Long-run quality performance dynamics and quality cost changes – new empirical evidence

Markus Plewa
University of Erlangen-Nuremberg

Gernot Kaiser
Nordhausen University of Applied Sciences

Evi Hartmann
University of Erlangen-Nuremberg
evi.hartmann@fau.de

Abstract
Dynamics in quality performance and quality cost have been discussed extensively in the Total Quality Management literature. However, besides anecdotal findings from case studies and simulation exercises, little empirical evidence is available. Using unique secondary data we set out to clarify the dynamic mechanics between quality cost and quality performance.

Keywords: Manufacturing, Quality performance, Secondary data analysis

Introduction

The influence of Total Quality Management (TQM) on company performance is well studied (e.g. Kaynak 2003; Nair 2006; Ebrahimi and Sadeghi 2013). Consequently, sustaining and improving quality while carefully watching the cost involved is a key objective in modern manufacturing in order to remain competitive and satisfy customers.

The research questions this paper strives to answer are based on the central axioms of TQM in general but also specifically examine the implications of continuous improvement and quality learning. First, do we observe increasing quality performance in the long run, pointing to the fruits of continuous improvement and the effects of quality learning? Is this also reflected in long run decreasing quality cost per sales volume? Moreover, do companies effectively reduce conformance cost without experiencing higher non-conformance cost in return? Finally, do we find a scale effect in quality learning in the sense that quality cost increases less than proportionately with sales volume?

Despite these being central questions for operations management practice and research, little empirical evidence is available apart from anecdotal findings from case studies and simulation exercises. The most notable studies include Ittner (1996) and Ittner *et al.* (2001).

The remainder of this article is organized as follows. In the next section we provide a brief literature review of the topics touched on as well as background for the motivation of this study. Then we formulate the research hypotheses. Afterwards we describe and discuss our methodology and data. Following this we provide results of the analysis and discuss them. The conclusion and implications close the paper.

Literature review and background

The foundations of TQM literature have been laid by Juran (1962) through his Quality Control Handbook. The first treatment of the topic of quality cost in particular is credited to Feigenbaum (1956). The latest review of the cost of quality literature was conducted by Schiffauerova and Thomson (2006). In addition to an overview on theoretical foundations, they list many case studies, some of which include (short) time series analysis. Williams *et al*. (1999) sum up managerial implications arising from the quality costing literature. Long-run developments however are mostly studied in simulation exercises which are usually based on case studies. Burgess (1996) and Chiadamrong (2003) are influential contributions, while Omar and Murgan (2014) is one of the most recent. Ittner *et al*. (2001) provide long-run empirical evidence but focus on a limited number of production sites of a single company.

An overview of the concepts for continuous improvement is provided by Bhuiyan and Baghel (2005). Ittner *et al*. (2001) establish a link between quality learning and related quality cost dynamics.

We set out to provide large sample evidence for the dynamic behavior of quality cost and quality performance within a narrowly defined group of manufacturing firms in order to find indications of quality learning and continuous improvement. To our knowledge this is the first study using secondary data featuring more than 400 observations.

Research hypotheses

In order to conduct concise empirical analysis we next state the research hypothesis. The first hypothesis is straightforward, and a central result of implementing TQM.

- Hypothesis 1: Quality performance increases in the long run.

The quality cost literature established that with increasing quality performance, quality cost for investments in prevention, for appraisal of intermediates and final products, and for product failure detected both in-house and after dispatch should diminish.

- Hypothesis 2: Quality cost per sales volume decreases in the long run.

To provide evidence for a central belief of TQM we formulate our next hypothesis. Due to lasting effects of investments to achieve better product quality, companies should be able to reduce their conformance cost without facing higher non-conformance cost (internal and external failure cost) as a result. That implies that conformance cost and internal and external failure cost move in the same direction in the long run.

- Hypothesis 3: There is a positive relationship between changes in conformance cost per sales and changes in non-conformance cost per sales.

Learning how to improve is not a question of time alone but rather of the volume produced over time. As quality cost is most commonly standardized by sales volume, a change in quality cost as a fraction of sales does not reveal whether sales growth itself may accelerate the developments in quality cost. We need to clarify whether quality cost grows less than proportionally to sales volume: The percentage change in quality cost should be smaller than the percentage change in sales. This is done by interpreting the empirically estimated elasticity of quality cost with sales volume.

- Hypothesis 4: The long-run elasticity for quality cost with sales volume is smaller than one.

Methodology and data

To address these questions we use a unique data set with 408 observations collected through a questionnaire. Supply chain, production, operations and executive managers were targeted mainly by announcements in a professional magazine. Participation was by self-selection and incentivized by the prospect of receiving an individual feedback report on one's relative position among the other participants.

The data set has two important characteristics of secondary data. First, none of the authors of this study was involved in gathering the data (therefore avoiding interviewer bias). Second, the topic of quality costs was only one of several topics on the comprehensive questionnaire (therefore avoiding common method bias). An introduction to secondary data analysis in operations management research is provided by Calantone and Vickery (2010).

Sower *et al.* (2007) legitimated the use of self-reported data in the context of quality cost research. They set an example of successful validation of self-reported quality related data and hence follow other publications using self-reported quality cost data (e.g. Ittner 1996; Ittner *et al.* 2001).

Some contributors have worried about the validity of comparing quality cost data across companies (e.g. Dale 2007; Williams *et al.* 1999). We are convinced that inter-company comparison of the quality cost data at hand is justified by the homogeneity of participants stemming from a limited range of industry branches, the calculation of within-company differences over time, and the use of a common reporting definition for quality cost.

Survey participants originate from a narrow band of manufacturing industries. In particular, the survey participants are producers of electrical equipment, machinery and other equipment, motor vehicles, trailers and semi-trailers, as well as other transport equipment (ISIC Divisions 27-30, United Nations publication 2008; and earlier versions). Moreover, all represent manufacturing facilities of German, Austrian or Swiss companies.

Each participant reports annual figures for the last year and for three years before that; therefore we dispose of 408 data points, resulting in 204 paired observations, as indicated in table 1. This allows us to work with long run within-company differences.

The common definition used for the reporting of quality cost components is the German norm DIN 55350 (German Institute for Standardization 2008; and earlier versions), which is very similar to the common British Standard 6143-2 (British Standards Institute 1990). On the questionnaire, quality cost components were defined as to include the labor cost that is directly attributable to quality activities. When reporting prevention costs, training, vendor certification, audit, Failure Mode & Effect Analysis, and Root Cause Failure Analysis were to be included. Appraisal costs comprise inspection of work in progress and finished goods, line technicians, test equipment, lab analysis, and related administration. Internal failure costs include cost of scrap, rework, waste, etc., while external failure costs include cost of warranty, returned material, and related administrative expenses. The details are in line with a frequently quoted definition by

Campanella (1990). Next to this four-part definition we use the more convenient terms *cost of conformance* to sum up prevention and appraisal and *cost of non-conformance* to sum up internal and external failure cost (Crosby 1979; but also e.g. Ittner *et al.* 2001). A detailed discussion on definitions of quality cost and its components has been published by Dale and Plunkett (1999, p. 47-66).

Table 1 – Distribution of observations

1997 & 2000	**’98 & ’01**	**’99 & ’02**	**’00 & ’03**	**’01 & ’04**	**’02 & ’05**	**’03 & ’06**	
7 & 7	7 & 7	6 & 6	12 & 12	15 & 15	19 & 19	17 & 17	
’04 & ’07	**’05 & ’08**	**’06 & ’09**	**’07 & ’10**	**’08 & ’11**	**’09 & ’12**	**’10 & ’13**	**total**
16 & 16	15 & 15	18 & 18	16 & 16	16 & 16	14 & 14	26 & 26	204 & 204 = 408

The data set was collected between 2001 and 2014. As we do not have longitudinal data available but cross-sectional data of paired observations with a two-year gap in between, we need to pool across the year groups. Composing the sample by pooling across time is justified by four important arguments.

First, no significant change across the data collection period is recognizable: Common testing techniques do not reveal a recognizable trend among the year groups. Neither the parametric ANOVA test nor the non-parametric Kruskal-Wallis and Jonckheere-Terpstra tests (Kruskal and Wallis 1952; Jonckheere 1954; Terpstra 1952) point to significant discrepancies between the year groups. Please refer to Appendix 1 for the details. Table 1 provides the distribution of the observations.

Second, using within-company differences eradicates firm-specific particularities. Moreover, the exceptionally large sample builds trust in statistic inference. Finally, all questionnaire responses have been manually checked on receipt for consistency and data errors.

Table 2 – Descriptive statistics (Second of paired observations)

	n	**Mean**	**Std. deviation**	**10%ile**	**Median**	**90%ile**
Prevention cost per sales	204	0.005	0.006	<0.001	0.003	0.013
Appraisal cost per sales	204	0.009	0.011	0.006	0.001	0.019
Conformance cost per sales	204	0.014	0.014	0.002	0.009	0.028
Internal failure cost per sales	204	0.006	0.006	0.001	0.004	0.013
External failure cost per sales	204	0.006	0.007	<0.001	0.003	0.016
Non-conformance cost per sales	204	0.012	0.010	0.002	0.010	0.027
Total quality cost per sales	204	0.026	0.020	0.007	0.020	0.051
Sales (k Euros)	204	164,972	198,923	33,435	115,360	512,000
Sales growth between paired observations	204	0.461	1.013	-0.162	0.305	0.913
Employees of plant	204	632	568	176	471	1,307
Defect rate %	168	0.044	0.089	0.002	0.020	0.085

On the questionnaire, quality cost components have been reported in Euros in absolute terms. We standardize them for most of the analysis by dividing them by the sales volume. On average, total quality cost sums up to 2.6 percent of sales. This figure is lower than what is regularly reported, for instance by Williams *et al.* (1999). In our sample, we find an equal split between conformance and non-conformance cost.

The defect rate is only available for a subsample of 336 observations. We standardize it by the natural logarithm following Ittner *et al.* (2001). It is defined as 100 less the first-pass yield of the production. Due to availability of data, this is different from the usual definition of errors per million opportunities.

Detailed descriptive statistics can be found in table 2. Table 3 states pairwise correlations among the relevant variables. Total quality cost per sales and its components are positively correlated among each other, negatively with sales volume and mostly positively with the log-Defect rate.

Table 3 – Pairwise correlations (Second of paired observations)

	Prev.	**Appr.**	**Conf.**	**Int. fail.**	**Ext. fail.**	**Non-conf.**	**Total**	**Sales**
Appraisal cost per sales	0.40***							
Conformance cost per sales	0.71***	0.93***						
Internal failure cost per sales	0.33***	0.29***	0.36***					
External failure cost per sales	0.08	0.24***	0.22***	0.04				
Non-conformance cost per sales	0.27***	0.36***	0.39***	0.65***	0.78***			
Total quality cost per sales	0.63***	0.83***	0.89***	0.57***	0.54***	0.76***		
Sales (k Euros)	-0.19***	-0.16**	-0.20***	-0.26***	-0.06	-0.20***	-0.24***	
log-Defect rate	0.11	-0.04	0.01	0.18**	0.09	0.18**	0.10	-0.07

Note: ***Significantly different from zero at 0.01 level, ** at 0.05 level, * at 0.10 level, two-tailed levels, respectively.

We set out to support the stated hypotheses using common test statistics. We use the parametric ANOVA and t-test as well as the non-parametric Kruskal-Wallis and Jonckheere-Terpstra tests. Moreover, we use ordinary least-squares (OLS) regression analysis to obtain empirical elasticity parameters and to assess the magnitude of correlations.

Empirical analysis and discussion

Having described the data set and the methods employed, we start off testing the stated hypotheses one by one and interpret our results as we go.

Hypotheses 1 and 2

In order to test the validity of hypotheses 1 and 2, we test the significance of the difference between the paired observations to be non-zero. That is, we test whether we observe a significant change between the first and the second observation of quality performance and quality cost across participants. The results are displayed in Table 4. Using both parametric and non-parametric tests, we find strong evidence rejecting the alternative of the values being equal to zero. Solely for prevention cost we do not find significant results. Nevertheless, median conformance cost is significantly below zero. Quality performance increases while quality cost per sales decrease in the long run in our sample.

Table 4 – Statistics related to hypotheses 1 and 2

	n	Mean	Median	Std. deviation
Δ Prevention cost per sales	204	>-0.0001	<0.0001	0.0053
Δ Appraisal cost per sales	204	-0.0006	-0.0003***	0.0050
Δ Conformance cost per sales	204	-0.0006	-0.0003**	0.0091
Δ Internal failure cost per sales	204	-0.0022***	-0.0009***	0.0054
Δ External failure cost per sales	204	-0.0012***	-0.0002***	0.0059
Δ Non-conformance cost per sales	204	-0.0034***	-0.0017***	0.0084
Δ Total quality cost per sales	204	-0.0040***	-0.0024***	0.0146
Δ log-Defect rate	145	-0.5884***	-0.5108***	0.7954

Note: ***Significantly different from zero at 0.01 level, ** at 0.05 level, * at 0.10 level, two-tailed levels, respectively. We use t-test (mean) and Wilcoxon signed-rank test (median)

Hypothesis 3

With our next step we build support for hypothesis 3. Table 5 provides the details for testing it. We find a strong and significant positive relationship between changes in conformance cost per sales and changes in non-conformance cost per sales. When companies achieve a decrease in internal or external failure cost, it does not happen at the expense of higher prevention and appraisal cost. On average, for one percentage point saved on non-conformance cost per sales, participating companies have been able to cut 0.43 percentage points of conformance cost per sales.

Table 5 – OLS regressions related to hypothesis 3

n=204	Δ Conformance cost per sales	Δ Conformance cost per sales
Δ Internal failure cost per sales	0.5887***	
Δ External failure cost per sales	0.2881***	
Δ Non-conformance cost per sales		0.4260***
Constant	-0.0011*	-0.0009
R-Squared	0.1711	0.1555
F-Statistic	20.74***	37.20***

Note: ***Significant at 0.01 level; ** at 0.05 level; * at 0.10 level; two-tailed levels, respectively, but F-Statistic one-tailed levels

Hypothesis 4

In our final step we aim to clarify whether the reduction in quality cost per sales is driven by the growth in sales. At the same time we check for a scale effect regarding quality cost. We estimate an empirical elasticity of quality cost regarding sales volume. Of course, we here use the quality cost data non-standardized by sales volume for the regression analysis. Results are given in table 6. We find a very consistent picture – Sales volume grows significantly faster than quality cost. Therefore, we find strong support for a cost-reducing scale effect in sales volume regarding quality cost: For each additional percentage point in long-run sales growth, we observe only 0.41 percentage points additional quality cost.

Table 6 – OLS regressions related to hypothesis 4

n=204	**Δ% Prevention cost**	**Δ% Appraisal cost**	**Δ% Conformance cost**	**Δ% Internal failure cost**	**Δ% External failure cost**	**Δ% Non-conformance cost**	**Δ% Total quality cost**
Δ% Sales	0.4356***	0.4542***	0.4512***	0.5150***	0.7266*	0.5249***	0.4107***
Constant	0.4332***	0.2134**	0.2153***	0.0826	0.4503*	0.0369	0.0756
R-Squared	0.0950	0.1323	0.1598	0.1027	0.0557	0.1404	0.2345
F-Statistic	21.19***	30.80***	38.43***	23.13***	11.91***	32.99***	61.87***

Note: ***Significant at 0.01 level; ** at 0.05 level; * at 0.10 level;
Δ% Sales: one-tailed levels against unit value. All coefficients are significantly different from zero at the 0.01 level;
Constant: two-tailed levels against zero; F-Statistic: one-tailed levels against zero

Conclusion and implications

In the course of our research we have collected strong evidence for how quality performance and quality cost evolve over time. First of all, we do observe increasing quality performance in the long run. This is in line with research findings on continuous improvement and the effects of quality learning. External failure costs per sales volume as well as the defect rate significantly decrease in the long-run, both reflecting increasing quality performance. Internal failure cost decreases significantly over time as well. The evidence for conformance cost reductions is weaker but still convincing: Cost of quality per sales for appraisal significantly decreased over the three-year period. Cost for prevention also decreases on average but not significantly so.

Further, we find a strong relationship between shrinking conformance cost and shrinking non-conformance cost in the long run. Reducing internal and external failure cost does not come with increasing prevention and appraisal expenses. That is, increasing internal quality performance (reducing internal failure cost) and increasing external quality performance (at the client) can indeed be achieved at shrinking expenses for product conformance.

Moreover, we find a significant scale effect in quality learning: Non-standardized quality cost grows significantly slower than sales volume. The long-run empirical elasticity of quality cost with respect to sales volume is positive but significantly smaller than one. Therefore, for the sample we have analyzed, sales volume growth has two effects on quality cost per sales. First, it accelerates learning by the bespoke scale effect. Second, it increases the base against which we standardize quality cost.

As the central result of our study we provide new and powerful evidence for the paradigm that “Quality is free” (Crosby 1979). We do achieve this by utilizing a unique secondary data set of quality cost that reflects long-run changes of quality performance over time.

However, even with more than 400 observations at hand, our results are based on correlations only. The identification of causal effects is beyond the possibilities of the data available. The nature of the data set is not panel data but high-level cost and sales data, in pooled cross-sections containing paired observations over time. Further analysis of the effect on, for instance, profitability is not reasonable, as too many unobserved effects may disturb the picture within the two-year gap between the observations.

Still, the striking managerial implication is that, despite the importance of achieving and maintaining high quality, in order to sustain competitiveness in manufacturing operations, prevention and appraisal costs can successfully be reduced in the long run without the burden of increasing failure costs.

References

Bhuiyan, N., A. Baghel. 2005. An overview of continuous improvement: from the past to the present. *Management Decision* **43**(5): 761–771.

British Standards Institute. 1990. *BS 6143-2: Guide to the Economics of Quality*. *Prevention, Appraisal and Failure Model*. British Standards Institute, London.

Burgess, T. F. 1996. Modeling quality-cost dynamics. *International Journal of Quality & Reliability Management* **13**(3): 8–26.

Campanella, J. (Ed.). 1990. *Principles of Quality Costs, 2nd ed*. ASQ Quality Press, Milwaukee.

Calantone, R. J., S. K. Vickery. 2010. Introduction to the special topic forum: using archival and secondary data sources in supply chain management research. *Journal of Supply Chain Management* **46**(4): 3–11.

Chiadamrong, N. 2003. The development of an economic quality cost model. *Total Quality Management & Business Excellence* **14**(9): 999–1014.

Crosby, P. B. 1979. *Quality is free: The art of making quality certain*. McGraw-Hill, New York.

Dale, B. G. 2007. Quality Costing. Dale, B.G., T. van der Wiele, T., J. van Iwaarden, eds. *Managing quality, 5th ed.* Blackwell publishing, Oxford, 176–199.

Dale, B. G., J. Plunkett. 1999. *Quality Costing, 3rd ed*. Gower, Aldershot.

Ebrahimi, M., M. Sadeghi. 2013. Quality management and performance: An annotated review. *International Journal of Production Research* **51**(18): 5625–5643.

Feigenbaum, A. V. 1956. Total quality control. *Harvard Business Review* **34**(6): 93–101.

German Institute for Standardization. 2008. *DIN 55350-11. Concepts for quality management – Part 11: Supplement to DIN EN ISO 9000:2005*. German Institute for Standardization, Berlin.

Ittner, C. D. 1996. Exploratory evidence on the behavior of quality costs. Operations Research **44**(1): 114–130.

Ittner, C. D., V. Nagar, M. V. Rajan. 2001. An Empirical Examination of Dynamic Quality-Based Learning Models. *Management Science* **47**(4): 563–578.

Juran, J. M. 1962. *Quality Control Handbook*. McGraw-Hill, New York.

Jonckheere, A. R. 1954. A distribution-free k-sample test against ordered alternatives. *Biometrika* **41**(1/2): 133–145.

Kaynak, H. 2003. The relationship between total quality management practices and their effects on firm performance. *Journal of Operations Management* **21**(4): 405–435.

Kruskal, W. H., W. A. Wallis. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* **47**(260): 583–621.

Nair, A. 2006. Meta-analysis of the relationship between quality management practices and firm performance – implications for quality management theory development. *Journal of Operations Management* **24**(6): 948–975.

Omar, M. K., S. Murgan. 2014. An improved model for the cost of quality. *International Journal of Quality & Reliability Management* **31**(4): 395–418.

Schiffauerova, A., V. Thomson. 2006. A review of research on cost of quality models and best practices. *International Journal of Quality & Reliability Management* **23**(6): 647–669.

Sower, V. E., R. Quarles, E. Broussard. 2007. Cost of quality usage and its relationship to quality system maturity. *International Journal of Quality & Reliability Management* **24**(2): 121–140.

Terpstra, T. J. 1952. The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae* **14**(3): 327–333.

United Nations. 2008. *International Standard Industrial Classification of All Economic Activities (ISIC), Rev. 4*. United Nations Publication, New York.

Williams, A. R. T., A. van der Wiele, B. G. Dale. 1999. Quality costing: a management review. *International Journal of Management Review* **1**(4): 441-460.

Appendix

ANOVA test on total cost of quality

Hypothesis: The means of the change of total cost of quality in the year groups are equal.

Result: Hypothesis cannot be rejected at common significance levels.
Test statistic: F = 1.06
$Pr(f \geq F)$ = 0.3986

Kruskal-Wallis test on total cost of quality

Hypothesis: The distributions of the change of total cost of quality in the year groups are equal.

Result: Hypothesis cannot be rejected at common significance levels.
Test statistic: K = 12.805 with 13 d.f.
$Pr(X^2 \geq K)$ = 0.4629 (corrected for ties)

Jonckheere-Terpstra test on total cost of quality

Hypothesis: The distributions of the change of total cost of quality in the year groups are equal.

Result: Hypothesis cannot be rejected at common significance levels.
Test statistic: J* = 0.765
$Pr(|Z| > |J^*|)$ = 0.4442 (ordered alternative in either direction)
$Pr(Z > J^*)$ = 0.7779 (descending ordered alternative)
$Pr(Z < J^*)$ = 0.2221 (ascending ordered alternative)

The test statistics displayed represent the alternative hypotheses for equality of medians in year groups, descending order of medians in year groups, ascending order of medians in year groups.