significance set at $\alpha=0.05$

A similar picture emerges for profitability (Table 2) with Certified firms enjoying
better profitability than Non-certified firms over the five year period with average
ROA being 8.20% compared to Non-certified firm’s 6.56%. Here three out of the
four years are statistically significant at or above 0.05 levels.

<table>
<thead>
<tr>
<th>Year</th>
<th>Non-certified</th>
<th>Certified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>5.50%</td>
<td>6.37%</td>
</tr>
<tr>
<td>1995</td>
<td>7.08%</td>
<td>*8.48%</td>
</tr>
<tr>
<td>1996</td>
<td>5.70%</td>
<td>*8.29%</td>
</tr>
<tr>
<td>1997</td>
<td>6.76%</td>
<td>*9.66%</td>
</tr>
<tr>
<td>1998</td>
<td>7.78%</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>*6.56%</td>
<td>*8.20%</td>
</tr>
</tbody>
</table>

* indicates that the t-test of means is significant at level of significance set at $\alpha=0.05$
certified firms have similar performance to Certified firms (T2: \( z \equiv y \)) which in turn implies that a T3 test \((x > z)\) would show similar gains to a T1 test \((x > y)\) so that the causal precedence is implied.

To see if these implications are sound we now examine the same data set but include in our findings the results for firms that will be certified during the five years. The findings for sales growth are shown in Table 3. Alongside the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-certified</td>
<td>13.88%</td>
<td>5.30%</td>
<td>11.77%</td>
<td>8.70%</td>
<td>40%</td>
</tr>
<tr>
<td>T1 Certified</td>
<td>*25.69%</td>
<td>*10.40%</td>
<td>ns 10.84%</td>
<td>ns 9.31%</td>
<td>*56%</td>
</tr>
<tr>
<td>T2 Not-yet-certified</td>
<td>*21.28%</td>
<td>*9.11%</td>
<td>*15.52%</td>
<td>*14.05%</td>
<td>*60%</td>
</tr>
<tr>
<td>T3 QCert difference</td>
<td>ns 4.4%</td>
<td>ns 1.29%</td>
<td>ns (4.72)%</td>
<td>ns (4.74)%</td>
<td>ns (4)%</td>
</tr>
</tbody>
</table>

* indicates that the t-test of means for Certified or Not-yet-certified firms’ data compared to Non-certified firms’ data is significant at level of significance set at \( \alpha = 0.05 \).
T3 QCert difference ns indicate no significant sales t-test difference between Certified and Not-yet-certified firms.

percentage sales growth is shown the results of the t-test significance for the attribution tests we described earlier for each of the years. Overall sales growth is significantly better for Certified and Not-yet certified than Not-certified firms (T1: \( x > y \) and T2: \( z > y \)) while gains from QCert are not statistically significant (T3: \( x \equiv y \)). This meets the conditions for reverse attribution, i.e. that firms had greater sales growth than their peers before QCert but achieve no additional significant gains from it. The results show clearly that firms with greater than average sales growth are more likely to adopt QCert but no significant additional sales gains are achieved post-certification.

The findings for profitability are shown in Table 4. Alongside the ROA percentage is shown the results of the t-test significance for the attribution tests described for

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-certified</td>
<td>5.50%</td>
<td>7.08%</td>
<td>5.70%</td>
<td>6.76%</td>
<td>7.78%</td>
<td>6.56%</td>
</tr>
<tr>
<td>T1 Certified</td>
<td>6.37%</td>
<td>*6.48%</td>
<td>*8.29%</td>
<td>*9.66%</td>
<td>*8.20%</td>
<td></td>
</tr>
<tr>
<td>T2 Not-yet-certified</td>
<td>*6.57%</td>
<td>*9.34%</td>
<td>*8.14%</td>
<td>*8.56%</td>
<td>*9.61%</td>
<td>*8.91%</td>
</tr>
<tr>
<td>T3 QCert difference</td>
<td>ns (2.97)%</td>
<td>ns 0.34%</td>
<td>ns (0.27)%</td>
<td>ns 0.05%</td>
<td>ns (0.71)%</td>
<td></td>
</tr>
</tbody>
</table>

* indicates that the t-test of means for Certified or Not-yet-certified firms’ data compared to Non-certified firm’s data is significant at level of significance set at \( \alpha = 0.05 \).
T3 QCert difference ns indicates that there are no significant ROA difference between Certified and Not-yet-certified firms.

each of the years. Overall ROA is significantly better for Certified and Not-yet certified than Not-certified firms (T1: \( x > y \) and T2: \( z > y \)) while ROA gains from QCert gains are not statistically significant (T3: \( x \equiv y \)). This, like the sales
attributed analysis, meets the conditions for reverse attribution, i.e. that firms had greater ROA growth than their peers before QCert but achieve no additional profitability gains from it. The results show clearly that firms with greater than average profitability are more likely to adopt QCert but no significant additional profitability gains are achieved post-certification.

If we compare these results with those in Table 1 and 2, we see a very different interpretation of the better results of QCert firms from those cited by the British Standards Institute (Breeze, 2004). We can see that the assumption of Not-yet-certified firms being similar to Non-certified firms ($z \equiv y$) that underlies their intuitive attribution of better sales and profitability is false. Our results indicate that for both sales and profitability, statistically significant ($z > y$) differences exist in all eight T2 tests and no significant effect magnitude resulting from QCert ($x > y$) can be detected in the eight T3 tests. Clearly, the better performance shown in the original research can now be attributed to a reverse causation mechanism.

Discussion

The literature that we reviewed earlier indicates that the most common benefits reported in the empirical research are increases in sales or market share. Our findings show that when we tested our data using cross-sectional analysis methods such as those found in the majority of the empirical literature on ISO 9000, we also found a significantly better sales growth in certified companies than in the control group of non-certified ones. However, using our attribution testing methods on the same data we found that none of these gains can be attributed to QCert. The tests indicate that there is no evidence to support any causal link between ISO 9000 registration and improvements in sales growth. Instead, we discovered that sales growth in Not-yet-certified firms was consistently better than Non-certified firms, and similar to Certified firms. Our tests thus show reverse causation, which indicates that firms with greater sales growth are more likely to pursue certification.

Our findings concerning profitability (ROA) follow a similar pattern. Our earlier cross-sectional study (Heras, et al, 2002) indicated that there was an association between profitability and certification. However, on testing for attribution, we found no evidence to support any causal link between ISO 9000 registration and improvements in profitability. Instead we discovered that profitability of the certified firms was consistently better than non-certified firms' pre and post their registration.

Our findings suggest that the many cross-sectional studies, such as Lloyds Register of Quality Assurance's survey (1996) that found certified companies' sales growth, profit margins, and return on capital employed were much better than the industry average, could well be implying an assumption of causality that is suspect. They are ignoring the influence of other causes and the propensity of more successful companies to pursue certification. Our findings on sales growth and profitability clearly illustrate how cross-sectional analysis can lead to erroneous inferences of causality, a question we will now discuss in greater detail.

As briefly mentioned earlier there is a problem in most studies of ISO 9000 certification and performance improvement since the direction of causation cannot be plausibly established using the research methods that are predominant in the empirical literature. Most of the research we reviewed were self-reported studies. Many relied solely on a T3 type test which shows effect magnitude, here good
financial performance may enable the pursuit of accreditation and may influence
the respondents’ perceptions concerning how much credit it deserves for the
current financial performance of the firm. Also common were methods that
contrasted the performance of firms by a snapshot cross-sectional split on a QCert
variable and thus are relying solely on a T1 type test. These studies can show
association with an effect or effect magnitude but the absence of a T2 test means
causation can never be proven. We have found only three longitudinal studies that
use methods that can cover all three tests. But even here in two out of the three
studies there seems to be too much emphasis given to the T1 and T3 test, with the
implications of the T2 findings indicating reverse causation not being fully explored
by the authors (Häversjö, 2000; Wayhan et al 2002). Only Corbett et al (2005)
study expressly deals with the T2 question through eliminating its influence in their
research design and here small but cumulatively significant effect magnitude gains
are indicated from QCert.

Our re-analysis of our earlier data shows only reverse causation which is consistent
with the results indicated by Häversjö (2000) and Wayhan, Kirche and Khumawala
(2002) data so at first sight these reverse causation conclusions appear to be at
odds with the finding of gains for QCert found by Corbett et al (2005). However,
Corbett et al results also report that the performance of Not-yet-certified firms are
greatly superior to Non-certified firms which lends support to the T2 test findings
from our study, Häversjö’s, (2000) and Wayhan et al (2002) data. We must
therefore conclude from our analysis that all the studies that can test attribution
find that reverse causation is a major factor in QCert performance attribution.
This shows clearly that where T1 or T3 tests indicate an effect related to QCert
probably only a modest proportion of the effect shown can be safely attributed to
the causal influence of QCert.

It is also worthy of note that our study and the three other studies that can show
causation have all used actual financial results, these should provide more reliable
evidence than self-reported results. All have shown the T2 cause test indicates
that firms with superior performance are more likely to have certification, not that
certified firms are more likely to have superior performance. This indicates an
underlying mechanism, a propensity amongst better performing firms to pursue
QCert earlier than their less successful peers. This has been found to be true in
three very different countries Denmark, Spain and the USA, which permits us to
generalize, as the reverse causation mechanism is clearly not just a local
phenomenon.

Therefore, what are the possible explanations that might underpin this reverse
causation mechanism? One possibility is that the quality system implementation
process takes place well before the certification date and so benefits accrue before
registration is achieved. An event study interpretation of the longitudinal data in
the four studies we have analyzed indicates that there is no evidence to support
this explanation since although year on year results vary there is no discernable
improvement in performance leading up to the registration year (Heras, Casadesús
and Dick, 2002). Another possible interpretation is that as the systems required by
ISO certification are costly to implement and maintain, profitable firms are more
likely to be ISO certified than less profitable firms. We noted earlier, that on
average the certified firms in our study were larger (this was also the case in the
Corbett et al study and is reported in other studies i.e. Simmons and White, 1999).
So could it be that the cost of accreditation is easier to bear for larger firms than
smaller ones, since they are likely to have more internal quality expertise and therefore less reliance on expensive consultants? Another possible interpretation is that all the studies examine the earlier years of adoption and subsequent growth in accreditation, so is it possible that these pioneer companies are characterized by having a greater exposure to international trade. Thus, these firms are more exposed to international standards of competition, and to compete they may already have in place many of the characteristics of “best practice” systems of quality management, prior to seeking accreditation. Therefore, pre and post certification business performance will not differ much, since gaining the “badge of quality” is only giving recognition for what were already good quality management systems.

Alternatively, could it be that there are latent common causes to QCert and better performance? An explanation could be that when firms already have in place good quality systems they are more likely to pursue certification early since their costs of implementation are lower and it is these extant quality systems that are the influence leading to their better than average performance? There appears to be adequate research to support this explanation as there is generally agreement in the literature on quality management system characteristics (the most dominant being improved conformance quality) that reduce internal costs, or are associated with business performance improvement (Maani, Putterill and Sluti 1994, Flynn, Schroeder and Sakakibara 1995; Flynn et al. 1997; Forker, Vickery and Droge 1996; Caruana and Pitt 1997; Adam et al. 1997; Samson and Terziovski, 1999; Hendriks and Singal, 2001; Kaynak, 2003; York and Miree, 2004). The explanation appears to be supported but once again this research can only indicate association, so it may be that this performance precedes the cause as York and Miree (2004) have suggested.

Extending this causal chain to its logical conclusion suggests that it may be the propensity of high performing firms to continually seek and learn from new practices/systems that can improve and sustain their capabilities that is ultimately the cause that explains their above average performance (on the role of learning see Naveh et al, 2004). These characteristics equate to those of Hayes and Wheelwright (1984) Stage Four companies where operations are creative and proactive in developing and adopting new practices and systems that relate to competitive performance (Flynn, Schroeder and Flynn, 1999).

So what are the implications of the reverse causation mechanism for the interpretation of the literature that can only infer causation? Clearly, the analyses cast doubt on any inference of attribution being drawn from the broad literature that finds an association of ISO 9000 accreditation with better business performance, since it indicates that the strongest direction of causality is that firms with superior performance are more likely to have certification, not that certified firms are more likely to have superior performance. Clearly, the theoretical model (Figure 1) needs modification to allow for the influence of reverse causation. Earlier, our review of the empirical literature revealed that

---

6 However, we need to note that there is little commonality in how they measure business performance or define quality (Sousa and Voss, 2002).
there is evidence in the field's empirical research to suggest benefits from QCert were uncertain unless firms had a developmental approach to QCert. Could this also be a false attribution of cause, since the possibility exists that these developmental characteristics are likely to be those of already high performing firms and may well contribute to that high performance? Thus, the developmental intermediate variable may be present in the already high performing firms and thus inflate its importance as an intermediate variable that explains achievement of QCert benefits.

Conclusions

In this research we have put forward three tests (Figure 2) that we suggest are necessary if attribution of performance is to be made. Before any attribution of performance can be made to QCert three conditions need to be satisfied. First, we need to establish that an effect exists in the presence of QCert that is absent when it is not present \([T1: x \cdot y]\). Second, we need to show that a cause that can be attributed to QCert exists \([T2: z = y]\) and finally, we need to demonstrate the magnitude of the effect caused by QCert is significant \([T3: x \cdot z]\). Thus attribution to QCert requires: \(x > y, z = y, x > z\). We are not suggesting that these tests are the ultimate solution in attribution testing, for that complex and costly research designs are required (i.e. Corbett et al 2005). Rather they represent the minimum needed to eliminate the potential influence of reverse causation (or a latent common cause) and provide logic for attribution of performance.

In this research we have used these three tests to analyze our earlier research data on sales growth and profitability of 800 firms divided into three samples: Certified, Not-yet-certified” and Non-certified over a period of five years. We have shown that the substantial difference between Certified and Non-Certified firms’ sales and profitability that we reported in our earlier research (Heras, et al, 2002) cannot be attributed to QCert. The findings of our attribution tests lead us to conclude that the superior performance of certified firms is due to firms with superior performance having a greater propensity to pursue ISO 9000 registration. Our evidence indicates that the direction of causality is that firms with superior performance are more likely to have certification. In other words a reverse causation mechanism is shown by our tests.

This finding of reverse causation has also been found to be the dominant causal mechanism from our analysis of the data of the other studies we could locate that could be tested for causation (Häversjö’s, 2000; Wayhan et al, 2002: Corbett et al, 2005). We do not preclude that benefits can be gained from QCert but suggest that its effect is not as strong as the reverse causation mechanism.

In the discussion, we put forward a range of possible reasons for this superior performance prior to accreditation to ISO 9000. Firstly, that the systems required by ISO certification are costly to implement and maintain, so more profitable firms are more likely to be able to afford ISO certification. Secondly, that the cost of accreditation is easier to bear for larger firms than smaller ones, since they are more likely to have internal quality expertise and therefore less reliance on expensive consultants. Thirdly, that the certified companies are characterized by having a greater exposure to international trade, and to compete they may have already emulated “best practice” systems of quality management prior to seeking accreditation. Thus, no great difference is found in pre and post certification performance. A fourth explanation is that there is a latent common cause. A
propensity of high performing firms to continually seek and learn from new practices/systems that can improve and sustain their capabilities; which ultimately causes their above average performance. Thus, better performance is not caused by any single system or practice but is the cumulative result of a process of continuous adoption, learning and adaptation of new management practices/systems.

For researchers the paper provides a logic for testing the influence of reverse causation on results and demonstrates the potential confusion of attribution in research designs that can only infer causation. The influence of reverse causation, we believe, profound implications for the interpretation of causation in the substantial literature that shows QCert is associated with improved business performance. Clearly, the evidence presented here for the presence of a strong reverse causation mechanism suggests that co-attribution or reverse attribution of performance benefits (and intermediate variables) must be considered when analyzing improvements in performance. Therefore, theories that model the influence of practices/systems on business performance need to include the possibility of reverse causation and use research designs that can control for the influence of reverse causation. Thus, the true level of causal inference can be established and in doing so enrich theory that explores the complexity of performance attribution.

For practitioners, our findings should give pause for thought. It is indeed tempting for managers to believe that ISO 9000 certification will lead to business benefits. After all, firms that they would like to emulate in terms of performance often have it! This is then reinforced by the seemingly pervasive believe (oft quoted as supported by research by certifying bodies e.g. Breeze (2004)) that a quality management system certified to ISO 9000 will increase sales and improve profitability through reduced costs. However, our findings indicate that it might be a wise decision to only pursue accreditation if there is a demand from customers for it, as we have found little evidence to suggest that sales or profitability improves after certification. Most reports indicate that certification is a major investment (Singapore Productivity and Standards Boards, 1999; Casadesús and Karapetrovic, 2005), yet our findings show that the money spent on certification has not adversely affected the profitability of our firms. This does suggest that cost benefits arising from certification are on average sufficient to offset the investment. Therefore, we are not suggesting to practitioners that certification is a bad investment, rather that inflated expectations of performance improvement are likely to be unfounded.

To summarize, we have explained a methodology that can be used to provide evidence of causation. Using these methods we have examined the results of our own and three other longitudinal studies and found evidence that superior performance cannot be attributed to certification since it was present prior to accreditation. In only one study from the four were any additional sales or profitability found that could be attributed to QCert. Thus, our findings cast doubt on any inference of attribution being drawn from the broad literature that finds an association of ISO 9000 accreditation with better business performance. Instead, our analysis indicates that the direction of causality is that firms with superior performance are more likely to have certification, not that certified firms are more likely to have superior performance. Overall, we have found little evidence that can prove that certification has any substantial influence on profitability or sales
growth. Instead, we have found that reverse causation is a major mechanism that explains the superior performance found in certified firms. We have provided explanations for this mechanism that indicate that existing causal theories of the origins of better business performance in management practices/systems may well be flawed in their causal assumptions.

References


Basque Autonomous Community, which is with Madrid and Cataluña considered to be one of the regions in Spain where ISO 9000 registrations are concentrated


