



ISO 9000/1994, ISO 9001/2000 and TQM: The performance debate revisited

Micaela Martínez-Costa^{a,*}, Thomas Y. Choi^{b,c,1}, Jose A. Martínez^{d,2},
Angel R. Martínez-Lorente^{d,3}

^a Universidad de Murcia, Facultad de Economía y Empresa, Campus de Espinardo, 30100 Murcia, Spain

^b Department of Supply Chain Management, W.P. Carey School of Business, Arizona State University, Tempe, AZ 85287-4706, United States

^c College of Business, Yonsei University, Seoul 120-749, Korea

^d Universidad Politécnica de Cartagena, Facultad de Ciencias de la Empresa, Paseo Alfonso XIII 50, 30203 Cartagena, Spain

ARTICLE INFO

Article history:

Received 13 December 2006

Received in revised form 2 April 2009

Accepted 10 April 2009

Available online 22 April 2009

Keywords:

Quality management

ISO 9000

TQM

Postal survey

ABSTRACT

The debate about the impact of ISO 9000/1994 on performance has been waging since its inception. While there is a general agreement regarding the positive impact of TQM on performance, there has been less agreement among the academics about the impact of ISO 9000/1994. Perhaps in response to such debate, the new ISO 9001/2000 has appeared purporting to be more in line with the TQM philosophy. As of now, how this 2000 version actually affects performance is yet to be explored. In this study, we compare the implementation of ISO 9000/1994 and ISO 9001/2000 as representing two different efforts to implement quality management practices. We evaluate its impact on company performance with a sample of 713 Spanish industrial companies. We also examine if the 2000 version of ISO is taking us closer to the implementation of TQM. Further, we depart from the past studies methodologically by considering performance as a formative construct rather than a reflective construct. Based on the mean and covariance structural (MACS) analyses, we conclude that ISO 9001/2000 certified companies do not perform noticeably better than ISO 9000/1994 or non-certified companies. However, we find that ISO 9001/2000 certified companies apply TQM at a higher level than ISO 9000/1994 certified companies, but whether they actually perform better is less clear.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

ISO 9000 series of standards first emerged in 1987 as the torch-bearer of the standards for doing business in Europe. According to the International Organization for Standardization (ISO), at the end of 2006 there were 897,866 certified companies in the world (www.iso.org).

With a growing annual rate of 20% for the period, 1995–2006, the number of ISO certified companies has been growing at a rate far higher than the economic growth. The academia has certainly taken notice of this phenomenon. As of late 2008, the ABI/INFORM database alone contains 2484 references that include ISO 9000 in their title or abstract. The bulk of these papers address managerial issues surrounding the ISO standard and its effects on company performance.

The *Journal of Operations Management* has published three seminal articles on the impact of ISO 9000 on company performance. The first paper appeared in 1997 (Terziovski et al., 1997) as one of the more rigorous studies analyzing this subject. In this paper, Terziovski et al. found that ISO 9000 had little or no impact on company

* Corresponding author. Tel.: +34 868887801; fax: +34 868 887537.

E-mail addresses: mili@um.es (M. Martínez-Costa),

Thomas.choi@asu.edu (T.Y. Choi), josean.martinez@upct.es

(J.A. Martínez), angel.martinez@upct.es (A.R. Martínez-Lorente).

¹ Tel.: +1 4809656135.

² Tel.: +34 968325776; fax: +34 968327008.

³ Tel.: +34 968325618; fax: +34 968327008.

performance. The authors used TQM as a moderating factor for the ISO–performance relationship. The second paper came in 2005 (Naveh and Marcus, 2005) addressing the same issue. They first reported ongoing debate regarding the effects of ISO standards on company performance. They then concluded that the way a company implements the standard is what introduces variations that distinguish one company from others in operating performance. The better a company uses the standard in daily practice and as a catalyst for change, the greater the operating performance from ISO implementation. The third paper is authored by Benner and Veloso (2008). They again highlight the existing debate between ISO 9000 and firm performance. They proposed a possible explanation for the previous contrasting results by differentiating the early adopters from late adopters. Financial performance advantages enjoyed by early adopters can disappear for the late adopters as more firms adopt and achieve similar generic improvements.

Clearly, the three articles have shed light on the question of the ISO's impact on company performance. However, all of them were built on the 1994 version of the standard. The study by Terziovski et al. (1997) appeared long before the 2000 version was instituted. Also, looking at the year of first submission (2002), the study by Naveh and Marcus (2005) also used data collected under the old ISO standards. The third and very recent paper (Benner and Veloso, 2008) uses data from 1988 to 1997. Presently, no more certification is being offered under the old version, ISO 9000/1994; the newest version of the ISO standard, ISO 9001/2000 (approved in December 2000) has now become compulsory for all applicants since the end of 2003.

A key departure reflected in the new ISO standard is a more extensive incorporation of TQM philosophy. Previously, ISO had been criticized for being incomplete in how it incorporated TQM practices into its requirements (Gotzamani and Tsiotras, 2001; Lee et al., 1999; Reimann and Hertz, 1996; Zhu and Scheuermann, 1999). The 2000 version is purported to be much more complete in its treatment of TQM; for instance, Biazzo and Bernardi (2003) observed that the changes that underlie the 2000 version have a much closer association with TQM principles compared to the 1994 version.

The main objective of this research is to extend the performance debate using the 2000 version of ISO. We do so by considering performance as a formative construct rather than a reflective construct (see Coltman et al. (2008))—all three papers we cite above (Terziovski et al., 1997; Naveh and Marcus, 2005; Benner and Veloso, 2008) used performance as a reflective construct. We examine if the incorporation of TQM concepts in the 2000 version is truly aligned with the TQM measures previously accepted in the academic literature. More explicitly, we pose three conceptual questions and one methodological question:

1. Has the 2000 version of the standard more impact on performance compared to the old one?
2. Are ISO 9001/2000 companies aligned more with TQM compared to ISO 9000/1994 companies?
3. Do the soft dimensions of TQM mark the primary improvement in ISO 9001/2000 from its earlier version of ISO 9000/1994?

4. Would framing performance as a formative construct contradict the results of earlier studies in any way that framed it as a reflective construct?

2. Literature review and hypotheses formulation

Since ISO 9000 series of standards first emerged in 1987, many authors found that ISO offered a reasonable first step toward implementing quality (Bradley, 1994; Claver et al., 2002; Escanciano et al., 2001; Gotzamani and Tsiotras, 2002; Skrabec, 1999; Stephens, 1997; Sun, 2000; Taylor, 1995; Tummala and Tang, 1996; Withers and Ebrahimpour, 2001; Yusof and Aspinwall, 2000). This initial version of ISO addressed quality issues in categories such as quality policy, quality documentation and quality planning. However, many authors also contended that the quality issues in ISO 9000 were addressed in a disjointed way scattered throughout the ISO document. The general impression at the time was that there did not seem to be an overarching TQM framework that guided the implementation of ISO requirements.

2.1. ISO and TQM

Many writers raised contentious issues about how well such a standard would compare with the overall TQM system (Goetsch and Stanley, 1998). It is true that there are common dimensions between ISO and TQM, and they rest largely on their emphases on process management and statistical tools. However, one consistent message that has come through in the writings of various authors (Gotzamani and Tsiotras, 2001; Lee et al., 1999; Reimann and Hertz, 1996; Zhu and Scheuermann, 1999) is that companies that were certified under ISO would still fall far short of implementing a comprehensive TQM system. Several authors (Goetsch and Stanley, 1998; Martinez-Lorente and Martinez-Costa, 2004) have presented a discussion of such comparisons between ISO and TQM.

One general consensus in the literature has been that TQM affects company performance significantly. For instance, Choi and Eboch (1998) showed how TQM practices have a significant impact on plant performance and customer satisfaction. Samson and Terziovski (1999) explained that behavioral factors in TQM were particularly strong predictors of performance. The question then remains whether the new standard, ISO 9001/2000, has in fact overcome the much debated shortcomings of the older version—that the 1994 version was lacking TQM practices and its subsequent mixed reviews regarding its impact on performance.

2.2. Implications of ISO for quality management and business performance

The old ISO 9000/1994 included 20 points that described categories ranging from management responsibility to statistical techniques. Contract reviews, design control, document and data control and purchasing were other categories. Under each of these 20 categories, there were more detailed descriptions of issues to be addressed. For instance, within the category of management respon-

sibility, the issues of quality policy, organization, and management review were addressed.

By comparison, the new ISO 9001/2000 version has four major sections that replace 20 points from the 1994 version. Each of these four sections is represented by four constructs and they are organized into one coherent framework. The constructs represented in the framework are (1) management responsibility, (2) resource management, (3) product and service realization, and (4) measurement, analysis and improvement. Further, they are organized in a way that one leads to another, as in a wheel, which ultimately improves customer satisfaction. Overall, the framework shows a dynamic model of a systems perspective, the bedrock of TQM philosophy. Indeed, this approach marks a radical departure from the previous version where the 20 categories were listed separately with no apparent link between them.

Authors such as Dobb (2004) and Schlickman (2003) echo these observations. They point out that the key change in the 2000 version is its adoption of the systems perspective inherent in TQM. Starting with the overall framework described in the preceding paragraph, ISO 9001/2000 has introduced the principle of continuous improvement that was conspicuously missing in the previous versions. The previous version was often referred to as a quality assurance system at best, but the 2000 version includes elements of TQM philosophy that separates it from a mere quality assurance system.

Finally, in supplanting the prevailing criticisms of the 1994 version as having an incomplete and disjointed quality management system, the new ISO 9001/2000 is poised to fill this gap and ultimately serve as a definitive step toward a higher level of quality management and performance. However, how much of what had been intended in the 2000 version would actually affect the quality management practices in the companies that adopt it still remains to be seen. For instance, Choi and Eboch (1998) pointed out that, when the motivation for implementation is “institutional” as is the case of ISO standards, the impact of the implementation might not be as immediate or direct as one would suspect. Nonetheless, comparatively speaking, companies that have adopted the 2000 version should show a higher average level of quality implementation compared to the companies that have not, considering the fundamental shifts that have occurred.

Most of the studies regarding TQM's influence on performance have concluded that companies applying it have reached better results (Adams et al., 1999; Choi and Eboch, 1998; Easton and Jarrell, 1998; Hendricks and Singhal, 1996, 2001a,b; Shetty, 1993; Terziovski and Samson, 1999, 2000). In contrast, there is still no general agreement regarding the effects of ISO 9000 on company performance. Although the predominant perception is that it has no positive influence (Aarts and Vos, 2001; Hua et al., 2000; Lima et al., 2000; Simmons, 1999; Singels et al., 2001; Terziovski et al., 1997; Wayhan et al., 2002), there are also some optimistic views (Docking and Downen, 1999; Gupta, 2000; Romano, 2000; Withers and Ebrahimpour, 2001). Nonetheless, none of these studies have taken into account the changes that the new standard has introduced.

We propose the following hypothesis. We intend to examine the impact of ISO 9000/1994 and ISO 9001/2000 on the company performance according to their different quality management levels.

H1. ISO 9001/2000 certified companies perform better than ISO 9000/1994 certified or non-ISO companies.

H2. ISO 9001/2000 certified companies show a higher level of improvements in TQM practices compared to ISO 9000/1994 certified or non-ISO companies.

2.3. Soft dimensions

In addition to the actual level of implementation of quality management practices, we were also interested in the impact of the 2000 version on the “soft” dimensions of TQM. In the TQM literature (Ahire et al., 1996; Black and Porter, 1996; Flynn et al., 1994; Saraph et al., 1989), there have been typically two types of TQM dimensions mentioned (Powell, 1995)—one of more “hard” or technical and the other of “soft” or human-related aspects. Examples of hard dimensions of TQM may include statistical process control and Ishikawa problem-solving tools. Examples of soft dimensions are discussed below.

As difficult as they might be to implement, the soft dimensions of TQM appear to be an integral part of the new ISO 9001:2000. Their presence is ubiquitous. Just to support this point, we present a list of soft dimensions of TQM contained in the 2000 version. Some of the completely new points occur in the areas of leadership and management commitment, continuous improvement, customer management, and human resource management (Dobb, 2004).

Leadership and management commitment: 5.5.3 (top management to ensure appropriate communication processes are established within the organization); 5.6 (the management review to assess opportunities for improvement); 6.1 (determine and provide the resources to continually improve effectiveness of the QMS).

Continuous improvement: 4.1 (requires an organization to continually improve the effectiveness of the quality management system); 5.4.1 (measurable quality objectives are established at relevant functions and levels within the organization).

Customers: 8.2.1 (monitor information relating to customer perception as to whether the organization has met customer requirements); 8.4a (analysis of data on customer satisfaction).

Human resource management: 5.5.3 (communication takes place regarding the effectiveness of the quality management system); 6.4 (organization to determine and manage the work environment needed to achieve conformity to product requirements).

This last point can be understood not only as a physical aspect but also as a desirable social atmosphere in which the workers can engage in their work (see ISO 9004/2000). Consequently, motivational issues could come into play, which are fundamental to TQM principles. The areas of

application can be as specific as ergonomic issues or can be as broad as the whole system of human resource management according to TQM principles. However, since this point does not specify how to achieve this “work environment,” the real application of the standard could be focused only on the physical aspects of work (Martínez-Costa and Martínez-Lorente, 2007). Additionally, there are requirements that have been expanded from the old ISO 9000/1994.

5.1 (4.1.1 in the previous version): Top management has to demonstrate involvement in the development and implementation of the QMS.

7.2.1b (before 4.3): Companies must anticipate customer needs and expectations; 7.2.3b: The companies must describe and have an appropriate method of communicating with customers.

5.2 (before 4.9): Companies must focus on enhancing customer satisfaction over and above simply conforming to requirements.

7.2.3c (before 4.14.2a): Customer feedback needs to be recorded as well as customer complaints.

6.2.2c and d (before 4.18): It is necessary to evaluate the effectiveness of training to be sure that workers performing the tasks affecting product quality are competent and that they are aware of the importance and relevance of their activities. This point is now integrated into a specific section of the norm previously; it was listed as an isolated point.

All the points listed above show numerous major changes that have taken place in the area of soft dimensions of TQM. Noting the 2000 version's significant departure from the 1994 version in the area of soft dimensions, we posit that the companies implementing the 2000 version of ISO would show significant improvement in this area. Therefore, we propose the following hypothesis.

H3. ISO 9001/2000 certified companies show improvements primarily in the soft dimensions of TQM compared to ISO 9000/1994 certified or non-ISO companies.

3. Methods

ISO 9001 was originally intended for manufacturing companies. Therefore, our sample includes only manufacturing companies. While the 2000 version adopts the terminologies for service companies, it also acknowledges that service companies show different characteristics than manufacturing companies (Beaumont et al., 1997; Brah et al., 2000; Gustafsson et al., 2003; McAdam and Canning, 2001). Furthermore, focusing on manufacturing companies would help us control for any confounding factors that might enter into our analyses by having both manufacturing and service companies.

3.1. Sampling and data collection

In 1986 Spain became a member of European Union. Spain's macroeconomic performance has remained remarkable: the country has experienced a 13th consecutive year of strong growth. This economic vitality has

had the effect of narrowing the gap in per capita GDP with the euro area average from 20% to under 12% over the past decade (OECD, 2007). Today the economy of Spain is the fifth largest in Europe, accounting for around 9% of EU output. Moreover, in terms of number of ISO 9001, Spain is the 4th country in the world by number of certificates, behind China, Italy and Japan (ISO, 2008).

Our sample was drawn from a Spanish financial database called *Sistema de Analisis de Balances Ibericos* (SABI). The data were collected from Spanish manufacturing organizations with more than 100 employees. Like other studies in quality management, the present study was designed to extend across industries. The total population showed 2986 companies. Many of them are engaged in international activity.

Data were gathered through a postal questionnaire sent to the companies in the population. The questionnaire was first pre-tested by experts from academia and industry. Questionnaires were then sent to the attention of the “quality manager.” Inside each envelope we included a presentation letter, a questionnaire and a pre-paid addressed envelope to be used when returning the completed questionnaire. In the cover letter, quality managers were promised a summary of the survey results.

The survey was sent to 2986 companies in March 2003. In May 2003, as advised by Frohlich (2002), the survey was sent again to 1500 randomly selected companies that had not responded the first time. We offered to quality managers who answered the questionnaire a report with the main findings of the study. From the original 2986 letters sent, 36 were returned due to unknown address and it was not possible to find them by other methods. Most likely, these companies either changed their address or were no longer in existence. There were 12 listed as manufacturing that turned out to be service companies. Finally, the population was made up of 2938 companies. The number of completed questionnaires was 713. It constituted a response rate of 24%. This percentage is considered a success with Spanish companies and is higher than the suggested minimum in the study done by Malhotra and Grover (1998).

3.2. Key variables

3.2.1. TQM

Flynn et al. (1994) defined and empirically tested a scale for measuring TQM. This scale was later used by the same authors for evaluation of TQM in the company performance (Flynn et al., 1995). This scale was chosen in order to measure TQM implementation. This is appropriate for manufacturing companies and also has been validated and accepted as a good measurement of TQM in the literature. Namely, Malhotra and Grover (1998) advise using previously tested scales in the literature and, as an example of a good scale to be used, they propose in their conclusions to use the Flynn et al. (1994) scale for measuring TQM practices in organizations.

3.2.2. ISO 9001

This is a categorical variable. It is a measure of whether the company has been certified by the ISO 9000 specifying which version is being applied.

3.2.3. Company performance

Company performance has been an elusive concept to measure. Some papers use financial performance such as sales (Corbett et al., 2005; Easton and Jarrell, 1998; Forker et al., 1996; Hendricks and Singhal, 2001a; Lima et al., 2000), market share (Forker et al., 1996), ROA or ROS (Corbett et al., 2005; Forker et al., 1996; Staw and Epstein, 2000), return on equity (Staw and Epstein, 2000), return on investment (Forker et al., 1996), income (Easton and Jarrell, 1998), income over assets and over sales (Lima et al., 2000). Many other papers use financial as well as non-financial performance. Powell (1995), for example, considers ROA, sales and others such as productivity. Product quality is another variable included in numerous studies and it is also measured in different ways (Curkovic et al., 2000; Choi and Eboch, 1998; Dow et al., 1999; Flynn et al., 1995; Forker et al., 1997; Forza and Filippini, 1998). Some measure performance with a combination of different operative results obtained by the company, together with financial and product quality measurements. In this sense, Terziovski et al. (1997) use operating performance (product quality, timeliness and productivity), customer satisfaction, employee morale, and business performance (cash flow, innovation, market share growth, sales growth, employee growth and export growth). Other papers using different measures of performance are Everett (1994), Flynn et al. (1995), Ebrahimpour et al. (1997), Tan et al. (1998), Martínez-Lorente et al. (2000) and Samson and Terziovski (1999).

The main objective of the papers above has been to analyse the effect of quality management in company performance. However, none of them use the same measures to capture performance, which makes it impossible to strictly compare their results. Given there is no clear agreement on the definition of performance in the literature, it appears that performance should not be limited just to the financial results (Quinn and Rohrbaugh, 1983; Venkatraman and Ramanujam, 1986) since it would only partially capture the business situation (Curkovic et al., 2000). Therefore, in our survey managers were asked about how their companies compared with their competitors. The dimensions used were: production costs, fast delivery, flexibility to change production volume and adapt stocks, cycle time, internal quality, external quality, customer satisfaction, market share, and employee satisfaction.

Such use of subjective data is widely used in the literature as exemplified by Powell (1995), Dow et al. (1999), Martínez-Lorente et al. (2000), and Douglas and Judge (2001). It is true that it may lead to a higher risk of perceptual bias, but it is also true that it allows controlling for the industry sector since the measurement is done relative to competitors. We asked for the company's position compared to its key competitors in last 3 years.

3.2.4. Performance as a formative construct

If the latent construct reflects the indicator measures, then it is referred as a reflective construct. In this case, the arrows of causality point from the latent construct to the indicators. However, if the latent construct is a manifestation of a set of indicator measures, then it is a formative

construct (Edwards and Bagozzi, 2000; Martínez and Martínez, 2008). In this case, this set of indicator measures is thought to “form” the latent construct. When considering performance measures such as high quality, fast delivery, and low cost, we discover that performance is actually a manifestation of these measures. These measures lead to or cause good performance rather than the other way around.

Coltman et al. (2008) delineate theoretical considerations for determining whether a reflective model or formative model is appropriate. In a formative model, a latent construct is a combination of its measurement indicators, the causal arrows move from the indicators to the latent construct, and the indicators constitute conceptually unrelated phenomenon. Contrarily, in a reflective model, a latent construct exists independent of the indicators, the causal arrows point from the latent construct to the indicators, and the indicators constitute a common theme or related phenomena. For instance, sense of well-being is a formative construct—job security, clean bill of health, and good personal relationships lead to the sense of well-being and they are conceptually independent from one another. Anger, on the other hand, is a reflective construct. Anger causes increase in heart beat, sweaty palms, and redness in face, and as a concept it exists independent of these indicators.

We measure performance using a formative approach. We define performance as a combination of several dimensions, as stated in the preceding section, and posit that performance does not exist independent of the measures used. Further, variation in the dimensions measured cause variation in performance, and dimensions do not need to share a common theme. For instance, they are not interchangeable; thus, adding or dropping an item necessarily changes the conceptual domain of the construct (see Coltman et al., 2008).

3.3. Empirical model and data analysis approach

To test for differences in performance variables between ISO 9000/1994-certified companies, ISO 9001/2000-certified companies and non-certified companies, mean and covariance structural analysis (MACS) was conducted (Sörbom, 1974). Numerous researchers have attested to the advantages of the MACS methodology compared to the traditional general linear model (GLM) (Lubke et al., 2003; Ployhart and Oswald, 2004; Raykov, 2001). The MACS methodology enables simultaneous modelling of latent variance–covariance and latent means within a single integrated statistical framework. This approach facilitates a comparison of different theoretical models to determine the model that best fits the data. Furthermore, MACS estimates latent mean differences, which accounts for measurement error variance and thus obviates the need for other methods (such as *t*-test or ANOVA).

The research model in the present study is shown in Fig. 1. It consists of nine dependent variables as dimensions of company performance and one independent variable with three levels (ISO 9000/1994 certified companies, ISO 9001/2000 certified companies, and non-certified compa-

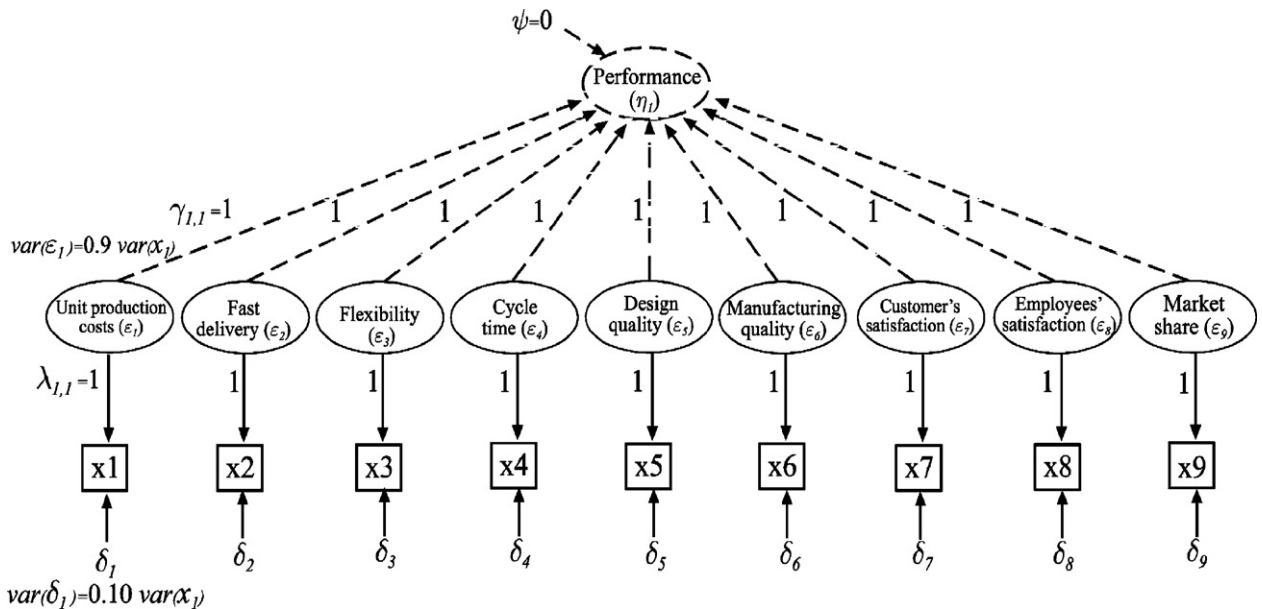


Fig. 1. Covariance structural model for three groups of companies. Note: Specification is only indicated for the unit production costs, using the common LISREL notation. The remaining latent variables and indicators are specified in a similar way.

gies). This represents a special case of GLM multivariable analysis of variance (MANOVA). Given the justified emphasis on the importance of reporting effect sizes (Cohen, 1990; Thompson, 2002), the MACS approach provides more reliable results and more appropriate interpretations because, by showing the magnitude of mean differences, it accounts for measurement error variance in latent variance estimations.

Our proposed model is shown as a formative model with a reflective latent respecification. These kinds of models are called spurious models (Edwards and Bagozzi, 2000) and they require respecifications of the formative model that permits a realist interpretation. In other words, we argue that the construct of interest (i.e. performance) is a perfect composite of several latent variables (in this case, nine dimensions) that are measured in a reflective way. This would permit accounting for the measurement error in the observable indicators of each dimension (Diamantopoulos, 2006).

The philosophical and practical aspects of reflective and formative models are still a matter of debate. As evidence, diverse perspectives on this subject matter appear in Howell et al. (2007), Bagozzi (2007), and Bollen (2007). While there are other references in the literature that explain the differences between them (Diamantopoulos and Winklhofer, 2001; Jarvis et al., 2003), a more profound theoretical discussion can be found in Edwards and Bagozzi (2000) or Borsboom et al. (2003). A middle position is proposed by Coltman et al. (2008) who delineate practical issues involved in theoretical and empirical considerations for determining whether a reflective model or formative model is appropriate. For instance, the theoretical consideration focuses on the nature of the latent construct and the direction of causality, and the empirical consideration covers how

items are interconnected and the issue of measurement errors.

Theoretically, the rationale for considering a formative model with a spurious respecification, as we have proposed, instead of a reflective model is straightforward. Reflective models must comply with the principle of local independence (Coltman et al., 2008); that is, if a latent construct underlies a number of observed variables, then conditionally that latent construct will render the observed variables statistically independent. This condition means that equally reliable indicators of a latent construct are interchangeable. If we are mainly interested in estimating the latent mean, each single indicator should estimate the mean, and each indicator should not differ significantly in mean value from the rest. In other words, high correlations between indicators are not enough to represent the latent mean correctly, because if indicators are interchangeable, the selection of an indicator with a low score would distort the true latent mean. However, as discussed above, formative indicators are not interchangeable and do not have to be necessarily correlated. In other words, all of them are necessary to form and determine the latent mean. From an operational view, a concept that is measured with formative indicators is merely an algebraic construction. There is no causal statement beyond the organizing principle of the representation (Markus, 2004), and there is no distinction between the construct and its measures because the construct is defined in terms of its measures.

In the same vein, the variable *performance* does not match the requirements of the reflective perspective, and the *nine proposed dimensions* do not necessarily have to be correlated. Therefore, companies can enhance their levels of performance through, for example, enhancing customers' satisfaction, without needing to enhance market

share. Further, companies with high scores on the nine dimensions will perform better than companies with high scores on only some of them. This is a logical statement, but it is *not* compatible with the reflective view because we can expect that if there are performance differences between companies, this will be reflected in all the dimensions, not only in some of them. Finally, the proposed dimensions of company performance are not interchangeable because if we disregard one of them, the meaning of the construct is clearly altered. Therefore, if the model is formative, then some widely used methods for testing the proposed scales from a reflective perspective (i.e. confirmatory factor analysis or two-step structural equation modelling) can be seriously misleading.

3.4. Possible sample bias

The companies surveyed were obtained from a database in which there is information about size (number of employees), industry sector, and financial results. We tested to see if there were important differences between sample and population.

The first difference analyzed was the sector distribution. Table 1 shows the number of companies in each industry (percentage in parentheses). To analyze the difference between the sample and the population we correlated the number of companies existing in each sector with the number of companies of each sector in the sample. The Pearson correlation was 0.957 and significant at the 1% level.

An analysis of variance (ANOVA) comparing the average of employees in the population and the sample was done in

order to test if there was any difference with respect to company size. The ANOVA did not reject the null hypothesis of equal means ($F = 0.034$, $p = 0.854$). The average ROA was also compared, and the results of the ANOVA did not show that means of population and sample were different ($F = 0.724$, $p = 0.395$).

We tested for the possible self-selection bias among the respondents. It might be that companies that responded had higher levels of quality management than non-responding companies. The total quality management averages of both the first 10 days respondents and the last 10 days was compared. No significant differences ($F = 2.821$; $p = 0.094$) were detected. Further, we compared the quality means of respondents in the first and second waves of the survey. We again did not find any difference ($F = 3.596$; $p = 0.058$). We also tested for the possible position bias in the company of the manager who answered the questionnaire. We compared the means of total quality management dimensions between all the possible respondents and did not find any difference ($F = 0.830$; $p = 0.563$).

4. Results

4.1. Performance

We first explored the intercorrelations between performance indicators. Bivariate correlations ranged from 0.218 to 0.653: there is no definite pattern of correlations between variables, suggesting lack of support for the reflective approach. To confirm, we conducted a confirmatory factor analysis by programming a reflective model (i.e. the performance latent variable underlying the nine performance indicators). The fit of the model was very poor: $SB\chi^2(df)$: 202.37 (27); $p = 0.000$. Again, this result shows lack of support for the reflective model because the principle of local independency does not hold. In addition, to examine whether indicators would be interchangeable as they would be under the reflective model, we estimated several confirmatory factor analyses, using 10 of the 126 possible different random samples of four of the nine performance indicators. These analyses would be similar to searching for TETRAD differences (Bagozzi and Fornell, 1989). Results indicated that 70% of tested models were misspecified, again exhibiting empirical evidences against the reflective model.

Therefore, we were confident that the results indicated performance was indeed a formative construct. We accept that the variable “company performance” is an index composed of nine dimensions and these nine dimensions cover the theoretical domain of the construct. We can then specify that company performance represents an exact weighted composite of k dimensions ($\gamma_{1,k} = 1$) as a single algebraic construction. Theoretically, there is no error variance for this composite ($\psi = 0$). The next stage is to specify the measurements of these nine dimensions.

In order to avoid halo effect and other method biases (Podsakoff et al., 2003), we used one observable indicator per dimension. This approach is in agreement with several authors who defend this form of measuring latent variables (Bergkvist and Rossiter, 2007; Hayduk and Glaser, 2000; Rossiter, 2002). In addition, we followed

Table 1
Industry distribution.

Sector	Population	Sample
Food and beverages	417 (14.8%)	97 (14 %)
Tobacco	7 (0.2%)	0 (0%)
Textiles	137 (4.8%)	16 (2.3%)
Confectionery	68 (2.4%)	6 (1%)
Leather	36 (1.3%)	67(0.9%)
Wood	59 (2.1%)	24 (3.5%)
Paper	96 (3.4%)	25 (3.6%)
Printing	164 (5.8%)	20 (2.9%)
Petrol	6 (0.2%)	3 (0.4%)
Chemical	270 (9.6%)	64 (9.3%)
Plastics	165 (5.8%)	49 (7.1%)
Non metallic minerals	245 (8.7)	65 (9.4%)
Iron	110 (3.9%)	31 (4.5%)
Metals	235 (8.3%)	60 (8.7%)
Machinery and mechanical equipment	185 (6.6%)	57 (8.2%)
Office and computer systems	9 (0.3%)	3 (0.4%)
Machinery and electric equipment	122 (4.3%)	38 (5.5%)
Electronic equipment. Radio, TV and communication equipment	47 (1.7%)	13 (1.9%)
Medical, surgery and optician equipment	28 (1%)	9 (1.3%)
Automobile industry	177 (6.3%)	45 (6.5%)
Other transportation industry	62 (2.2%)	17 (2.5%)
Furniture	89 (3.2%)	25 (3.6%)
Recycling	6 (0.2%)	2 (0.3)
Production and distribution of electric energy, gas and hot water	23 (0.8%)	8 (1.2%)
Water distribution	46 (1.6%)	4 (0.6%)
Building	14 (0.5%)	3 (0.4%)

the recommendations of Hayduk (1996) with regard to fixing the error variance of each observable indicator. In other words, assuming each indicator is the best indicator of each dimension, we were able to fix the meaning of each theoretical dimension. By fixing the error variance of each indicator, we accepted a certain level of reliability of each observable measure.

4.2. ISO and performance

Having explained the rationale of our model, we were now ready to undertake the MACS analysis. As a prerequisite for testing group mean differences, MACS requires an assessment of measurement invariance (construct equivalence) between groups. The relationship between latent variables and observable indicators in two groups must be similar to ensure that latent mean differences are not biased by the existence of disparate patterns of causal relationships between constructs and indicators (Lubke et al., 2003).

Because measurement invariance can occur at different levels, three types of invariance must be analyzed: (1) configural invariance (i.e. similar measurement models and the same pattern of factor loadings exist between groups); (2) metric invariance (i.e. indicators have the same causal relationships with their respective constructs across groups – that is, the factor loadings are identical); and (3) error invariance (i.e. indicators are equally reliable across groups – that is, the measurement error is invariant for the two groups considered). The last of these is the most restrictive condition because factor loadings and indicator reliabilities must be statistically equal across groups. According to Steenkamp and Baumgartner (1998), the minimum required level of invariance depends on the aim of the study; however, given that full measurement invariance is an infrequent occurrence, at least partial measurement invariance should exist (Byrne et al., 1989). The weaker condition of partial measurement invariance is traditionally accepted for MACS analysis.

Once measurement invariance is established, the next step is to undertake a latent mean difference test and estimate the associated effect size. Both the latent mean difference test and the effect size will account for indicator reliability and covariance relationships between model variables. At the beginning of the analysis, non-ISO 9000-certified companies were selected as the reference group. ISO 9000/1994-certified companies and ISO 9001/2000-certified companies were respectively taken as the second and third group. Cases with missing data were dropped, given the characteristics of formative models (Rossiter, 2008). The final sample sizes for the three groups were 104, 249 and 224, respectively. Subsequently, the restricted model (M0_{G1}) was tested. The error variances of the k indicators, $Var(\delta_k)$, were fixed to 10% of the observable indicator variance (Hayduk, 1996). This means that the reliability coefficient of the measures is equal to 90%. In addition, variances of the k latent dimensions, $Var(\varepsilon_k)$, were also fixed to a specific value, following the partition of variance Eq. (1):

$$Var(x_k) = \lambda_{k,k}^2 Var(\varepsilon_k) + Var(\delta_k) \quad (1)$$

Given $Var(x_k)$ is the observed variance of the k indicators and assuming independence on the right hand side of the equation, we can fix both $Var(\varepsilon_k)$ and $Var(\delta_k)$ as a function of certain level of reliability, in this case 90%. Note that $\lambda_{k,k}$ was also fixed to 1, because we also assigned the same metric to the latent variables and observed indicators.

The LISREL 8.80 robust maximum likelihood method was applied (Jöreskog and Sörbom, 2006), because non-severe departures of multivariate normality were encountered. Results showed a perfect fit of the model, $SB\chi^2(df)$: 0.000 (9); $p = 1.000$. Models M0_{G2}; M0_{G3} were also tested following the same analytic approach. As expected, both models fitted: $SB\chi^2(df)$: 0.059 (9); $p = 1.000$, $SB\chi^2(df)$: 0.027 (9); $p = 1.000$, respectively. The specification of our model directly implies that configural, factor, and error invariance hold, given that factor loadings were fixed to be the same and all indicators were fixed to be equally reliable across groups. This was an inherent assumption derived from the use of one single indicator per latent variable.

Multivariate latent mean and covariance structure analysis was then performed. The series of nested models proposed by Ployhart and Oswald (2004) was adapted for the present study. Latent mean contrasts (M2 and M3) thus were achieved. Model M2 established that latent mean differences did not exist: it was built by adding the indicator means to the model of each respective group and constraining the latent means to achieve equality. This model fulfilled the requirements of error invariance. The fit of this model was adequate ($SB\chi^2$: 54.936 (45); $p = 0.147$). However, Model M3 yielded a perfect fit $SB\chi^2$: 0.004 (27); $p = 1.000$. This model allowed latent means to differ, which led to the conclusion that differences existed in latent means across groups, because the nested model comparison favoured M3: $p(\Delta SB\chi^2) = 0.000$, such that the scaled chi-square difference (Satorra and Bentler, 1999) yielded a significant result.

Model M3 was programmed with non-ISO 9000-certified companies as the reference group for the comparison. Latent means for the other two groups were permitted to vary. Additional two competing models were also evaluated: M3₂ and M3₃. Model M3₂ considered that latent means of non-ISO 9000-certified companies and ISO 9000/1994-certified companies were equal, with the latent mean of the ISO 9001/2000-certified companies being different. The fit of this model was also very good— $SB\chi^2$: 24.195 (36); $p = 0.933$. Finally, Model M3₃ considered that latent means of the two groups of certified companies were equal, with the mean of the first group being different. Again, a very good fit was encountered at $SB\chi^2$: 31.714 (36); $p = 0.673$.

These results clearly indicate that (1) performance of certified companies is different from non-certified companies in at least one of the nine dimensions and (2) ISO 9000/1994 and ISO 9001/2000-certified companies also differ in at least one of the nine dimensions. This latter conclusion is derived from the nested model comparison achieved between M3₂, M3₃ and M3. As Table 2 shows, both comparisons yielded significant results. We then considered M3 as the best model. In addition, a final model was estimated (M4) in order to test if latent variances were homogeneous. This analysis yielded a poor fit: $SB\chi^2$:

Table 2

Fit Indices for the model-testing sequence (ISO-performance).

Model	Test	$SB\chi^2(df)$	p -Value	$p(\Delta SB\chi^2)$
M0 _{G1}	Baseline model	0.000 (9)	1.000	
M0 _{G2}	Indicators equally reliable	0.059 (9)	1.000	
M0 _{G3}	Indicators equally reliable	0.027 (9)	1.000	
M2	Latent means fixed to be equal	54.936 (45)	0.147	
M3	Latent means differ	0.004 (27)	1.000	M2-M3 = 0.000*
M3 ₂	Latent means of groups 2 and 3 fixed to be equal and mean of group 1 differs	24.195 (36)	0.933	M3 ₂ - M3 = 0.000*
M3 ₃	Latent means of groups 1 and 2 fixed to be equal, and mean of group 3 differs	31.714 (36)	0.673	M3 ₃ - M3 = 0.000*
M4	Latent means of the groups 2 and 3 fixed to be equal, and homogeneity of variances	41.063 (27)	0.040*	
M5	Latent means of the groups of certified companies differ	0.000 (18)	1.000	

* $p < 0.05$.

41.063 (27); $p = 0.040$, rendering support for the heterogeneity of variances.

The differences in each dependent variable are shown in Table 3. The results indicate that the possession of ISO certification had a significant effect on at least one of the nine performance variables.

Having demonstrated a statistical difference among the three groups, the effect size was then considered to ascertain the substantive differences. Utilizing the variances of the performance dimensions and the latent means differences, Cohen's (1977) d effect size and the approximate 95% confidence interval for d were calculated using Hedges and Olkin's (1985) procedure.

The results showed that ISO 9000/1994 certification had a significant effect only on one performance dimension: flexibility. This effect could not be considered large, because confidence interval of the effect size included small and medium values, in accordance with Cohen's (1988) conventions. In addition, ISO 9001/2000 certification yielded a significant effect on two performance dimensions: fast delivery and design quality. The direction of the difference for the former was contrary to what we expected: certification had a negative effect on fast delivery. The strength of the effects for both dimensions (fast delivery and design quality) was not large.

Table 3

Latent mean differences along performance dimensions. Non-ISO certified companies are the reference group.

	Latent mean differences (standard error)	Cohen's d effect size ^a (95% CI) 90% reliability	Cohen's d effect size ^a (95% CI) 80% reliability
Unit production costs			
ISO 1994	-0.130 (0.141)	-0.113 (-0.343; 0.116)	-0.120 (-0.349; 0.109)
ISO 2000	0.110 (0.137)	0.103 (-0.13; 0.335)	0.109 (-0.124; 0.342)
Fast delivery			
ISO 1994	-0.200 (0.134)	-0.182 (-0.411; 0.047)	-0.193 (-0.422; 0.036)
ISO 2000	-0.290 (0.133)*	-0.272 (-0.505; -0.038)	-0.288 (-0.522; 0.055)
Flexibility			
ISO 1994	0.340 (0.142)*	0.298 (0.069; 0.528)	0.317 (0.087; 0.547)
ISO 2000	0.240 (0.141)	0.219 (-0.014; 0.452)	0.233 (-0.001; 0.466)
Cycle time			
ISO 1994	0.140 (0.135)	0.128 (-0.101; 0.357)	0.136 (-0.093; 0.365)
ISO 2000	0.130 (0.133)	0.125 (-0.108; 0.358)	0.133 (-0.100; 0.365)
Design quality			
ISO 1994	0.030 (0.143)	0.026 (-0.203; 0.255)	0.027 (-0.201; 0.256)
ISO 2000	0.320 (0.140)*	0.292 (0.058; 0.526)	0.310 (0.076; 0.543)
Manufacturing quality			
ISO 1994	-0.070 (0.124)	-0.069 (-0.298; 0.160)	-0.073 (-0.302; 0.156)
ISO 2000	0.060 (0.123)	0.061 (-0.171; 0.294)	0.065 (-0.168; 0.298)
Customers' satisfaction			
ISO 1994	-0.150 (0.111)	-0.148 (-0.392; 0.066)	-0.173 (-0.402; 0.057)
ISO 2000	0.130 (0.110)	0.148 (-0.085; 0.381)	0.157 (-0.076; 0.390)
Market share			
ISO 1994	0.040 (0.133)	0.037 (-0.191; 0.266)	0.040 (-0.189; 0.269)
ISO 2000	0.180 (0.136)	0.167 (-0.066; 0.400)	0.177 (-0.056; 0.410)
Employees' satisfaction			
ISO 1994	0.200 (0.142)	0.171 (-0.058; 0.401)	0.182 (-0.047; 0.411)
ISO 2000	0.220 (0.142)	0.195 (-0.038; 0.428)	0.207 (-0.026; 0.440)

^a d effect size conventions (Cohen, 1988): small: 0.20; medium: 0.50; large: 0.80.

* $p < 0.05$.

Table 4

Latent mean differences along performance dimensions between both groups of certified companies. ISO 9000:1994 certified companies are the reference group.

	Latent mean differences (standard error)	Cohen's <i>d</i> effect size ^a (95% CI) 90% reliability	Cohen's <i>d</i> effect size ^a (95% CI) 0% reliability
Unit production costs	0.240 (0.108)*	0.216 (0.035; 0.397)	0.229 (0.048; 0.410)
Fast delivery	−0.090 (0.109)	−0.080 (−0.261; 0.101)	−0.085 (−0.266; 0.096)
Flexibility	−0.100 (0.103)	−0.094 (−0.275; 0.087)	−0.100 (−0.281; 0.081)
Cycle time	−0.010 (0.102)	−0.009 (−0.190; 0.171)	−0.010 (−0.190; 0.170)
Design quality	0.290 (0.108)*	0.261 (0.080; 0.442)	0.277 (0.096; 0.458)
Manufacturing quality	0.130 (0.098)	0.129 (−0.052; 0.310)	0.137 (−0.044; 0.318)
Customers' satisfaction	0.280 (0.091)*	0.299 (0.118; 0.480)	0.317 (0.135; 0.499)
Market share	0.140 (0.103)	0.133 (−0.048; 0.314)	0.141 (−0.040; 0.322)
Employees' satisfaction	0.020 (0.113)	0.017 (−0.163; 0.197)	0.018 (−0.162; 0.198)

^a *d* Effect size conventions (Cohen, 1988): small: 0.20; medium: 0.50; large: 0.80.

* $p < 0.05$.

Regarding the differences between ISO 9000/1994 and ISO 9001/2000 certified companies along performance dimensions, these may be also analyzed using the non-ISO certified companies as a reference. As Table 3 shows, in all performance dimensions, differences statistically exist for certain dimensions. In order to increase the power of the test, we tested a new model, M5, considering only two groups of certified companies. ISO 9000/1994 certified companies were taken as the reference group. Reinforcing the results of previous contrasts, M5 fitted very well: $SB\chi^2$: 0.000 (18); $p = 1.000$, so latent means differed. Effect sizes are showed in Table 4.

Taken these results together, we observe that small-medium differences exist between both groups of companies in some of the performance dimensions. ISO 9001/2000 certified companies had better scores on unit product costs, design quality and customer's satisfaction, but again these differences could not be considered large. Interpretation of results in the remaining dimensions is not so clear because confidence interval of effect sizes included zero, and consequently, positive or negative small differences could exist in these dimensions.

4.3. ISO and TQM dimensions

To test for differences in TQM dimensions between ISO 9000/1994-certified companies, ISO 9001/2000-certified companies and non-certified companies, MACS analysis was again conducted. The research model is similar to the prior model shown in Fig. 1. In this case, the model consisted of seven dependent variables (the TQM dimensions of customers, suppliers, workforce, information, design, process management and leadership), and one independent variable with three levels (ISO 9000/1994 certified companies, ISO 9001/2000 certified companies, and non-certified companies).

Again we considered our model as a formative model with a reflective latent respecification. Using the same reasoning as presented below, we considered TQM dimensions as formative instead of reflective and followed the same analysis procedure as previously discussed. As each TQM dimension was measured using the Flynn et al.'s (1994) scale, we created an index for each dimension. In other words, those dimensions were composite latent variables. Items measuring each dimension composed an

index for their respective dimensions. And finally, all dimensions composed the TQM latent variable.

Traditionally, Flynn et al.'s (1994) scale has been considered as reflective, but we regard it as being formative. For example, for the supplier relationship dimension, such following items as “we strive to establish long-term relationships with suppliers” and “we rely on a small number of high quality suppliers” are measuring two distinct things: long-term relationships and high quality suppliers, respectively. Both items do not necessarily have to be correlated and they are not interchangeable. The same occurs for the items measuring the remaining TQM dimensions. For example, for the process management dimension, the items such as “processes in our plan are designed to be foolproof” and “our plant is kept clean at all times” are again measuring two disparate things: foolproof processes and cleanness. Consequently, items form their respective dimensions in accordance with the formative view. As we did in the case of performance, we conducted several empirical analyses that not supported the reflective perspective.

We applied MACS analysis following the same steps as before. Table 5 shows the model testing sequence. Interpretation of the LISREL models is analogous to the results shown in Table 3. The results indicate that (1) ISO certified companies scored higher in at least one of the TQM dimensions than non-certified companies, and (2) ISO 9000/1994 and ISO 9001/2000-certified companies also differ in at least one of the seven dimensions.

The differences in each dependent variable are shown in Table 6. Having demonstrated a statistical difference

Table 5

Fit indices for the model-testing sequence (ISO-TQM).

Model	$SB\chi^2$ (df)	<i>p</i> -Value	$p(\Delta SB\chi^2)$
M0 _{G1}	0.736 (7)	0.998	
M0 _{G2}	0.059 (7)	1.000	
M0 _{G3}	2.451 (7)	0.931	
M2	82.205 (35)	0.000*	
M3	11.840 (21)	0.953	M2 – M3 = 0.000*
M3 ₂	56.220 (28)	0.001*	M3 ₂ – M3 = 0.000*
M3 ₃	29.661 (28)	0.380	M3 ₃ – M3 = 0.000*
M4	48.165 (21)	0.000*	
M5	8.085 (14)	0.088	

* $p < 0.05$.

Table 6

Latent mean differences along TQM dimensions. Non-ISO certified companies are the reference group.

	Latent mean differences (standard error)	Cohen's <i>d</i> effect size ^a (95% CI) 90% reliability	Cohen's <i>d</i> effect size ^a (95% CI) 80% reliability
Customers			
ISO 1994	0.041 (0.026)	0.198 (−0.031; 0.427)	0.210 (−0.019; 0.439)
ISO 2000	0.054 (0.026) [*]	0.270 (0.037; 0.503)	0.286 (0.052; 0.520)
Suppliers			
ISO 1994	−0.011 (0.022)	−0.061 (−0.290; 0.168)	−0.064 (−0.293; 0.165)
ISO 2000	0.024 (0.022)	0.141 (−0.092; 0.374)	0.150 (−0.083; 0.383)
Workforce			
ISO 1994	−0.024 (0.024)	−0.125 (−0.354; 0.104)	−0.133 (−0.362; 0.096)
ISO 2000	0.006 (0.024)	0.032 (−0.201; 0.265)	0.034 (−0.199; 0.267)
Information			
ISO 1994	0.129 (0.027) [*]	0.599 (0.366; 0.832)	0.636 (0.402; 0.870)
ISO 2000	0.156 (0.028) [*]	0.708 (0.469; 0.947)	0.751 (0.511; 0.991)
Design			
ISO 1994	0.003 (0.021)	0.018 (−0.211; 0.247)	0.019 (−0.210; 0.248)
ISO 2000	0.051 (0.021) [*]	0.330 (0.096; 0.564)	0.350 (0.116; 0.584)
Process management			
ISO 1994	0.000 (0.021)	0.000 (−0.229; 0.229)	0.000 (−0.229; 0.229)
ISO 2000	0.041 (0.020) [*]	0.261 (0.028; 0.494)	0.276 (0.042; 0.510)
Leadership			
ISO 1994	−0.008 (0.023)	−0.044 (−0.273; 0.185)	−0.047 (−0.276; 0.182)
ISO 2000	0.039 (0.023)	0.276 (0.042; 0.510)	−0.076 (−0.309; 0.157)

^a *d* effect size conventions (Cohen, 1988): small: 0.20; medium: 0.50; large: 0.80.^{*} *p* < 0.05.

between the three groups, the effect size was thus again ascertained the substantive difference. The results showed that ISO 9000/1994 certification had a significant effect only on one TQM dimension: information. This effect was considered medium-large, in accordance with Cohen's (1988) conventions. In addition, ISO 9001/2000 certification yielded a significant effect on five TQM dimensions: customers, information, design, process management and leadership. The size of the effect for information was noticeable.

Regarding the differences between ISO 9000/1994 and ISO 9001/2000 certified companies along the TQM dimensions, these were again analyzed using the non-ISO certified companies as a reference. As Table 6 shows, in all TQM dimensions, statistical differences exist for certain dimensions, but they are relatively small from a practical viewpoint. We again conducted a focused test (M5), in order to increase power. Effect sizes are showed in Table 7.

We conclude that small-medium differences exist between both groups of companies in some of the TQM dimensions. ISO 9001:2000 certified companies show better scores on suppliers, design, process management and leadership, but again these differences cannot be considered large. Interpretation of results in the remaining dimensions is more difficult because confidence interval of effect sizes included zero. However, compared to the performance dimensions, the TQM dimensions provide a context where ISO 9001/2000 companies outperform ISO 9000/1994 companies more clearly.

4.4. Sensitivity analysis

We had to be confident about the theoretical meaning of the latent variables was strong because reliability coefficient for indicators was fixed to 90%. In other words, we must examine the sensitivity of our analysis results to

Table 7

Latent mean differences along TQM dimensions between both groups of certified companies. ISO 9000:1994 certified companies are the reference group.

	Latent mean differences (standard error)	Cohen's <i>d</i> effect size ^a (95% CI) 90% reliability	Cohen's <i>d</i> effect size ^a (95% CI) 80% reliability
Customers	0.013 (0.018)	0.069 (−0.112; 0.250)	0.073 (−0.108; 0.254)
Suppliers	0.035 (0.016) [*]	0.208 (0.027; 0.389)	0.217 (0.036; 0.398)
Workforce	0.030 (0.018)	0.161 (−0.020; 0.342)	0.171 (−0.010; 0.352)
Information	0.027 (0.021)	0.124 (−0.057; 0.305)	0.132 (−0.049; 0.313)
Design	0.048 (0.015) [*]	0.323 (0.141; 0.505)	0.343 (0.161; 0.525)
Process management	0.041 (0.016) [*]	0.254 (0.073; 0.435)	0.269 (0.088; 0.450)
Leadership	0.047 (0.016) [*]	0.277 (0.096; 0.458)	0.294 (0.113; 0.475)

^a *d* effect size conventions (Cohen, 1988): small: 0.20; medium: 0.50; large: 0.80.^{*} *p* < 0.05.

verify the reasonableness of this assumption. This approach is common in other methodologies such as system dynamics (Sterman, 2002), and it is also desirable in the context of structural equation modelling (Hayduk (1996)). Given the small effect sizes encountered, it is expected that results will be affected when decreasing reliability.

Therefore, we conducted the latent means difference test fixing the reliability of each indicator to 80% (i.e. measurement error was fixed to 20%). Results are shown in Tables 2, 3, 5 and 6. As expected, effect sizes increased but they increased by a tiny amount. In other words, the interpretation of results is not significantly affected.

Combining the empirical evidences shown by the MACS analyses, we conclude that ISO certified companies do not perform noticeably better than non-certified companies. The results suggest that ISO 9001/2000 certified companies apply TQM in a more powerful way than ISO 9000/1994 certified companies. However, how this higher level of TQM manifests in performance is less clear.

5. Discussion and conclusions

Our results extend the debate involving the ISO 9000's impact on company performance. Previous studies published in JOM analyzed the effect of ISO 9000/1994 in different performance variables. However, a new version of the standard was born in 2000 with apparent characteristics that would bring companies nearer to a TQM system—the latest version of the standard, the ISO 9001/2008, has not implied any important change in its philosophy. Scholars (Lee et al., 1999; Zhu and Scheuermann, 1999) have argued that TQM offers a much more complete quality management system with clear implications for benefits, whereas the previous version of ISO 9000 reflects a mere quality assurance system. If this is true, companies that adopted the ISO 9001/2000 version of the standard should lead companies to a higher level of TQM practices and, consequently, better performance.

Our hypotheses were developed along the lines of such reasoning. However, our empirical analyses do not corroborate exactly what was initially expected. Firstly, what we found was that companies certified by the 1994 and the 2000 version of the standard showed barely noticeable differences in the performance dimensions. ISO 9001/2000 certified companies had slightly better scores on unit product costs, design quality and customer's satisfaction, but these differences could not be considered significant. These differences were also small in comparison with non-ISO companies.

In 2006, JOM published a special issue encouraging replication research underlying OM models and theories (volume 24, 2006). In that issue, Kaynak and Hartley (2008) affirmed that replication implies that results can be generalized with confidence, building the foundation for theory. Many previous studies have empirically analyzed the impact of ISO 9000 implementation on the company performance. At the moment, there is still lack of agreement about its benefits. In this regard, this paper attempted to shed light on this issue and push it a little further.

Our results in essence support the study of Terziovski et al. (1997) since it does not find significant differences between certified and non-certified firms. Terziovski et al. posited in their conclusion that ISO 9000 could contribute to organizational performance if a climate of change was created, which was not happening at the time of their study. This climate of change was subsequently examined in Naveh and Marcus (2005). They found in their longitudinal study that, just as Terziovski et al. predicted, the climate of change was an important moderating variable in the ISO 9000 and performance relationship. Companies that implemented the standard as a catalyst for change would get the most from it. Other studies have also confirmed motivation as a moderating variable (Martínez-Costa et al., 2008).

The version of the standard created in 2000 introduced a few key aspects of quality largely neglected in the previous version such as continuous improvement or customer orientation. These two aspects are of great importance, and without them, as was the case under the 1994 version, the companies would be limited to implement quality management at a superficial level. Companies certified under the 2000 version are motivated to explicitly define ways to continuously improve quality and have the customer satisfied. The next logical question then becomes: is the 2000 version being implemented as a catalyst for change? The answer based on our results is not encouraging. Companies that implemented the 2000 version do not show significantly large improvements compared with companies with the 1994 version or non-certified ones.

Benner and Veloso (2008) posited that, with regard to the 1994 version of the standard, the early adopter might have gained some benefits from the implementation. Similarly, we tried to take into account that the ISO 9001/2000 companies analyzed in our sample were the first ones in implementing the 2000 version. However, we are not able to support such early adopter argument based on our results. We suspect that the 1994 version might have developed an "aura" about it so that it was believed to be useful only for making some sales or as a requisite condition for gaining access to potential customers. If this aura has preceded the 2000 version, then it seems that the 2000 version does not lead to real changes in companies certified under the 2000 version.

An important contribution of our paper is that we empirically studied if the changes in the standard, which are more in line with the TQM philosophy, were in fact leading companies to the higher levels in the TQM dimensions. We had stated that the new ISO 9001/2000 should present higher levels in TQM and, particularly in those considered as "soft dimensions" by previous literature, mainly, leadership and management commitment, continuous improvement, customer management, and human resource management. Our findings were that ISO 9001/2000 certified companies had better scores on suppliers, design, process management and leadership. As it can be seen, from the improvements of the soft variables included in the new standard identified in the literature review only leadership has slightly improved. The other two dimensions in which our results showed differences

are considered of a technical nature or “hard dimensions” in the literature.

Our results show that ISO 9001/2000 certified companies apply TQM at a higher level than ISO 9000/1994 certified companies. However, the superiority of the 2000 version is not clearly manifested in performance. Possible reasons are, firstly, that previous literature in the field has identified that the TQM dimensions that most affect performance are the “soft ones”. If the 2000 version of the standard does not affect those dimensions significantly, then performance will not improve significantly. Secondly, it is important to note that, the data collection for the present study occurred only after a few years of the implementation of the 2000 version. Therefore, a longitudinal study would be necessary to make a definitive statement about the impact of this standard on TQM dimensions and company performance.

Recommendations to managers from our analyses regarding this 2000 version of the standard are in line with the previous studies published in JOM that analyzed the 1994 version of ISO 9000. Whatever its version is, ISO standard is not a guarantee of quality or better performance. There are practical implications of this realization involving the 1994 and 2000 versions of ISO. One, the selection of suppliers or business partnership should not depend solely on the possession of an ISO 9001 certifica-

tion. Two, when a company wants to improve its quality management system, the application in isolation of the ISO 9001 model is not a secure way of getting this improvement.

Our study has limitations. The variability of the effect size estimates (e.g. confidence interval are not narrow) does not allow us to draw clearer conclusions with respect to the importance of the differences. Future studies should use larger samples to reduce such variability. Nevertheless, an important contribution of our research is providing effect sizes, so that other researchers could replicate our findings and conduct a meta-analysis. We consider meta-analysis a highly desirable approach in our field of operations management, so that we could continue to expand and build empirically based theories and would know about the empirical distribution of the effects we find. Therefore, our conclusions could be revisited and reinforced or rebuffed after the analysis of several replications in other test situations (Hitchcock, 2002).

Acknowledgements

We would like to thank the Fundación Séneca and Fundación Cajamurcia for financial support. We also acknowledge two anonymous referees and the associate editor for valuable comments and insights.

Appendix A

Mark your level of agreement/disagreement in these statements (1: completely agree; 5: completely disagree).

A.1. Dimension 1: top managers support

Quality leadership:

All major department heads within our plant accept their responsibility for quality	1	2	3	4	5
Plant management provides personal leadership for quality products and quality improvement	1	2	3	4	5
The top priority is evaluating plant management in quality performance	1	2	3	4	5
Our top management strongly encourages employee involvement in the production process	1	2	3	4	5

Quality improvements rewards:

Workers are rewarded for quality improvements	1	2	3	4	5
We pay a group incentive for quality improvement ideas	1	2	3	4	5
Our plant has a annual bonus system based on plant productivity	1	2	3	4	5
Non-financial incentives are used to reward quality improvement	1	2	3	4	5

A.2. Dimension 2: quality information

Process control:

Processes in our plant are designed to be “fool proof”	1	2	3	4	5
A large percent of the equipment or processes on the shop floor are currently under statistical quality control	1	2	3	4	5
We make extensive use of statistical techniques to reduce variance in processes	1	2	3	4	5
Charts showing defect rates are posted on the shop floor	1	2	3	4	5
Charts showing schedule compliance are posted on the shop floor	1	2	3	4	5
Charts plotting the frequency of machine breakdowns are posted on the shop floor	1	2	3	4	5
Workers are always told when they do a good job	1	2	3	4	5
Information on productivity is readily available to employees	1	2	3	4	5
My manager always comments about the quality of my work	1	2	3	4	5

A.3. Dimension 3: process management

Cleanliness and organization:

Our plant emphasizes putting all tools and fixtures in their place	1	2	3	4	5
We take pride in keeping our plant neat and clean	1	2	3	4	5
Our plant is kept clean at all times	1	2	3	4	5
I never have trouble finding the tools I need	1	2	3	4	5
Our plant is disorganized and dirty	1	2	3	4	5

A.4. Dimension 4: product design

New product quality:

New product designs are thoroughly reviewed before the product is produced and sold	1	2	3	4	5
Customer requirements are thoroughly analyzed in the new product design process	1	2	3	4	5
New product's quality is more important than cost	1	2	3	4	5
Quality is more important than schedule concerns in the new product development process	1	2	3	4	5

Interfunctional design process:

Direct labor employees are involved to a great extent before introducing new products or making product changes	1	2	3	4	5
There is little involvement of manufacturing and quality people in the early design of products, before they reach the plant	1	2	3	4	5
We work in teams, with members from a variety of areas to introduce new products	1	2	3	4	5

A.5. Dimension 5: workforce management

Selection for teamwork potential:

We use ability to work in a team as a criterion in employee selection	1	2	3	4	5
We use problem solving ability as a criterion in selecting employees	1	2	3	4	5
We use work values and ethics as a criterion in employee selection	1	2	3	4	5

Teamwork:

Our plant is organized into permanent production teams	1	2	3	4	5
During problem solving sessions, we make an effort to get all team members' opinions and ideas before making a decision	1	2	3	4	5
Our plant form teams to solve problems	1	2	3	4	5
In the past 3 years, many problems have been solved through small group sessions	1	2	3	4	5

A.6. Dimension 6: supplier involvement

Supplier relationship:

We strive to establish long-term relationships with suppliers	1	2	3	4	5
Our suppliers are actively involved in our new product development process	1	2	3	4	5
Quality is our number one criterion in selecting suppliers	1	2	3	4	5
We rely on a small number of high quality suppliers	1	2	3	4	5

A.7. Dimension 7: customer involvement

We frequently are in close contact with our customers	1	2	3	4	5
Our customers often visit our plant	1	2	3	4	5
Our customers give us feedback on quality and delivery performance	1	2	3	4	5

Please, select in which competitive position is your company *in comparison with its competitors* relating to these performance measurements (1: much worse; 7: much better).

Unit production costs	1	2	3	4	5	6	7
Fast deliveries	1	2	3	4	5	6	7
Flexibility to change production volume and inventories	1	2	3	4	5	6	7
Cycle time	1	2	3	4	5	6	7
Design quality	1	2	3	4	5	6	7
Manufacturing quality	1	2	3	4	5	6	7
Customers satisfaction	1	2	3	4	5	6	7
Employees satisfaction	1	2	3	4	5	6	7
Market share	1	2	3	4	5	6	7

References

- Aarts, F., Vos, E., 2001. The impact of ISO registration on New Zealand firms performance: a financial perspective. *The TQM Magazine* 13 (3), 180–191.
- Adams, G., McQueen, G., Seawright, K., 1999. Revisiting the price impact of quality awards. *Omega* 27, 595–604.
- Ahire, S.L., Golhar, D.Y., Waller, M.A., 1996. Development and validation of TQM implementation constructs. *Decision Sciences* 27 (1), 23–56.
- Bagozzi, R.P., 2007. On the meaning of formative measurement and how it differs from reflective measurement: comment on Howell, Breivik, and Wilcox. *Psychological Methods* 12 (2), 229–237.
- Bagozzi, R.P., Fornell, C., 1989. Consistency criteria and unidimensionality: an attempt at clarification. *Advances in Consumer Research* 16, 321–325.
- Beaumont, N.B., Sohal, A.M., Terziovski, M., 1997. Comparing quality management practices in the Australian service and manufacturing industries. *International Journal of Quality and Reliability Management* 14 (8), 814–833.
- Benner, M.J., Veloso, F.M., 2008. ISO 9000 practices and financial performance: a technology coherence perspective. *Journal of Operations Management* 26, 611–629.
- Bergkvist, L., Rossiter, J.R., 2007. The predictive validity of multiple-item vs. single-item measures of the same constructs. *Journal of Marketing Research* 44 (2), 175–184.
- Biazio, S., Bernardi, G., 2003. Process management practices and quality system standards. Risks and opportunities of the new ISO 9001 certification. *Business Process Management Journal* 9 (2), 149–169.
- Black, S.A., Porter, L.J., 1996. Identification of the critical factors of TQM. *Decision Sciences* 27 (1), 1–21.
- Bollen, K.A., 2007. Interpretational confounding is due to misspecification, not to type of indicator: comment on Howell, Breivik, and Wilcox. *Psychological Methods* 12 (2), 219–228.
- Borsboom, D., Mellenbergh, G.J., van Heerden, J., 2003. The theoretical status of latent variables. *Psychological Review* 110 (2), 203–219.
- Bradley, M., 1994. Starting Total Quality Management from ISO 9000. *The TQM Magazine* 6 (1), 50–54.
- Brah, S.A., Wong, Y.L., Madhu Rao, B., 2000. TQM and business performance in the service sector: a Singapore study. *International Journal of Operations and Production Management* 20 (11), 1293–1312.
- Byrne, B.M., Shavelson, R.J., Muthén, B., 1989. Testing for equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychological Bulletin* 105 (May), 456–466.
- Cohen, J. (Ed.), 1977. *Statistical Power Analysis for the Behavioral Sciences* (revised edition). New York: Academic Press.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioural Sciences*, 2nd ed. Erlbaum, Hillsdale, NJ.
- Cohen, J., 1990. Things I have learned (so far). *American Psychologist* 49 (12), 997–1003.
- Coltman, T., Devinney, T.M., Midgley, D.F., Venaik, S., 2008. Formative versus reflective measurement models: two applications of formative measurement. *Journal of Business Research* 61, 1250–1262.
- Corbett, C., Montes-Sancho, M.J., Kirsch, D.A., 2005. The financial impact of ISO 9000 certification in the United States: an empirical analysis. *Management Science* 51 (7), 1046–1059.
- Curkovic, S., Vickery, S., Dröge, C., 2000. Quality-related action programs: their impact on quality performance and firm performance. *Decision Sciences* 31 (4), 885–905.
- Choi, T.Y., Eboch, K., 1998. The TQM paradox: relations among TQM practices, plant performance, and customer satisfaction. *Journal of Operations Management* 17, 59–75.
- Claver, E., Tari, J.J., Molina, J.F., 2002. Areas of improvement in certified firms advancing towards TQM. *International Journal of Quality and Reliability Management* 18(8&9): 1014–1036.
- Diamantopoulos, A., 2006. The error term in formative measurement models: interpretation and modeling implications. *Journal of Modeling in Management* 1 (1), 7–17.
- Diamantopoulos, A., Winklhofer, H., 2001. Index construction with formative indicators: an alternative to scale development. *Journal of Marketing Research* 38 (2), 269–277.
- Dobb, F.P., 2004. *ISO 9001: 2000. Quality Registration Step by Step*, 3rd ed. Elsevier, Oxford.
- Docking, D.S., Downen, R., 1999. Market interpretation of ISO 9000 registration. *The Journal of Financial Research* 22 (2), 147–160.
- Douglas, T.J., Judge, W.Q.J., 2001. Total quality management implementation and competitive advantage: the role of structural control and exploration. *Academy of Management Journal* 44 (1), 58–169.
- Dow, D., Samson, D., Ford, S., 1999. Exploding the myth: do all quality management practices contribute to superior quality performance? *Production and Operations Management* 8 (1), 1–27.
- Easton, G.S., Jarrell, S.L., 1998. The effects of total quality management on corporate performance: an empirical investigation. *Journal of Business* 71 (2), 253–307.
- Ebrahimipour, M., Withers, B.E., Hikmet, N., 1997. Experiences of US and foreign-owned firms: a new perspective on ISO 9000 implementation. *International Journal of Production Research* 37 (2), 567–576.
- Edwards, J.R., Bagozzi, R.P., 2000. On the nature and direction of relationships between constructs and measures. *Psychological Methods* 5, 155–174.
- Escanciano, C., Fernández, E., Vázquez, C., 2001. ISO 9000 certification and quality management in Spain: Results of a national survey. *The TQM Magazine* 13 (3), 192–200.
- Everett, E.A., 1994. Alternative quality improvement practices and organization performance. *Journal of Operations Management* 12, 27–44.
- Flynn, B.B., Schroeder, R.G., Sakakibara, S., 1994. A framework for quality management research and an associated measurement instrument. *Journal of Operations Management* 11, 339–366.
- Flynn, B.B., Schroeder, R.G., Sakakibara, S., 1995. The impact of quality management practices on performance and competitive advantage. *Decision Sciences* 26 (5), 659–691.
- Forker, L.B., Mendez, D., Hershauer, J.C., 1997. Total quality management in the supply chain: what is its impact on performance? *International Journal of Production Research* 35 (6), 1681–1701.
- Forker, L.B., Vickery, S.K., Droge, L.M., 1996. The contribution of quality to business performance. *International Journal of Operations & Production Management* 16 (8), 44–62.
- Forza, C., Filippini, R., 1998. TQM impact on quality conformance and customer satisfaction: a causal model. *International Journal of Production Economics* 55, 1–20.
- Frohlich, M.T., 2002. Techniques for improving response rates in OM survey research. *Journal of Operations Management* 20, 53–62.
- Goetsch, D.L., Stanley, B.D., 1998. *Understanding and Implementing ISO 9000 and ISO Standards*. Prentice Hall, New Jersey.
- Gotzamani, K.D., Tsiotras, G.D., 2001. An empirical study of the ISO 9000 standards' contribution towards total quality management. *International Journal of Operations and Production Management* 21 (10), 1326–1342.
- Gotzamani, K.D., Tsiotras, G.D., 2002. The true motives behind ISO 9000 certification. Their effect on the overall certification benefits and their long term contribution towards TQM. *International Journal of Quality and Reliability Management* 19 (2), 151–169.

- Gupta, A., 2000. Quality management practices of ISO vs non-ISO companies: a case of Indian industry. *Industrial Management & Data Systems* 100 (9), 451–455.
- Gustafsson, A., Nilsson, L., Johnson, M.D., 2003. The role of quality practices in service organizations. *International Journal of Service Industry Management* 14 (2), 232–244.
- Hayduk, L.A. (Ed.), 1996. LISREL Issues, Debates and Strategies. John Hopkins University Press, Baltimore and London.
- Hayduk, L.A., Glaser, D.N., 2000. Jiving the four-step, waltzing around factor analysis, and other serious fun. *Structural Equation Modeling* 7, 1–35.
- Hedges, L.V., Olkin, I., 1985. *Statistical Methods for Meta-analysis*. Orlando: Academic Press.
- Hendricks, K.B., Singhal, V.R., 1996. Quality awards and the market value of the firm: an empirical investigation. *Management Science* 42 (3), 415–436.
- Hendricks, K.B., Singhal, V.R., 2001a. Firm characteristics, total quality management and financial performance. *Journal of Operations Management* 19, 269–285.
- Hendricks, K.B., Singhal, V.R., 2001b. The long-run stock price performance of firms with effective TQM programs. *Management Science* 47 (3), 359–368.
- Hitchcock, C. (Ed.), 2002. Probabilistic Causation. *Stanford Encyclopedia of Philosophy*.
- Howell, R.D., Breivik, E., Wilcox, J.B., 2007. Reconsidering formative measurement. *Psychological Methods* 12 (2), 205–218.
- Hua, H., Chin, K.S., Sun, H., Xu, Y., 2000. An empirical study on quality management practices in Shanghai manufacturing industries. *Total Quality Management* 11 (8), 1111–1122.
- ISO, 2008. ISO annual report.
- Jarvis, C.B., MacKenzie, S.B., Podsakoff, P.M., 2003. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research* 30 (2), 199–218.
- Jöreskog, K., Sörbom, D. (Eds.), 2006. LISREL 8.80. Scientific Software International, Inc..
- Kaynak, H., Hartley, J.L., 2008. A replication and extension of quality management into the supply chain. *Journal of Operations Management* 26, 468–489.
- Lee, T.Y., Leung, H.K.N., Chang, K.C.C., 1999. Improving quality management on the basis of ISO 9000. *The TQM Magazine* 11 (2), 88–94.
- Lima, M.A.M., Resende, M., Hasenclever, L., 2000. Quality certification and performance of Brazilian firms: an empirical study. *International Journal of Production Economics* 66, 143–147.
- Lubke, G.H., Dolan, C.V., Kelderman, H., Mellenberg, G.J., 2003. On the relationship between sources of within- and between-group differences and measurement invariance in the common factor model. *Intelligence* 31, 543–566.
- Malhotra, M.K., Grover, V., 1998. An assessment of survey research in POM: from constructs to theory. *Journal of Operations Management* 16, 407–425.
- Markus, K.A., 2004. Varieties of causal modeling: how optimal research design varies by explanatory strategy. In: van Montfort, K., Oud, J., Satorra, A. (Eds.), *Recent Developments on Structural Equation Models: Theory and Application*. Kluwer Academic Publishers, Dordrecht.
- Martínez-Costa, M., Martínez-Lorente, A.R., 2007. ISO 9000:2000. The key to quality? An exploratory study. *The Quality Management Journal* 14 (1), 7–18.
- Martínez-Costa, M., Martínez-Lorente, A.R., Choi, T.Y., 2008. Simultaneous consideration of TQM and ISO 9000 on performance and motivation: an empirical study of Spanish companies. *International Journal of Production Economics* 113 (1), 23–39.
- Martínez-Lorente, A.R., Dewhurst, F.W., Gallego-Rodríguez, A., 2000. Relating TQM, marketing and business performance: an exploratory study. *International Journal of Production Research* 38 (14), 3227–3246.
- Martínez-Lorente, A.R., Martínez-Costa, M., 2004. ISO 9000 and TQM: substitutes or complementaries? An empirical study in industrial companies. *International Journal of Quality and Reliability Management* 21 (3), 260–276.
- Martínez, J.A., Martínez, L., 2008. Building better causal models to measure the relationship between attitudes and customer loyalty. *International Journal of Market Research* 50 (4), 437–447.
- McAdam, R., Canning, N., 2001. ISO in the service sector: perceptions of small professional firms. *Managing Service Quality* 11 (2), 80–92.
- Naveh, E., Marcus, A., 2005. Achieving competitive advantage through implementing a replicable management standard: installing and using ISO 9000. *Journal of Operations Management* 24, 1–26.
- OECD (Eds.), 2007. *Economic Survey of Spain*. OECD.
- Ployhart, R.E., Oswald, F.L., 2004. Applications of mean and covariance structure analysis: integrating correlational and experimental approaches. *Organizational Research Methods* 7 (1), 27–68.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P., 2003. Common method variance in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88, 879–903.
- Powell, T.C., 1995. Total quality management as competitive advantage: a review and empirical study. *Strategic Management Journal* 16, 15–37.
- Quinn, R.E., Rohrbaugh, J., 1983. A spatial model of effectiveness criteria towards a competing values approach of organizational analysis. *Management Science* 29 (3), 363–377.
- Raykov, T., 2001. Testing multivariable covariance structure and means hypothesis via structural equation modelling. *Structural Equation Modeling* 8 (2), 224–256.
- Reimann, C.W., Hertz, H.S., 1996. The Baldrige Award and ISO 9000 registration compared. *Journal for Quality and Participation* (January & February), 12–19.
- Romano, 2000. ISO 9000: what is its impact on performance? *Quality Management Journal* 7 (3), 38–56.
- Rossiter, J.R., 2002. The coarse procedure for scale development in marketing. *International Journal of Research in Marketing* 19, 305–335.
- Rosster, J.R., 2008. Content validity of measures of abstract constructs in management and organizational research. *British Journal of Management* 19 (4), 380–388.
- Samson, D., Terziovski, M., 1999. The relationship between total quality management practices and operational performance. *Journal of Operations Management* 17, 393–409.
- Saraph, J.V., Benson, P.G., Schroeder, R., 1989. An instrument for measuring the critical factors of quality management. *Decision Sciences* 20 (4), 810–829.
- Satorra, A., Bentler, P.M., 1999. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 66, 507–514.
- Schlickman, J., 2003. ISO 9001:2000 Quality Management System Design. Artech House, Boston.
- Shetty, Y.K., 1993. The quest for quality excellence: lessons from the Malcom Baldrige National Quality Award. *SAM Advanced Management Journal* 58 (42), 34–40.
- Simmons, B.L., 1999. The relationship between ISO 9000 and business performance: does registration really matter? *Journal of Managerial Issues* 11 (3), 330–343.
- Singels, J., Ruel, G., Van der Water, H., 2001. ISO 9000 series. Certification and performance. *International Journal of Quality & Reliability Management* 18 (1), 62–75.
- Skrabec, Q. R. J., 1999. Quality assurance revisited. *Industrial Management*, November&December: 6–9.
- Sörbom, D., 1974. A general method for studying the differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology* 27, 229–239.
- Staw, B.M., Epstein, L.D., 2000. What bandwagons bring: effects of popular management techniques on corporate performance, reputation and CEO pay. *Administrative Science Quarterly* 45, 523–556.
- Steenkamp, J.B., Baumgartner, H., 1998. Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research* 25 (June), 78–90.
- Stephens, B., 1997. Implementation of ISO 9000 or Ford's Q1 award: effect on organizational knowledge and application of TQM principles and quality tools. *The TQM Magazine* 9 (3), 190–200.
- Sterman, J., 2002. All models are wrong: reflections on becoming a systems scientist. *System Dynamics Review* 18 (4), 501–531.
- Sun, H., 2000. Total Quality Management. ISO 9000 certification and performance improvement. *International Journal of Quality & Reliability Management* 17 (2), 168–179.
- Tan, K.C., Handfield, R.B., Krause, D.R., 1998. Enhancing the firm's performance through quality and supply base management: an empirical study. *International Journal of Production Research* 36 (10), 2813–2837.
- Taylor, W.A., 1995. Organizational differences in ISO 9000 implementation practices. *International Journal of Quality & Reliability Management* 12 (7), 10–27.
- Terziovski, M., Samson, D., 1999. The link between total quality management practice and organizational performance. *International Journal of Quality & Reliability Management* 16 (3), 226–237.
- Terziovski, M., Samson, D., 2000. The effect of company size on the relationship between TQM strategy and organizational performance. *The TQM Magazine* 12 (2), 144–148.
- Terziovski, M., Samson, D., Dow, D., 1997. The business value of quality management systems certification. Evidence from Australia and New Zealand. *Journal of Operations Management* 15, 1–18.
- Thompson, B., 2002. Statistical, practical and clinical: how many kinds of significance do counselors need to consider? *Journal of Counseling and Development* 80, 64–71.

- Tummala, V.M.R., Tang, C.L., 1996. Strategic quality management. Malcolm Baldrige and European quality awards and ISO 9000 certification: Core concepts and comparative analysis. *International Journal of Quality & Reliability Management* 13 (4), 8–38.
- Venkatraman, N., Ramanujam, V., 1986. Measurement of business performance in strategy research: a comparison of approaches. *Academy of Management Review* 1 (4), 801–814.
- Wayhan, V.B., Kirche, E.T., Khumawala, M., 2002. ISO 9000 certification: the financial performance implications. *Total Quality Management* 13 (2), 217–231.
- Withers, B.E., Ebrahimpour, M., 2001. Impact of ISO 9000 registration on European firms: a case analysis. *Integrated Manufacturing Systems* 12 (2), 139–151.
- Yusof, S., Aspinwall, E., 2000. TQM implementation issues: Review and case studies. *International Journal of Quality & Reliability Management* 20 (6), 634–655.
- Zhu, Z., Scheuermann, L., 1999. A comparison of quality programmes: total quality management and ISO 9000. *Total Quality Management* 10 (2), 291–297.