An economic evaluation of the Baldrige National Quality Program.

By: Albert N. Link & John T. Scott

Link, A., & Scott, J. (January 01, 2006). An economic evaluation of the Baldrige National Quality Program. *Economics of Innovation and New Technology, 15*, 1, 83-100.

Abstract:

All federal programs are accountable for their use of public funds. This paper presents conservative estimates of the net social benefits associated with the Baldrige National Quality Award Program, established within the National Institute of Standards and Technology in 1987. On the basis of survey data from members of the American Society for Quality, we estimate cost savings benefits to members, extrapolate those benefits to the economy as a whole, and compare the benefits to the social costs associated with the Program. Our estimation method implies that the ratio of economy-wide benefits to social costs probably exceeds 207:1, supporting the hypothesis that the public investments in quality-standards infrastructure are worthwhile.

Keywords: evaluation methods | public program evaluation | social benefits and costs | economics | innovation | federal programs

Article:

1 Introduction

In response to the productivity decline in the non-farm US economy in the mid-1970s and early-1980s, a number of economic policy initiatives were introduced in the early- and mid-1980s in an effort to reverse the downward productivity trend by stimulating innovative activities within firms. One such initiative was the Malcolm Baldrige National Quality Improvement Act of 1987 (PL 100-107) that declared:1

> [T]he leadership of the United States in product and process quality has been challenged strongly (and sometimes successfully) by foreign competition, and our Nation's productivity growth has improved less than our competitors over the last two decades; … a national quality award program … in the United States would help improve quality and productivity by—
>
> (A) helping to stimulate American companies to improve quality and productivity for the pride of recognition while obtaining a competitive edge through increased profits,

(B) recognizing the achievements of those companies which improve the quality of their goods and services and providing an example to others,

(C) establishing guidelines and criteria that can be used by businesses, industrial, governmental, and other organizations in evaluating their own quality improvement efforts, and

(D) providing specific guidance for other American organizations that wish to learn how to manage for high quality by making available detailed information on how winning organizations were able to change their cultures and achieve eminence.

[And] there is hereby established the Malcolm Baldrige National Quality Award.

Physically and administratively located at the National Institute of Standards and Technology, the Program has been supported by federal and private funding since its establishment. On the federal side, support for the Program has increased from $ 200,000 in 1988 to $ 5,344,000 in 2000.

Albeit a small program by fiscal standards, the Baldrige National Quality Program may be one of the better-known competitiveness programs sponsored by the government. Certainly, the Malcolm Baldrige National Quality Award is one of the most widely publicized of all public sector sponsored performance awards.

Regardless of the size or visibility of the Program, it is like any federal program accountable for its use of public funds. The Government Performance and Results Act (GPRA) of 1993—to improve the confidence of the American people in the performance capability of the federal government and to improve federal program effectiveness—requires each federal program to develop a process of identifying and quantifying the economic benefits of the program's outcomes.

The purpose of this paper is to present estimates of the net social benefits associated with the Baldrige National Quality Program. In Section 2, various institutional aspects of the Program are overviewed including a discussion of the Award criteria, trend in applications, and anecdotal information related to the social benefits associated with the Program. In Section 3, systematic approaches to program evaluation are discussed, and our counterfactual approach is contrasted with traditional approaches. In Section 4, results from the application of the counterfactual evaluation method are presented. Fundamental to our application of the counterfactual evaluation method is our use of very detailed survey response data that, as a practical matter, will always come from a relatively small sample. Given the smallness of our sample of members of the American Society for Quality (ASQ), we use a statistical procedure that controls for selection into the small sample and uses the standard errors of our estimates to ensure that we arrive at a conservative estimate of the Program's social benefit-to-cost ratio. That conservatively estimated benefit-to-cost ratio is 207:1. Finally, in Section 5, we offer concluding observations.

2 An overview of the Baldrige National Quality Program

Although the Program's federally funded budget has increased since the Program's inception, as shown in Figure 1, the number of applicants to the Program for the Award has not, as shown in Figure 2. There has been speculation that the decline in the number of applicants since 1991 reflects the increasing opportunity costs for organizations to conform to the Baldrige Criteria for Performance Excellence (Tab. I). The increase in applicants for 1999 and 2000 is because the Program's scope was broadened to include awards for educational organizations and health care providers.

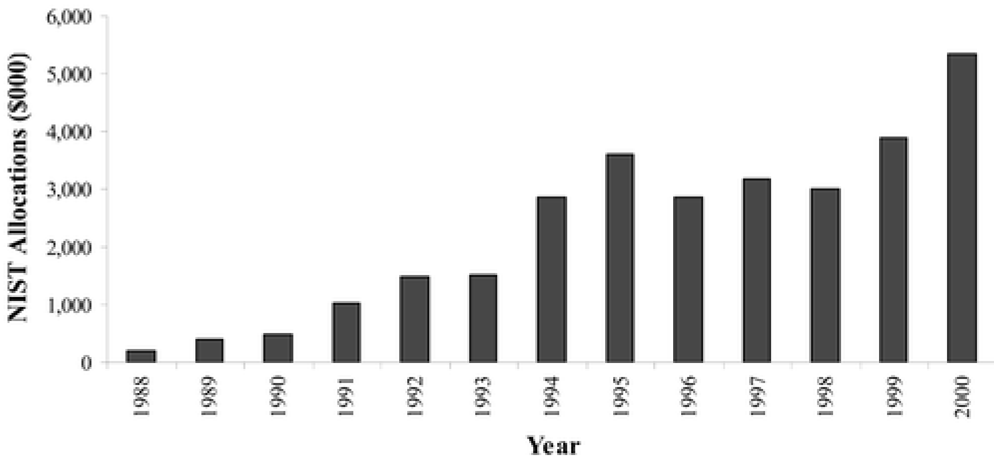

Figure 1 NIST Allocations to the Baldrige National Quality Program (thousands of dollars: $ 000), 1988–2000.

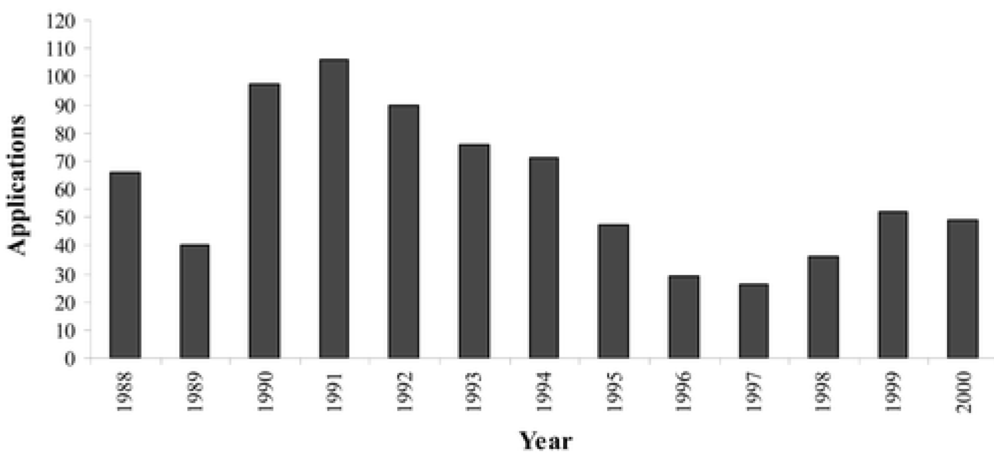

Figure 2 Applications to the Baldrige National Quality Program, 1988–2000.

Table 1 is omitted from this formatted document.

The companies that received the Malcolm Baldrige National Quality Award through 2000 are listed in Table II. At that time, 43 Awards had been announced.

Table II Malcolm Baldrige National Quality Award winners, 1988–2000.

Year	Recipients
1988	Motorola Inc., Westinghouse Electric Corporation—Commercial Nuclear Fuel Division, Globe Metallurgical Inc.
1989	Milliken & Company, Xerox Corporation—Business Products and Systems
1990	Cadillac Motor Car Company, IBM Rochester, Federal Express Corporation, Wallace Co., Inc.
1991	Solectron Corporation, Zytec Corporation, Marlow Industries, Inc.
1992	AT&T Network Systems Group—Transmission Systems Business Unit, Texas Instruments Incorporated—Defense Systems and Electronics Group, AT&T Universal Card Services, The Ritz-Carlton Hotel—Company, Granite Rock Company
1993	Eastman Chemical Company, Ames Rubber Corporation
1994	AT&T Customer Communications Services, GTE Directories Corporation, Wainwright Industries, Inc.
1995	Armstrong World Industries Inc.—Building Products Operation, Corning Incorporated—Telecommunications Products Division
1996	ADAC Laboratories, Dana Commercial Credit Corporation, Custom Research, Inc., Trident Precision Manufacturing, Inc.
1997	3M Dental Products Division, Solectron Corporation, Merrill Lynch Credit Corporation, Xerox Business Services
1998	Boeing Airlift and Tanker Programs, Solar Turbines, Incorporated, Texas Nameplate Company, Inc.

Table II Malcolm Baldrige National Quality Award winners, 1988–2000.

Year	Recipients
1999	STMicroelectronics, Inc.-Region Americas, BI, The Ritz-Carlton Hotel Co., LLC, Sunny Fresh Foods
2000	Dana Corp.-Spicer Driveshaft Division, KARLEE Company, Inc., Operations Management International, Inc., Los Alamos National Bank

Source: Available online at: http://www.quality.nist.gov.

One important motivation for selecting the Baldrige National Quality Program for an evaluation study was the rich empirical literature related to investments in quality that strongly suggested that there were measurable spillover benefits associated with the Baldrige Criteria, and thus with the Award. For example, George and Weimerskirch (1994, pp. 5–6) champion the Baldrige Criteria as the leading model of total quality management with the following observations:

No other model has gained such widespread global acceptance. As evidence, consider these facts.

- Since the Baldrige Program was introduced in 1988, the National Institute of Standards and Technology has distributed more than a million copies of the criteria. It estimates that people have made at least that many copies for their own use.
- More than half the states in the country now have state quality award programs based on the Baldrige criteria.
- Several countries, including Argentina, Australia, Brazil, Canada, and India, are developing or have implemented quality award programs based on the Baldrige criteria.
- The criteria for the European Quality Award, first presented in 1992, are patterned after the Baldrige criteria.
- Companies, such as Honeywell, Intel, IBM, Carrier, Kodak, and AT&T, have adopted the Baldrige criteria as their internal assessment tool and criteria for their corporate quality awards. Many other large companies are asking suppliers to assess their organizations by the Baldrige criteria.

From the mid-1980s to the mid-1990s, the service sector of the US economy grew faster than the non-service sector by an order of magnitude (Scott, 1999). The rapidly evolving service sector is using the Baldrige Criteria to ensure comprehensive management of quality; Blodgett (1999, p. 74) thus observes:

> Service organizations are adopting the criteria in two main ways: They are conducting self-assessments against this robust organizational management model to help identify their strengths and opportunities for improvement, and they are applying for the increasing number of Baldrige-based quality awards in place at the state and local level.

In addition to these general observations about aspects of the social benefits associated with the Program, the Criteria have been adopted by states as a foundation or benchmark for their own quality award programs, thus signifying one dimension of spillover benefits (Tab. III).

Table 3 is omitted from this formatted document.

Table IV provides a concise perspective of the extant empirical economic and management literature related to measurable firm performance effects associated with all aspects of the Program. Although it is beyond the scope of this evaluation paper to discuss the literature in detail, the table does show that scholars have identified several aspects of the social benefits associated with the Program, the Award, and the underlying Criteria. Therefore, our a priori expectations are that the net social benefits associated with the Program are substantial.

Table IV Empirical literature related to measurable firm performance effects associated with the Baldrige National Quality Program.

Aspects of performance	**Authors**
Award winners have stronger financial performance and	Wisner and Eakins (1994)
greater increases in their market value	Helton (1995)
	Lawler *et al*. 1995)
	Huselid and Becker (1996)
	Easton and Jarrell (1998)
	NIST (1996, 1997a, 2000)

Table IV Empirical literature related to measurable firm performance effects associated with the Baldrige National Quality Program.	
Aspects of performance	**Authors**
	Hendricks and Singhal (2001)
Application of Baldrige Criteria improved employee relations,	GAO (1991)
lowered costs, improved customer satisfaction	
Application of Baldrige Criteria improved competitiveness	Council on Competitiveness (1995)
	Banker *et al*. 1998)
Application of Baldrige Criteria increased worker productivity	Ichniowski *et al*. 1995)
	Black and Lynch (1996a,b)

3 Systematic approaches to program evaluation

With any publicly funded program, in principle, the government has an economically justifiable role in supporting investment because of market failures stemming from the public-good nature of the investments associated with the private sector's inability to appropriate returns to the investments or to accept their risks.2 When the public-good nature of investments provides a justifiable role for government in a publicly funded program, systematic program evaluation will demonstrate that the program's social benefits exceed its social costs.

3.1 Traditional economic evaluation methods

Griliches (1958) and Mansfield et al. 1977) pioneered the application of fundamental economic insight to the development of measurements of private and social rates of return to innovative investments. Streams of investment costs generate streams of economic benefits over time. Once identified and measured, these streams of costs and benefits are used to calculate such performance metrics as social rates of return and benefit-to-cost ratios.

For example, for a process innovation adopted in a competitive market, using the traditional framework, the publicly funded innovation being evaluated is thought to lower the cost of producing a product to be sold in a competitive market. As the innovation lowers the unit cost of production, consumers will actually pay less for the product than they paid before the innovation and less than they would have been willing to pay—a gain in consumer surplus. The social benefits from the innovation include the total savings that all consumers receive as a result of producers adopting the cost-reducing innovation. Thus, the evaluation question that can be answered from this traditional approach is ‘given the investment costs and the social benefits, what is the social rate of return to the innovation?’

Asking the question in the foregoing way is not the most appropriate approach from a public accountability perspective. Certainly, the approach allows the evaluation to show the benefits of a socially useful innovation, as intended. However, for publicly funded and publicly performed research, the procedure ignores consideration of the cost effectiveness of the public sector undertaking the research as opposed to the private sector. In other words, the procedure ignores the efficiency with which social benefits are being achieved. Is the public performance less costly than performing the research in the private sector? For publicly funded and privately performed research, the procedure does not by itself distinguish the private rates of return with and without public funding from the social rate of return. As a result, the benefits from the public funding are not identified.

In our opinion, the following ‘counterfactual’ evaluation method is more appropriate for publicly funded and publicly performed infrastructure research and development (R&D) (as well as related operations and maintenance investments in the infrastructure more generally) than the traditional economic approaches.3

3.2 The counterfactual evaluation method

When publicly funded and publicly performed investments are being evaluated, holding constant the economic benefits that the Griliches–Mansfield model measures, and making no attempt to measure that stream, the relevant counterfactual question to ask is ‘what would the private sector have had to invest to achieve those same benefits in the absence of the public sector’s investments?’

The answer to this question yields the benefits of the public investments, namely, the private sector’s costs avoided through the public’s investments plus the benefits from the public sector’s investments that industry would be unable or unwilling to duplicate.4 With those benefits—obtained in practice through extensive interviews with administrators, federal research scientists, and those in the private sector who would have to duplicate the investments in the absence of public performance—counterfactual rates of return and benefit-to-cost ratios can be calculated.

These metrics answer the fundamental evaluation question: are the public investments a more efficient way of generating the technology than the private sector investments would have been?

The answer to this fundamental question aligns with the public accountability issues implicit in GPRA, and certainly addresses a key question of public-sector stakeholders who may doubt the appropriateness of government having a role in the innovation process in the first place. Further, in the context of investments with a public-good nature, the hypothesized answer to the fundamental evaluation question is yes; the counterfactual method tests that hypothesis.

3.3 Evaluation method applicable to the Baldrige National Quality Program

In a broad sense, the Baldrige National Quality Program is a measurement-and-standards infrastructure R&D investment program, with the associated investments in operations and maintenance. Publicly funded and publicly performed infrastructure R&D and related operations and maintenance investments occur within the Program in the sense that therein the Baldrige Criteria were originally developed and therein, through the Baldrige Award process, appropriate applications of the criteria for performance excellence are evaluated. In this broad sense, the Baldrige National Quality Program is similar to a NIST laboratory that performs infrastructure technology R&D investments and sets performance standards (i.e. the Baldrige Criteria) and then continually calibrates bench standards used in private-sector laboratories to achieve a predetermined level of performance (i.e. the Baldrige Award process).

Thus, we apply the counterfactual evaluation method to the evaluation of the Baldrige National Quality Program. Benefits to the economy from the Program are systematically quantified in terms of the cost savings organizations realized by having the Baldrige Criteria to follow as opposed to organizations, on their own, developing and testing comparable criteria.

Benefit data were collected through surveys to selected members of the ASQ and then extrapolated to the aggregate economy as discussed in the following section. Cost data were provided by the Baldrige National Quality Program Office at NIST. The relevant evaluation metric is a benefit-to-cost ratio, with all benefits and all costs referenced to year 2000.

4 Application of the counterfactual evaluation method to the Baldrige National Quality Program

4.1 American society for quality5

The ASQ agreed to a request from the management of the Baldrige National Quality Program Office on our behalf to have a mail survey distributed to its 875 US private-sector companies and public-sector organizations (hereafter 'members').6

ASQs stated mission is to advance individual and organizational performance excellence on a worldwide basis by providing members opportunities for learning, quality improvement, and knowledge exchange. As stated at its web site, the Society's objectives for 2000 are as follows:

- to be our members' best resource for achieving professional and organizational excellence;
- to be a worldwide provider of information and learning opportunities related to quality;
- to be the leader in operational excellence and delivering customer value;
- to be the recognized leader worldwide for advancing individual and organizational performance excellence.

The Society was formed on 16 February 1946.

4.2 Social costs of operating the Baldrige National Quality Program

The public source of funds for the Baldrige National Quality Program is an annual allocation from the NIST budget. Column (2) of Table V shows the Program's annual allocations from NIST by fiscal year beginning with its first year of operation, 1988.

Table V Baldrige National Quality Program operating costs.

(1) Fiscal year	(2) NIST allocations ($)	(3) Foundation allocations ($)	(4) Company reimbursed examiner expenses ($)	(5) Examiner time (h)	(6) Total operating costs ($2000)
1988	200,000	600,000	190,000	37,995	3,689,349
1989	408,000	600,000	190,000	37,995	3,910,205
1990	488,000	600,000	190,000	37,995	3,951,030
1991	1,018,000	600,000	190,000	46,510	5,059,093
1992	1,482,000	600,000	190,000	49,763	5,750,259
1993	1,525,000	600,000	190,000	46,223	5,516,050
1994	2,860,000	728,973	190,453	45,944	7,072,918
1995	3,611,000	694,669	188,137	51,259	8,092,820
1996	2,865,000	652,017	160,230	44,143	6,683,663

1997	3,174,000	778,600	171,803	44,090	7,073,404
1998	3,010,000	808,713	157,879	43,662	6,840,293
1999	3,877,000	1,159,337	186,052	51,735	8,553,566
2000	5,334,000	1,187,543	160,363	51,349	9,891,218

Notes: Column (2): NIST allocation data were provided by the Award office. For inclusion in column (6), these data were inflated to $ 2000 using the chain-type price index for gross domestic product from Table B-7, ‘Chain type-price indexes for gross domestic product, 1959–2000’ Council of Economic Advisers (2001, p. 284). Column (3): Foundation allocation data were provided by the Award office for 1994–2000. The upper-bound on pre-1994 data was estimated (italics), with advice from the Award office. For inclusion in column (6), these data were inflated to $ 2000 using the chain-type price index for GDP in Council of Economic Advisers (2001, p. 284). Column (4): Foundation reimbursements of 70% were paid in 1999 and 2000 for examiners in the education and health care areas; all other examiners were reimbursed at 60% of their expenses. From these data, provided by the Award office, company reimbursed expenses were calculated for 1994–2000. The upper-bound on pre-1994 company costs was estimated (italics), with advice from the Award office. For inclusion in column (6), these data were inflated to $ 2000 using the chain-type price index for GDP in Council of Economic Advisers (2001). Column (5): Examiner time was provided by the Award office. The upper-bound on pre-1990 examiner time was estimated (italics), with advice from the Award office. Based on the management background of the numerous examiners involved in the program, the Award office estimates that the current fully burdened value of a man-year of examiner time is $125,000 ($2000 based on 2000 h per year). The estimated value of examiner time is included in column (6) without additional adjustment.

The Malcolm Baldrige National Quality Improvement Act of 1987 states that:

> The Secretary [of Commerce] is authorized to seek and accept gifts from public and private sources to carry out the program.

In addition to the public funding through NIST, there are private sources of funds. The Program was initially endowed by private industry with $ 10 million. A Foundation was established to manage these funds and to allocate the interest earned to the Program for award ceremonies, publication costs, and partial training and travel costs for examiners whose companies would not pay for such expenses. In column (3) of Table V are the Program’s annual allocations from the Foundation. In column (4) are annual estimates of company expenditures for examiner travel that were not reimbursed by the Foundation through the Program.7

Industry also supports the Program through volunteer examiners during the application and evaluation process. In column (5) of Table V are the total man-hours of examiner time devoted to training, application review, and site visits.

Column (6) of Table V reports the estimated Program costs in year-2000 dollars (year-2000 dollars will be denoted as '$ 2000' from this point on), by year. The present value of these costs, brought forward at the real social rate of return of 7% to account for the social opportunity costs of these funds following the guidelines of OMB (1992), is $118,617,000.

Thus, $119 million (rounded in $2000) is used to represent the present value of the total social costs (to date) associated with the Baldrige National Quality Program.

4.3 Social benefits associated with the Baldrige National Quality Program

A five-step approach is used to estimate the net social benefits associated with the Baldrige National Quality Program. Each step is discussed subsequently in detail, but here is a brief overview.

Benefit data were collected by survey from a sample of the membership of ASQ. These benefit data were extrapolated first to the ASQ membership as a whole and then to the economy as a whole.

The present value of the conservative estimate of the net private benefits received by the ASQ members as a result of the Baldrige National Quality Program is $2.17 billion (rounded in $2000).

If the entire economy benefits to the same extent as the ASQ members, the present value of the conservative estimate of the net social benefits associated with the Baldrige National Quality Program is $24.65 billion (rounded in $2000).

The net private benefits to ASQ members and net social benefits were estimated as follows.

Step 1: Estimating the probability of survey response from ASQ members

As noted earlier, the ASQ agreed to a request from the management of the Baldrige National Quality Program Office at NIST on our behalf to distribute a survey administered by the Program Office to its 875 US members. Sixty-five organizations returned completed or partially completed survey instruments.8

Step 1 quantifies the probability that an ASQ member who received a survey would respond to the survey. Obviously, the average probability of response is 65 returned surveys out of 875 sent surveys, or a 7.43 response rate. However, for the statistical analysis, an estimated probability of response for each of the 875 members is needed as a control variable used in Step 2.9

The probability of a member responding to the survey is estimated using an industry effects model represented as

$$\text{Prob(response)} = F\,(\text{2-digit SIC industry variables})$$

dependent variable used to estimate Eq. (1) equals 1 if the member returned a completed or partially completed survey and 0 otherwise, and where the 2-digit Standard Industrial Classification (SIC) industry variable categories are as described in the note to Table VI.10 Equation (1) then posits that the probability of a member responding can be predicted on the basis of the industry in which that member produces. The probit results from Eq. (1) are in Table VI.

Table VI Probit results for probability of response to the survey (n=859).

Variable	Estimated coefficient
dnonmin	0.743
	(1.46)
dchempet	−0.008
	(−0.03)
dmcneqin	−0.076
	(−0.37)
dtrcomut	0.020
	(0.06)
dwholret	0.035
	(0.11)
dfire	−0.047
	(−0.12)
dserv	−0.586**
	(−2.02)
dbusser	0.350
	(1.49)
dhealth	0.795**
	(2.07)
dpubadm	−0.215

	(−0.75)
Intercept	−1.418*
	(−8.73)
Log likelihood	−220.297
Psuedo-R 2	0.043
hi 2 (10)	19.94**

Note: The 16 observations in the miscellaneous category (members who could not be assigned to a 2-digit SIC industry or who were assigned to miscellaneous manufacturing) were dropped because the miscellaneous category predicted non-response perfectly. $\mathrm{dnonmin} = 1$ for the agriculture, forestry, fisheries, minerals, and construction industries, and 0 otherwise; includes $\mathrm{SICs} < 20$. $\mathrm{dchempet} = 1$for chemicals, petroleum, and rubber, and miscellaneous plastics, and 0 otherwise; includes SICs 28, 29, and 30. $\mathrm{dmcneqin} = 1$ for machinery and equipment, both non-electric and electric and electronic, and instruments, and 0 otherwise; includes SICs 35, 36, 37, and 38. $\mathrm{dmats} = 1$ for the remaining manufacturing SICs, and 0 otherwise; includes SICs 20 through 27 and SICs 31 through 34; observations with $\mathrm{dmats} = 1$ are in the intercept. $\mathrm{dtrcomut} = 1$ for transportation, communications, and utilities, and 0 otherwise; includes all 2-digit SICs greater than 39 and less than 50. $\mathrm{dwholret} = 1$ for wholesaling and retailing, and 0 otherwise; includes all 2-digit SICs greater than 49 and less than 60. $\mathrm{dfire} = 1$ for finance, insurance, and real estate, and 0 otherwise; includes all 2-digit SICs greater than 59 and less than 70. $\mathrm{dserv} = 1$ for other services other than business services and health services, and 0 otherwise; includes all 2-digit SICs greater than 69 and less than 90 except for SIC 73 and 80. $\mathrm{dbusser} = 1$ for business services, and 0 otherwise; includes SIC 73. $\mathrm{dhealth} = 1$ for health services, and 0 otherwise; includes SIC 80. $\mathrm{dpubadm} = 1$ for public administration, and 0 otherwise; includes 2-digit SICs greater than 89 and less than 100. Asymptotic t-statistics in parentheses.

*Significant at 0.01 level, **significant at 0.05 level, and ***significant at 0.10 level.

For each of the 875 surveyed members, Eq. (1) produces a predicted value for the probit index, z, for the probability of response.11 12

Step 2: Estimating the probability of self-assessment for responding members

Step 2 quantifies the probability that an ASQ member who received a survey conducted a quality-based self-assessment. A probability of self-assessment is needed in the estimation of net benefits. First, a probability of self-assessment model is estimated, and secondly, a prediction of the probability of self-assessment for each ASQ member is calculated in Step 3.

The probability of a member having conducted a self-assessment in the past, given that the member returned a completed or partially-completed survey, is estimated using a model written as

$$\text{Prob(self-assessment)} = F(\text{2-digit SIC industry variables, competitiveness variables, control variables})$$

where the dependent variable used to estimate the model equals 1 if the member responded in the affirmative to at least one of the following survey statements, and 0 otherwise:

> Has your organization performed a self-assessment using the Baldrige Criteria for Performance Excellence or related criteria (and by related criteria we mean criteria informed or derived by the Baldrige Criteria)? If yes, in what year(s)?
>
> Has your organization applied for the Malcolm Baldrige National Quality Award? If yes, in what year(s)?
>
> Has your organization applied for a state quality award? If yes, in what year(s)?

and where the competitiveness variables noted in Eq. (2) are defined in terms of a member's Likert responses (7= strongly agree to; 1 = strongly disagree) with the following two survey statements:13

> 1. the possibility or threat of new competition is significant (comp);
>
> 2. our customers have a significant ability to bargain on the price of our primary products (barg);

and where the relevant control variables are based on estimates of the probability of response (probres) to the survey from Eq. (1). We introduce a control for the probability of response to the survey because our model of the probability of assessment is exploratory and unlikely to be complete with just the variables other than the probability of response. We do not believe the response variable is simply controlling for the effect of a correlation in random errors in the model of response and the complete model of the probability of self-assessment. Instead, we view the variable probres as capturing substantive effects of the complete model that otherwise would be left in the error term and that are related to the probability of responding to the survey. Thus, probres completes our substantive model, capturing systematic effects on the probability of self-assessment that vary with the characteristics of the ASQ members that are associated with their probability of response. Those ultimate characteristics may not be those in our response model, but associated with them and therefore with response.

Twenty-three of 65 members had performed a self-assessment.

The probit results from Eq. (2) are in Table VII.14 15

Table VII Probit results for probability of self-assessment (n=60).

Variable	Estimated coefficient
dwholret	0.899
	(1.33)
dpubadm	1.932*
	(2.46)
comp	−0.189
	(−1.36)
barg	0.234***
	(1.80)
probres	4.248
	(1.25)
Intercept	−1.276
	(−1.40)
Log likelihood	−32.096
Psuedo-R^2	0.124
hi^2 (5)	9.11***

Note: There are 65 observations available to estimate the model in Eq. (2); however, the 2-digit industry variables, *dtrcomut* and *dfire*, are dropped along with the five observations where they equal 1 because they predict assessment perfectly. Thus, the results mentioned earlier are based on 60 observations. Asymptotic *t*-statistics in parentheses. *Significant at

Table VII Probit results for probability of self-assessment (n=60).	
Variable	**Estimated coefficient**
0.01 level, **significant at 0.05 level, and ***significant at 0.10 level.	

Step 3: Predicting the probability of self-assessment for members of ASQ

The statistical output from this Step 3 is an estimate of the probability of conducting a self-assessment for each of the 875 members of ASQ using the results from Eq. (2) presented in Table VII.

With reference to Eq. (2), a probit index for each of the 875 members is estimated by multiplying the actual value of each independent variable for each member by the estimated probit coefficient reported in Table VII.16 17

Step 4: Estimating the net social value of the Baldrige National Quality Program to ASQ members

Of the 23 members of ASQ that performed a self-assessment, 14 responded to the following survey statement:

> In the absence of the Malcolm Baldrige National Quality Award—and therefore without the information and assistance that it provides about performance management/quality improvement assessments and therefore with the need to incur expenditures to develop and acquire such knowledge and assistance from other sources—what expenditures (fully burdened) would your organization have incurred to achieve the same level of expertise in performance management/quality improvement that you now have? $___per year over the previous___years.

As discussed earlier with reference to the counterfactual evaluation method, members' responses to this statement represent credible time-specific estimates of the benefits (i.e. the costs avoided reported in $2000) associated with the Baldrige National Quality Program. Thus, for each of the 14 responding members, a time series of real benefits received is formulated.

Regarding costs to compare to this time series of benefits, each of the 14 members responded to the following two questions:

> If your organization has been an award applicant, what was the total economic cost (fully burdened) to your organization to obtain, understand, collect relevant information, and

comply with the Baldrige Criteria or state application requirements? $___per year during the year(s)___.

and,

If your organization did not apply for the Malcolm Baldrige National Quality Award or state award, but nonetheless performed a self-assessment using the Baldrige Criteria or related criteria, what was the total economic cost (fully burdened) to your organization to perform the self-assessment? $___per year during the year(s)___.

Thus, for each of the 14 responding members, a time series of real ($ 2000) costs incurred to make the Baldrige Criteria operational is also developed.18

The net present value of each member's benefits is calculated using these survey data by first calculating the present value (referenced to the earlier of the first year of benefits or the first year of costs, hereafter the base year) of each member's benefits and each member's costs. The discount rate for this calculation is r = (k-0.03)/(1+0.03), where k is each member's reported hurdle rate and where the prevailing rate of price inflation over the reported time intervals is estimated at 3%.19 Thus, net present value is the difference between the present value of benefits less the present value of costs, both referenced to the base year. Each member's net present value of benefits is then re-referenced to 2000 using a 7% growth rate to account for the social opportunity costs of these moneys (OMB, 1992).

The following model is estimated using the 14 calculated net present values:20

$$\mathrm{NPV}_{2000} = F(\text{2-digit industry variables, size variables})$$

member size was provided by ASQ for 874 of the 875 members. The least-squares results from Eq. (3) are in Table VIII.21

Table VIII Least-squares results for net present value of benefits (n=14).

Variable	Estimated coefficient
size	−83844.49**
	(−2.48)
size2	13.33**
	(2.27)
dtrcomut	4.90e+07***
	(2.10)

Intercept	9.45e+07**
	(2.71)
F (3, 10)	3.51***
R^2	0.513

Note: The explanatory member-size variable is measured in millions of dollars, whereas the dependent variable for value is measured in dollars. t-statistics in parentheses. *Significant at 0.01 level, **significant at 0.05 level, and ***significant at 0.10 level.

The estimated coefficients in Eq. (3) are used to forecast the net present value of benefits for each of the 874 members of ASQ for which member size was available.

The predicted values from Eq. (2) represent point estimates for the probability of each member of ASQ conducting a self-assessment. The predicted values from Eq. (3) represent point estimates of the net present value of benefits associated with the Baldrige Program conditional on a member conducting a self-assessment. The product of these two estimates gives a point estimate of the expected net present value from the Baldrige Program for a member of ASQ. Using the standard errors of our predictions from Eqs. (2) and (3), we shall control to the extent possible for the relatively small sample of members that provided the detailed information about their net benefits from the Program.

Thus, in an effort to present conservative estimates of the net present value of benefits associated with the Baldrige Program to members of ASQ, the following adjustments are made.

First, regarding the predicted values of the probability of a self-assessment from Eq. (2), a 0.4142 confidence interval is calculated for each member of ASQ, and the lower-bound on that interval is used as the relevant predicted value of the probability of self-assessment for that member. The lower-bound on a 0.4142 confidence interval implies that there is a 0.7071 probability that the true value of the probability of self-assessment is greater than the value being used.22

Secondly, regarding the predicted value, conditional on self-assessment, of the net present value of benefits associated with the Baldrige Program from Eq. (3), a 0.4142 confidence interval is calculated for each ASQ member using the standard errors for the linear combination of the estimated coefficients and for the error in equation. The lower-bound on that interval is then used as the conservative net present value conditional on self-assessment by the member.

The product of the lower-bound of the probability of self-assessment from Eq. (2) and the lower-bound of the net present value of benefits from Eq. (3) yields for each member an estimate of net present value of benefits. That estimate may be lower or higher than the true value of the net present value of benefits. The true value has greater than a 50% probability 0.7071 x 0.7071 =

0.50 of being larger than the value being used as the estimate, because the probability that both estimates multiplied are exceeded by their true value is 0.50. Of course, in some cases where the true value of one but not the other of the two estimates being multiplied falls short of the lower-bound, the true value of net present value benefits may still exceed the estimate used. Hence, the true value has more than a 50% probability of being greater than the one used.

The sum of the conservative, lower-bound derived value of net benefits for ASQ members is $ 2.17 billion.23

Thus, if it is assumed that there is no value associated with the Baldrige National Quality Program other than that received by the ASQ members, the conservative present value for net private benefits is $ 2.17 billion. When compared with the present value of the total social cost associated with the Program of $ 119 million, the ratio of ASQ benefits to social costs is 18.2:1.

Step 5: Estimating the aggregate net social value of the Baldrige National Quality Program

If the entire economy benefits from the Baldrige National Quality Program to the same extent as the ASQ members,24 then total social benefits can be forecast using the following formula:

$$\text{Economy Value} = \frac{\text{Value for ASQ}}{\text{Proportion taken by the ASQ members in the 50 represented industrial sectors}}$$

where the denominator is calculated to be 0.0880285.25

Thus, under this assumption, the conservative present value of social benefits is $ 24.65 billion.26 When compared with the present value of the total social cost associated with the Program of $ 119 million, the ratio of economy-wide benefits to social costs is 207:1.27

4.4 Ratio of net social benefits to social costs associated with the Baldrige National Quality Program

As derived in the previous section, the conservative estimate of the present value of aggregate economy-wide net social benefits associated with the Program through 2000 is $ 24.65 billion (rounded in $ 2000). As also explained earlier, the present value of the social costs to operate the Program through 2000 is $ 119 million (rounded in $ 2000). From an evaluation perspective, these values yield a benefit-to-cost of 207:1.

5 Concluding remarks

This paper reports the findings from an economic evaluation of Baldrige National Quality Program. Extrapolating from the ASQ membership to the entire economy—under the assumption that the entire economy benefits from the Program to the same extent as ASQ members—implies a social benefit-to-cost ratio of 207:1. Yet the organizations outside of the ASQ may benefit even more than the ASQ membership. The ASQ members represent by member size 8.8% of the 50 industrial sectors with ASQ members. Through 2000, 11 of the 43 Baldrige Awards were received by ASQ members. Further, on the basis of requests for Baldrige application materials and criteria, as well as the many winners from outside the ASQ, many companies outside of the ASQ are using and benefiting—conceivably even more than ASQ members—from the Baldrige Criteria. Thus, extrapolating from the net social benefits of the Program for ASQ members to the economy as a whole may underestimate the true social benefits associated with the Program.28 In that case, the social benefit-to-cost ratio derived in this paper would understate, even beyond the conservative estimation procedure used in this study, the true benefits of the Program.

Certainly, the estimated benefit-to-cost ratio of 207:1 supports the hypothesis that the public's investment in quality-standards infrastructure is worthwhile. The Baldrige National Quality Program at NIST provides another NIST standards-infrastructure investment, although in contrast to NIST's investments in infratechnologies focused on engineering, measurement, and science, the Baldrige Program is focused on management. These public investments in management standards appear to be worthwhile when evaluated using the benefit-to-cost ratio. However, although a benefit-to-cost ratio greater than 1.0 logically implies that the Program is worthwhile given the standard assumptions behind evaluation analysis, the assumptions are important.29 As Scott (2000) explains, the benefit-to-cost ratio assumes that we really do know the opportunity costs of the resources invested by the public in a program. Although the social benefits greatly outweigh the costs as measured, there may be even higher yields on other potential uses of the public's funds. We have simply assumed, albeit following OMB (1992), that a yield of 7% covers the opportunity costs for the public's invested funds. Further, the value of resources has been measured by the preferences given the current distribution of income. Finally, the public may want to use its investments to promote goals, such as diversity, that are not measured by values of resources as determined by demands for goods and services and the costs of technologies available for providing them.

The economics and management literature cited in Table IV describes benefits from the Baldrige Program because it leverages private sector investments in quality management. Thus, the literature about the Program and the Award suggests that the large benefit-to-cost ratio for the Baldrige Program is realistic and not unexpected. Taken together, the literature about the Program and the benefit-to-cost ratio estimated in this paper support the hypothesis that NIST's National Quality Award Program is an efficient infrastructure investment in standards that are important for the effective operation of organizations.

Acknowledgements

We appreciate the comments and suggestions of Daniel Barton, Lee G. Branstetter, Charles C. Brown, Joseph E. Cooper, Susan Cozzens, Barry Diamondstone, Harry Hertz, Steven Martin, Troy J. Scott, Gregory Tassey, and two anonymous referees on aspects of earlier versions of this paper. Any remaining errors or omissions are our responsibility.

Notes

1 As Townsend and Gebhardt (1996) explained, the origins of the Baldrige Award grew from ‘alarm over the Japanese challenge to the American economy’ (p. 6), and they concluded (p. 13):

[T]he Baldrige will retain its position of importance, a position earned by being perhaps the major factor in positioning American business for the 21st century. The Baldrige didn’t just shift the paradigm for American business—it defined a whole new way to go about doing things. As a result, business communities throughout the world once again can look to America to learn how to get things done.

2 The origin of this view can be traced at least to Bush (1945); Link and Scott (2001) place it in a specific policy context.

3 For discussion and illustrations of the method, as well as an alternative method for use with publicly funded but privately performed R&D investments, see Link and Scott (1998, 2000, 2001).

4 In the extreme case where industry would not have made the investments at all, there are no private-sector costs avoided. However, because the private-sector performance shortfall is complete, the entire, traditional Griliches-Mansfield-like (whether their cost-reducing innovations or surplus-creating innovations more generally) stream of returns to the R&D investments is valued as benefits. In that special case, the Link–Scott approach is identical to the Griliches–Mansfield approach except that it has the advantage of having pointed out that government could do the work more efficiently—in this special case because industry would not do it at all; see Link and Scott (1998) for more details about the counterfactual evaluation method. Consistent with what our respondents have told us, and further to be conservative in our estimate of the benefits of the Baldrige Program, we assume throughout this paper that the private sector could—for the additional costs identified in our survey—have replicated the results of the Program.

5 Available online at: http://www.asq.org.

6 In addition to these US organizational members, there are over 200 international organizational members plus over 120,000 individual members.

7The Foundation reimburses between 60% and 70% of examiner travel costs, and the remainder is paid by the examiner’s company or organization.

8 ASQ sent an electronic reminder to each survey recipient 3 weeks after the initial mailing. No member-specific information is reported herein to ensure confidentiality.

9 As explained in the discussion of Step 2, we include a variable for response in the belief that it will capture substantive effects of the complete model of the probability of self-assessment.

10 ASQ provided the 2-digit industry for 75% of its members. Public domain information was used to determine the remaining classifications, including the Thomas Register and other Internet search mechanisms. The simple industry effects model is significant; more elaborate models that add other available characteristics of the members have no greater explanatory power—the additional variables are not statistically significant.

11 There are 16 cases that were assigned to a miscellaneous category because either a member could not be matched uniquely to a 2-digit SIC industry or was assigned to miscellaneous manufacturing. None of those 16 members responded. Consequently, the categorical variable for the group predicted non-response perfectly, and the 16 observations were dropped from the sample used to estimate the model and assigned a probability of response of 0.

12 On the basis of Eq. (1), the hazard rate is also computed as $h(z)=F'(z)/[1-F(z)]$, where $F(z)$ is the probability of response given the probit index z (hence, it is the cumulative density function for the standard normal variable at the value z) and $F'(z)$ is the density of the standard normal variable at z for each observation. The hazard rate is the conditional probability of response for a small increase in z. Conditional on no response for the observation, the probability of response for a small increment in z is $F'(z)d\,z/[1-F(z)]$. A 'non-selection' hazard rate used in traditional two-step methods to control for selection is defined analogously.

13 The mean value of comp (n=65)=5.6. The mean value of barg (n=65)=4.6. The inclusion of these competitiveness variables follows from the economic and management literatures related to quality shown in Table IV. Firms facing greater competitive pressures or buyers with greater bargaining strength are expected to be more likely to invest heavily in quality management; see for example Lau (1996) who develops information about his responding firms' competitive environments, including the possibility or threat of new competition.

14 When the hazard rate is included in Eq. (2) in place of the probability of response, the estimated probit model performed almost identically to the model reported in Table VII. Those results are available from the authors on request. Further, other available, potential explanatory variables were insignificant and did not add importantly to the model's explanatory power.

15 The model in Eq. (2) is estimated with 65 observations, however the 2-digit industry variables, dtrcomut and dfire, are dropped along with the five observations where they equal 1 because they predict assessment perfectly. Thus, the results in Table VII are based on 60 observations.

16 As noted with reference to the estimation of Eq. (2), data are available for 65 members on comp and barg. The mean value of these two variables (n=65) is imputed to the other 810 (875−65) ASQ members for predicting the probability of self-assessment.

17 The mean value of the probit index (n=810)=−0.7041409, corresponding to a probability of assessment=0.2602325. In the following calculations, a lower-bound probit index is used rather than the predicted value averaged here. Note from the foregoing footnote that there are 65 ASQ members that responded to the survey. Also there are by happenstance 65 of 875 members where dtrcomut and dfire equal 1, so there is no probit index for them from the estimation of Eq. (2)—recall from an earlier footnote that those two categories are perfect predictors of assessment—and hence n=810. In the following calculations, rather than imposing a probability of self-assessment of 1.0 on each of the additional 65 members in the perfect prediction categories, the average lower-bound probability of self-assessment from Eq. (2) is imputed to them; thus, producing in these instances, a more conservative estimate. The average lower-bound probability, as contrasted with the average probability, is explained subsequently.

18 Such costs are often referred to as pull costs; see Link and Scott (1998).

19 Regarding the hurdle rate, each member was asked to respond to the following statement: \begin{quote} What is your company's hurdle rate for investments (the minimum rate of return that your company must anticipate if it is to consider new investment worthwhile)? \noindent \underline{ } percent. \end{quote} The real rate of return will be r=(k−a)/(1+a), where a is the anticipated rate of inflation. If one invests X and receives Y, the nominal return for the period is k such that X(1+k)=Y and k=(Y−X)/X. Given an anticipated rate of inflation a, the real rate of return r is such that X(1+a)(1+r)=Y as that yields the rate of return r in constant dollars: X(1+r)=Y/(1+a). As X(1+a+r+ra)=X(1+k), then k=(a+r+ra) and r=(k−a)/(1+a). The mean value of k=0.1821.

20 The mean value of NPV2000 (n=14)=$ 17.7 million.

21 Other available, potential explanatory variables, including various hazard rates or associated probabilities and other sector effects, were insignificant and did not add importantly to the explanatory power of the model.

22 Each tail in a 0.4142 confidence interval contains 0.2929 of the distribution, so there is 0.7071 probability (0.4142+0.2929) that the true value is greater than the value being used.

23 The mean value of the conservative estimate of value (n=874)=$2,478,039.

24 This extrapolation is similar in procedure to that used by Scherer (1982).

25 The size data for industrial sectors were assembled using information in US Census (1997) and Council of Economic Advisers (2001). Size data for 1997 were inflated using the chain-type price index for gross domestic product from Table B-7, 'Chain-type price indexes for gross domestic product, 1959–2000' Council of Economic Advisers (2001, p. 284) to be comparable with the ASQ 1999 sales data. When 1997 sector size data were unavailable, 1992 data were used and then inflated to 1999.

26 $ 2.17 billion/0.088025= $24.65 billion.

27 All but a few ASQ members could be separated into the manufacturing sector and the service sector. Recalculating, using only these two broad industrial categories and omitting industrial categories where there are very few members (SIC<20 sectors with only eight ASQ members), yields a conservative estimate of the aggregate manufacturing sector's net benefits of $7.6 billion and a conservative estimate of the aggregate service sector's net benefits of $13.0 billion. Thus, when the sum of these estimates is compared with total social costs of $119 million, the resulting benefit to cost ratio is 173:1.

28 There is, on the other hand, a possible upward bias from extrapolation because the ASQ population has a proportionally greater number of Baldrige Award winners than does industry as a whole.

29 Of course, a benefit-to-cost ratio greater than 1.0 does not by itself justify the use of public money to support this program. A discussion of elements of market failure, which the Baldrige Program seeks to overcome, is in Link and Scott (2005).

References

1. Banker, R. D., Khosla, I. and Sinha, K. K. 1998. Quality and Competition. Mangement Science, 44(9): 1179–1192.

2. Black, S. E. and Lynch, L. M. 1996a. Human Capital Investments and Productivity. American Economic Review, : 263–267.

3. Black, S. E. and Lynch, L. M. 1996b. "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity". Harvard University and US Department of Labor

4. Blodgett, N. 1999. Service Organizations Increasingly Adopt Baldrige Model. Quality Progress, : 74–76.

5. Bush, V. 1945. Science—the Endless Frontier, Washington, DC: US Government Printing Office.

6. Council on Competitiveness. 1995. Building on Baldrige: American Quality for the 21st Century, Washington, DC: Council on Competitiveness.

7. Council of Economic Advisers. 2001. Economic Report of the President, Washington, DC: US Government Printing Office. January 2001

8. Easton, G. S. and Jarrell, S. L. 1998. The Effects of Total Quality Mangement on Corporate Performance: an Empirical Investigation. The Journal of Business, 71(2): 253–307.

9. George, S. and Weimerskirch, A. 1994. Total Quality Management: Strategies and Techniques Proven at Today's Most Successful Companies, New York: John Wiley & Sons, Inc.

10. Griliches, Z. 1958. Research Costs and Social Returns: Hybrid Corn and Related Innovations. Journal of Political Economy, 66: 419–431.

11. Helton, B. R. 1995. The Baldie Play. Quality Progress, 28(2): 43–45.

12. Hendricks, K. B. and Singhal, V. R. 2001. The Long-Run Stock Price Performance of Firms with Effective TQM Programs. Management Science, 47(3): 359–368.

13. Huselid, M. A. and Becker, B. E. 1996. High Performance Work Systems and Firm Performance: Cross-Sectional versus Panel Results. Industrial Relations, 35(3): 400–423.

14. Ichniowski, C., Shaw, K. and Prennushi, G. 1995. "The Effects of Human Resource Management Practices on Productivity". NBER Working Paper No. 5333

15. Lau, R. S.M. 1996. A Survey of Competitiveness and Manufacturing Practices. South Dakota Business Review, 54(1): 4–9.

16. Lawler, E. E., III, Mohrman, S. A. and Ledford, G. E. Jr. 1998. Creating High Performance Organizations: Practices and Results of Employee Involvement and Total Quality Management in Fortune 1000 Companies, San Francisco: Jossey-Bass.

17. Link, A. N. and Scott, J. T. 1998. Public Accountability: Evaluating Technology-Based Institutions, Norwell, Mass: Kluwer Academic Publishers.

18. Link, A. N. and Scott, J. T. 2000. "Estimates of the Social Returns to Small Business Innovation Research Projects". In The Small Business Innovation Research Program, Edited by: Charles, W. Wessner. 275–290. Washington, DC: National Academy Press.

19. Link, A. N. and Scott, J. T. 2001. Public/Private Partnerships: Stimulating Competition in a Dynamic Market. International Journal of Industrial Organization, 19: 763–794.

20. Link, A. N. and Scott, J. T. 2005. Evaluating Public Research Institutions: The U.S. Advanced Technology Program's Intramural Research Initiative, London: Routledge.

21. Mansfield, E., Rapoport, J., Romeo, A., Wagner, S. and Beardsley, G. 1977. Social and Private Rates of Return from Industrial Innovations. Quarterly Journal of Economics, 91: 221–240.

22. NIST. 1996. "Study Finds 'Quality Stocks' Yield Big Payoff". In Update February 5, Available online at http://www.quality.nist.gov/qualstok.htm

23. NIST. 1997a. "Results of Baldrige Winners' Common Stock Comparison, Third NIST Stock Investment Study". Available at http://www.quality.nist.gov/96stok.htm

24. NIST. 1997b. “Malcolm Baldrige National Quality Award: 1997 Criteria for Performance Excellence. Gaithersburg, Maryland: National Institute of Standards and Technology”. Available online at http://www.quality.nist.gov/

25. NIST. 1998. Ten Years of Business Excellence for America, Gaithersburg, MD: National Institute of Standards and Technology.

26. NIST. 2000. “‘Baldrige Index’ Outperforms S&P 500 by Almost 5 to 1. Commerce News”. February 25, Available at http://www.nist.gov/public_affairs/releases/g00-26.htm

27. Office of Management and Budget (OMB). 1992. “Circular No. A-94”. In Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs, Washington, DC: Office of Management and Budget.

28. Scherer, F. M. 1982. Interindustry Technology Flows in the United States. Research Policy, 11: 227–245.

29. Scott, J. T. 1999. The Service Sector’s Acquisition and Development of Information Technology: Infrastructure and Productivity. Journal of Technology Transfer, 24: 37–54.

30. Scott, J. T. 2000. The Directions for Technological Change: Alternative Economic Majorities and Opportunity Costs. Review of Industrial Organization, 17: 1–16.

31. Townsend, P. and Gebhardt, J. 1996. The Importance of the Baldrige to the US Economy. Journal for Quality and Participation, : 6–13.

32. US General Accounting Office (GAO). 1991. Management Practices: U.S. Companies Improve Performance Through Quality Efforts Washington, DC: General Accounting Office GAO/NSIAD-91-190

33. US Census Bureau. 1997. Economic Census, 1987 SIC Basis, Sales, Receipts or Shipments in 1997 Available online at http://www.census.gov/epcd/ec97sic/E97SUS.HTM along with related material at http://www.census.gov/epcd/www/sic.html

34. Wisner, J. D. and Eakins, S. G. 1994. A Performance Assessment of the U.S. Baldrige Quality Award Winners. International Journal of Quality and Reliability Management, 11(2): 8–25.