Planning Report 01-3

Economic Evaluation of the Baldrige National Quality Program

Prepared by:
Albert N. Link
and
John T. Scott

for

National Institute of Standards & Technology

Program Office
Strategic Planning and
Economic Analysis Group

October 2001

U.S Department of Commerce
Technology Administration

Economic Evaluation of the

Baldrige National Quality Program

by

Albert N. Link
Department of Economics
University of North Carolina at Greensboro

and

John T. Scott
Department of Economics
Dartmouth College

Final Report Submitted to the National Institute of Standards and Technology
Program Office

October 2001

This report has benefited from the comments and suggestions of a number of individuals including Daniel Barton, Lee G. Branstetter, Charles C. Brown, Joseph E. Cooper, Barry Diamondstone, Harry Hertz, Steven Martin, Troy J. Scott, and Gregory Tassey.

Table of Contents

Executive Summary

This study is an economic evaluation of the Baldrige National Quality Program. Specifically, the analysis assesses:

- the net private benefits associated with the Baldrige National Quality Program to the U.S. private and public sector organizational members of the American Society for Quality,
- via generalization, the net social benefits of the Program in the aggregate, and
- the relationship between economy-wide net benefits traceable and the social costs associated with operating the Program.

Based on information collected from a mail survey of the U.S. organizational members of the American Society for Quality (ASQ), the conservative estimate of the present value (in constant 2000 dollars) of the net private benefits associated with the Baldrige National Quality Program is $2.17 billion.

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

Based on information provided by the Baldrige National Quality Program, the present value (in constant 2000 dollars) of social costs associated with the Program to date is $119 million.

Therefore, from an evaluation perspective for the economy as a whole, the benefit-to-cost ratio characterizing the Baldrige National Quality Program is conservatively 207-to-1.

I

Introduction and Overview of the Baldrige National Quality Program

A. Purpose of the Study

The National Institute of Standards and Technology (NIST) regularly conducts assessments of the economic impacts of its programs in order to provide important programmatic and decision-relevant information to NIST managers, Executive Branch officials, and to the Congress. The results of such studies are also of interest to U.S. industry.[1]

The purpose of this study, as summarized in this report, is an economic evaluation of the Baldrige National Quality Program (the Program).

More specifically, this study considers:

- the net private benefits associated with the Baldrige National Quality Program to the U.S. private and public sector organizational members of the American Society for Quality (ASQ),
- via generalization, the social benefits of the Program in the aggregate, and
- the relationship between economy-wide benefits traceable and the social costs associated with operating the Program.

These are important considerations because the findings will provide valuable information to NIST about the economic impacts of its Baldrige National Quality Program, and because it will provide accountability information to the U.S. Congress (under the umbrella of the Government Performance and Results Act of 1993) about the efficiency with which the Program utilizes public moneys.

B. History of the Baldrige National Quality Program

Productivity in the non-farm U.S. economy fell in the early-1970s and then fell again in the early to mid-1980s.[2] Associated with these declines was the loss of world market shares by firms in many critical industries. In response, a number of economic policy initiatives were introduced in the early-1980s in an effort to reverse the downward productivity trend by stimulating innovative activities within firms. These initiatives included the Stevenson-Wydler Innovation Act of 1980 to encourage technology transfer from federal laboratories to the private sector, the Economic Recovery and Tax Act of 1981

[1] A review of previous studies is in Tassey (1999).

[2] Most economists measure this decline in terms of total factor productivity.

that contained provisions for a R&E tax credit, and the National Cooperative Research Act of 1984 to encourage collaborative research activity among firms.

Further, Congress declared as part of the Malcolm Baldrige National Quality Improvement Act of 1987 (P.L. 100-107) that:[3]

> … the 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 …

C. Outline of the Report

The remainder of this report is outlined as follows. In an effort to place the Malcolm Baldrige National Quality Improvement Act of 1987 and the Malcolm Baldrige National Quality Award in a performance management/quality improvement perspective, alternative concepts of quality are described in Section II.

In Section III, a conceptual model of investments in quality by the firm is presented, and the associated literature regarding such investments is reviewed in the context of the conceptual model.

---

[3] As Townsend and Gebhardt (1996) explain, the origins of the Baldrige Award grew from "alarm over the Japanese challenge to the American economy" (p. 6), and they conclude (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.

In Section IV, the key empirical hypothesis about the benefits of the Program is stated.

In Section V, the history of applications to the Program is overviewed.

In Section VI, an economic evaluation of the Baldrige National Quality Program and the related Baldrige Criteria for Performance Excellence (Baldrige Criteria) is offered.

Finally, the report concludes in Section VII with summary remarks.

II

Concepts of Quality

A. Evolution of Ideas about Quality

Although a long history of managerial interest in the concept of quality exists; the systematic and comprehensive view of quality, in the sense of the Baldrige National Quality Program, is of recent origin. Juran (1995) provides one historical treatment of concepts related to the management of quality in a number of countries—reaching far back into history with chapters on quality management in ancient China, ancient Israel, ancient Greece, and ancient Rome, as well as other early civilizations, yet documenting quality management in more recent times including in twentieth-century Japan and the United States. Juran (1995, p. 630) concludes from his historical overview that new forces emerged in the twentieth century and that these forces have required a revolution in quality management. These new forces are:

- Greater complexity and precision of products
- Threats to human safety and health, and to the environment
- Government regulation of quality
- The rise of the consumerism movement
- Intensified international competition in quality

Despite a history of quality management reaching back into antiquity, in the twentieth century (Juran, 1995, p. 630):

> Technological measures of quality did exist on the shop floors, but managerial measures of quality did not exist in the boardrooms. So, except in Japan, the needed quality revolution did not start until very late in the twentieth century. To make that revolution effective throughout the world economies will require many decades—the entire twenty-first century. So, while the twentieth century has been the Century of Productivity, the twenty-first century will be known as the Century of Quality.

Reimann and Hertz (1993) capture the shift from narrow to comprehensive views of quality and the necessary managerial efforts to ensure it. They explain that the globalization of markets and associated international competition provided the impetus for the shift toward comprehensive quality management strategies that integrate company-wide pursuit of operational excellence rather than relegating quality issues to traditional forms of product inspection. Reimann and Hertz (1993, p. 43) observe:

> The United States and much of the world are in the early stages of a major transition in how work and quality are managed—moving from quality as a narrowly-defined,

> separate function managed at lower organizational levels [e.g., on shop floors], to quality integrated within all work units and with overall business management [e.g., in boardrooms]. The success of this transition depends on numerous factors not well addressed within traditional quality: executive-level leadership and involvement, empowerment, rapid deployment of changing requirements, management of innovation, management of diversity, and customer-focused planning and operations.

Table II-1 summarizes the information in Juran (1995) in a manner that clearly illustrates a descriptive time line for quality management.

The Council on Competitiveness (1995, p. 3) asked, "What is Quality?" as part of their report on the Baldrige National Quality Award:

> Today quality is best understood as principles and methods to improve the performance of organizations in achieving their objectives. From the broadest perspectives, we can think of quality management as a system to facilitate organizational [adaptation] to a changing environment. *While there are no fixed definitions of quality in the United States* [authors' emphasis], several factors are widely accepted as central to a successful quality program:
>
> 1) Customer satisfaction
> 2) Executive-level leadership
> 3) Employee involvement.

B. Formalizing Ideas about Quality

The evolution of ideas about the *sources* of quality mirrors the marked change in quality management that occurred in boardrooms during the last two decades of the twentieth century.

In the economics literature, quality is viewed in one of two general ways. Quality is viewed as a dimension of an industry's product. As such, when this dimension changes (e.g., the product improves through focused investment activity), the industry demand for the product increases. The simplest of economic analyses conceptualizes the results of investments in quality in terms of an increase in an industry demand schedule. The more complex analyses, which fall under the rubric of the "quality ladder" literature, envision industry demand increases over time in response to continued investments in quality, and such sequential shifts are, for mathematical convenience, indexed by numerical intervals (hence the concept of a ladder). Alternatively, quality is viewed as an indirect activity that affects industry productivity. With either view, the economics literature emphasizes sources of investments that lead to quality.

B.1. Traditional sources of quality improvement

The more traditional sources of quality improvement for private firms, with quality measured abstractly as a dimension of the product, are investments in research and development (R&D) or investments in product differentiation. Product differentiation can be achieved through modification of the physical characteristics of the product (with its locational attributes in both the product space and geographic space), or through advertising and other promotional efforts that can develop customers' awareness of the product's characteristics and thereby affect its image and customers' appreciation of the product. When quality is measured indirectly with productivity indices, many sources of the productivity are identified, including capital and R&D investments, but also investments in workplace practices, human capital, and information technology (Black and Lynch 1996). These latter explanations of productivity of course mirror the new comprehensive view of quality and its management.

In the management literature, quality is typically less abstract than in the economics literature in the sense that it is related to attributes of a firm or its products that are more activity-specific compared to more abstract or aggregate impact measures such as a single shift in an industry demand schedule or "quality ladder" interval shift, or even a productivity index such as output per worker. Thus, in the quality management literature, quality typically encompasses detailed operational performance characteristics such as the prompt delivery of a commodity or service to the firm's customers. For example, Link, Quick, and Tassey (1991, p. 473) find that the firms typically agree that quality is defined by "performance levels, performance stability, reliability and longevity." Simply put, quality is "conformance and/or fitness for use" and, in the quality management literature, the measures of performance and fitness are non-conceptual operational business indicators.

While the economics literature typically links quality to R&D investment, the management literature typically focuses not only on the amount of the R&D investment, but also on the manner in which the R&D investment is implemented. The economics literature typically assumes that the method of implementation is the optimal one and then proceeds to work out the implications of the investment, such as R&D, for economic performance. In contrast, the quality management literature typically works to understand what in fact would be the optimal method of implementation for R&D investment.

B.2. Broad-based sources of quality improvement

Further, the quality management literature frequently looks at a much broader set of quality investments than does the economics literature. Thus, while the economics literature focuses on R&D investments, the quality management literature considers the investments in quality within the several major budget categories of the firm. For example, Link, Quick, and Tassey (1991) discover from their survey of firms in the optical fiber industry the relative importance of various functional objectives of quality assurance programs. U.S. firms dominated the optical fiber industry in the early 1990s, and they allocated substantially more—three to four times as much, as a percentage, of their budgets—of their

operations, capital, R&D, and overhead budgets for quality assurance than did their foreign competitors, with the proportions of these budgets allocated for quality ranging from somewhat less than a fifth to somewhat more than a fourth. The functional objectives, in order of their average percentage of the quality budget for the successful U.S. firms, are (i) improving manufacturability and (ii) improving product performance, with each of these leading objectives taking somewhat under a third of the average quality budget, (iii) increasing product reliability and (iv) reducing attribute variability, with each of these objectives taking somewhat less than a sixth of the average quality budget. No other objective received more than a twentieth of the average quality budget. In all, it appears (Link, Quick, and Tassey 1991, pp. 473-474) that the successful U.S. firms are "pursuing a broader quality assurance strategy … of 'building quality in' through both product design and control of the production process (i.e. through a concurrent engineering process) as well as the more traditional post-production testing."

Reimann and Hertz (1993, p. 46), in the spirit of the quality management literature, emphasize that quality as captured in the criteria for judging the Malcolm Baldrige National Quality Award encompasses "continuous improvement in overall operations." Their view that the Baldrige Criteria capture and define "quality" is supported by Juran (1995, pp. 649-650), who states that the Baldrige Criteria define quality and the process of total quality management:

> By the 1980s it was becoming clear to upper managers that quality leadership could not be achieved by pecking away—by bringing in this or that tool or technique. Instead, it was necessary to apply the entire array of quality know-how (the "quality disciplines") throughout the entire company—to all functions and all levels—and to do so in a coordinated way. One shorthand expression for this comprehensive approach is the term *total quality management*, or TQM. (The usual Japanese term is company-wide quality control.) … At the outset there was no agreed standard definition for TQM, so communication became confused—among company departments, in their training courses, and in the general literature. This confusion has since been reduced by the publication of the criteria used by the American National Institute for Standards and Technology (NIST) to judge the applications for the United States' Malcolm Baldrige National Quality Award (Baldrige Award) … Those criteria have been widely disseminated—NIST has filled over a million requests for application forms.[4] While there have been relatively few applications for the award, many companies have conducted self-audits against the criteria. In addition, as national quality awards have proliferated, many have used the Baldrige Award criteria as inputs to their own list of criteria. By the early 1990s, this wide exposure had made the Baldrige Award criteria the most widely accepted definition of what is included in TQM.

Hertz (2000) distinguishes clearly the difference between the comprehensive scope of the Baldrige Criteria and the narrow focus of quality management tools such as "six sigma" or the ISO 9000 quality standard. Hertz observes that adoption of such management tools certainly does not mean that a company would thereby easily win the Baldrige Award. He observes (Hertz, 2000, p. 2):

[4] Updating the original quotation, note that by 2001, NIST has filled over two million requests for application forms.

> Unlike the requirements for six sigma or ISO 9000, organizations that use the Baldrige performance excellence criteria get a true systems perspective to their overall organizational performance.

Table II-2 is an overview of the criteria of the Malcolm Baldrige National Quality Award, as published by the Baldrige National Quality Program (NIST 2001).

Ray Stata (then chairman and CEO of Analog Devices), who served on the Malcolm Baldrige National Quality Award Board of Overseers, encapsulates the definition and theory of quality management succinctly (Council on Competitiveness, 1996, p. 1):

> By my definition, quality management is a company-wide system to accelerate the rate of learning and improvement in all aspects of a company's performance. So far, quality management has focused on customer satisfaction, continuous improvement, management by fact and management by process. Management by fact is just that: fact versus opinion. It's dealing with statistics and data as opposed to winging it. Management by process is where every person, every employee, has a customer and deliverables. It's the process by which they work to make those deliveries. The challenge is to understand and continuously improve these processes.

At the start of the new century, performance excellence and performance management encompass continuous improvement of all processes as contrasted with prioritized improvement efforts and response to changing business needs. Rather than focusing mainly on processes, as is the case with earlier views of total quality management, Baldrige performance excellence has a systems perspective about results and processes.

Table II-1
The Evolution of the Concept of Quality

Time Period	Concepts of Quality
Late 19th century	The Taylor System: "scientific management" increasing production without increasing skilled craftsmen by separating the planning of production (by engineers) from the execution by supervisors and workers. Juran (1973), Juran (1995, p. 555).
Early 20th century	Independent Inspection Departments: The Taylor System damaged human relations and had a negative impact on craftsmanship and quality. Central inspection departments were created to restore balance. Materials and goods were sampled in process with the results determining whether or not a lot of goods would be used. Finished goods were inspected in detail. Quality came to be seen as the responsibility of the inspection department. Juran (1995, p. 555-556).
Mid-1920s	Early Statistical Quality Control (SQC): Sampling inspection was grounded in probability theory. Juran (1995, pp. 556-557), AT&T (1989).
1940s and 1950s	Second wave of SQC and ASQC: Production needs and delivery deadlines required during World War II brought new interest in SQC. Eventually the American Society for Quality Control was created. New impetus for SQC resulted in quality control engineering and quality control departments to supervise the inspection department. Eventually functions of inspection, testing, quality control and reliability engineering were housed in the "quality department" headed by the "quality manager" usually reporting to the vice president for manufacturing. Juran emphasizes the deficiencies of the system in which quality was the top priority of just the quality department rather than the entire organization. Juran (1995, pp. 558-562), Working (1945), Grant (1953), AT&T (1989), Grant (1991), Juran (1991), Wareham and Stratton (1991).
1960s and 1970s	The big forces for change in the concept of quality Juran (1995, p. 630): "Greater complexity and precision of products," "Threats to human safety and health, and to the environment." "Government regulation of quality." "The rise of the consumerism movement." "Intensified international competition in quality." In response to the quality crisis brought on by the forces for change, piecemeal strategies emerged, including (Juran, 1995, pp. 583): "Exhortation of the workforce." "Organization and training of quality circles." "Statistical process control." "Awareness training for managers and supervisors." "Computation of the cost of poor quality." "Project-by-project quality improvement." "Preparation of complete manuals of procedure." "Revision of organization structure." "Incentives for quality." "Automation inspection and test." "Automation and robotics." Juran (1995, pp. 562-581, 630-634.)
1980s	In the face of a major quality crisis, U. S. firms focused primarily on three strategies: "exhortation, project-by-project quality improvement, and statistical process control." Juran, (1995, pp. 584-586.)
1990s	Poor strategies and poor execution of valid ones caused largely disappointing results for most quality initiatives in the 1980s, but some firms, including the winners of the NQA, "attained quality leadership … and thereby became the role models for the rest of the American economy." Juran (1995, p. 586). The core list of strategies, embodying the lessons learned in the 1980s about what worked and what did not, for successful pursuit of quality are captured by the NQA criteria described above. Juran (1995, pp. 649-650), Reimann and Hertz (1993, p. 46), George (1992), George and Weimerskirch (1994). The NQA criteria define "a model of integration" that demonstrate "how all of a company's processes and people can be focused on meeting customer requirements and improving operating performance." George and Weimerskirch (1994, p. v.)

Source: See Juran (1995, pp. 553-655).

Table II-2
Overview of the Baldrige Criteria for Performance Excellence

The Baldrige Criteria for Performance Excellence provide organizations with an integrated, results-oriented framework for implementing and assessing processes for managing all operations. The Baldrige Criteria are also the basis for making the Baldrige Award and providing feedback to applicants. The Baldrige Criteria consist of seven categories:

- Leadership: The company's leadership system, values, expectations, and public responsibilities.
- Strategic Planning: The effectiveness of strategic and business planning and deployment of plans, with a strong focus on customer and operational performance requirements.
- Customer and Market Focus: How the company determines customer and market requirements and expectations, enhances relationships with customers, and determines their satisfaction.
- Information and Analysis: The effectiveness of information collection and analysis to support customer-driven performance excellence and marketplace success.
- Human Resource Focus: The success of efforts to realize the full potential of the work force to create a high performance organization.
- Process Management: The effectiveness of systems and processes for assuring the quality of products and services.
- Business Results: Performance results, trends, and comparison to competitors in key business areas—customer satisfaction, financial and marketplace, human resources, suppliers and partners, and operations.

Source: NIST (2001).

III

Academic and Professional Literature Related to Investments in Performance Excellence

A. Fragmented Nature of the Literature

The academic and professional literature related to corporate investments in quality, that is sources of quality, is highly fragmented. In other words, authors tend to examine only selected aspects of the concept of quality. While the characterization of the literature as highly fragmented is a criticism in the best sense of the word, such fragmentation is understandable. From Table II-1 it is clear that the holistic view of quality and quality management as an all-encompassing total company effort to improve performance is a *new* view, and possibly a twenty-first century view, for which a consensus is still emerging.

There are a number of ways to summarize the literature on the broadly-defined subject of quality, and each could be useful in its own right. One could segment the literature by discipline, meaning that one would review the relevant economics literature, production literature, organizational literature, management literature, and so on. That approach has the advantage of grouping scholarship by the questions explored, because the choice of questions to be investigated is for the most part discipline specific. However, such a discipline-segmentation approach has the disadvantage of not unveiling areas of possible overlap in the questions being asked by scholars and not placing the subsequently identified questions into a general framework of analysis. One could alternatively segment the literature by general methodology, meaning that one could review the case-based literature, the survey-based literature, the theoretical literature, and so on. That approach may unveil some areas of possible overlap in the questions being asked, but it too falls short in the sense that the areas of overlap that are revealed might be neither systematic nor complete. Rather, the areas of overlap revealed could be the result of the pre-selected methodological groupings. Neither approach is effective for the purposes of this economic evaluation.

A careful review of the academic and professional literature reveals that there is not a general framework for considering investments in quality or how investments in quality affect firm performance; scholars have identified important elements that might be included in a general framework, but to date have failed to provide such a framework. Therefore, a conceptual model of firm performance is set forth in this section, and it will form the basis of a taxonomy for characterizing the extant literature.

B. Conceptual Model of Firm Performance

The conceptual model of firm performance is depicted in Figure III-1. The discussion below emphasizes specific aspects of the model that are related to overall firm performance. These specific aspects are illustrated as labeled boxes and their causal flows are illustrated by directional arrows.

Shown in Figure III-1 is one firm within a supply chain. The represented firm, simply put, allocates resources toward the production of a good or service. Once produced, the good or service competes in the market and generates value added. Fundamental to this production process is the use of purchased capital and labor, enhanced by the internal (i.e., self-financed and self-generated) innovation process of the firm. The good or service produced by the firm can be purchased by another firm as an input in its production process or it can be purchased by a final consumer.

More specifically, through strategic planning, the firm depicted in Figure III-1 decides what good or service to produce and how to produce it. The production decision is complicated; the firm has at its disposal at any given time a market from which to purchase current technology capital and current technology labor. The term "current technology" is used to emphasize that at a given time the market offers capital and labor of given vintage and of given technical capabilities. These inputs become more productive through the internal innovative activities of the firm. Academics generally think of an innovation as a process of applying new technical knowledge. New technology originates from a number of sources. These alternative sources for new technology are imbedded in the box simply labeled "Technology." Such sources include primarily, at least for the purposes of this review of academic models, proprietary research and development (R&D) activity.[5] The Baldrige Criteria define innovation more broadly: "Innovation means making meaningful change to improve an organization's products, services, and processes …(NIST 2001, p. 3).

To illustrate the dynamics of this conceptual model, consider the following hypothetical production issue. The firm, that has purchased other firms' capital equipment (i.e., other firms' product technology capital) and that employs a heterogeneous mix of labor, decides to produce a widget. The choice of producing a widget, as opposed to something else, is based on management's perception of market need. The widget will have certain performance attributes determined by (i) the current technology embodied in the capital equipment; (ii) the ability of the labor force; and (iii) the managerial expertise of the firm to efficiently coordinate capital and labor, introduce the product into the market place, and gain market acceptance. The firm is engaged in process R&D to enhance, via innovation, the productivity of the capital equipment that it purchased; it is engaged in product R&D to enhance the performance attributes of the widget.

[5] Other sources, not considered specifically for this study, are generic technology that originates from, for example, collaborative research arrangements; infratechnology that originates from, for example, the research agenda of the national laboratory system; and the science base that is fed through the basic research in, for example, universities. A more complete model is in Tassey (1997). Therein, Tassey refers to these other sources as "technical infrastructure."

Each of the eight boxes in Figure III-1 represents a purposeful aspect of the process of producing a good or service; hence each represents one aspect of firm performance. To anticipate the conclusions of this literature review, previous scholarly efforts have not yet asked, much less answered, all of the relevant questions related to the broadly-defined concept of performance excellence that is posited in Section II and that is captured in the quality attributes specified in the Malcolm Baldrige National Quality Award. Previous scholarly efforts have focused on only selected aspects of firm performance (specific boxes in Figure III-1). Previous inquiries have overlooked critical interactions within the firm that are fundamental not only to firm performance but also to a complete understanding of investments in performance excellence as a strategic decision.

C. Investments in Quality and Firm Performance

The literature relating investments in quality to firm performance addresses various aspects (i.e., specific activities or boxes) of the conceptual model in Figure III-1. Some analyses focus on the strategic planning that chooses the product or service to be produced. Others focus on the efficient use of current technology capital and current technology labor in production. Some studies analyze the innovation process or alternatively the technology base from which innovation is derived. Marketing and the penetration of markets are the focus of other studies, while still others analyze the management practices that enhance value added.

Figure III-2 captures a broad sweep of the literature by characterizing it in two dimensions. On the vertical axis is quality derived from an investment in the organizational culture of the firm, broadly defined to include the managerial tools as well as the attitudes toward numerous aspects, from key personnel elements to financial elements, to managing the firm's resources. Stated alternatively, the vertical axis characterizes the quality management literature as summarized in Section II. On the horizontal axis is quality resulting from an investment in product attributes, broadly defined to include the multitude of ways product attributes can be influenced, such as the R&D to develop new product characteristics, the production processes for products and services produced, the inspection of products to ensure desired characteristics, and the marketing and servicing of the products. Stated alternatively, the horizontal axis characterizes the economics literature as summarized in Section II.

The Baldrige Award, and the underlying Baldrige Criteria, emphasize the quality leaders and role models whose investments in both organizational culture and product attributes place them in the upper right hand quadrant of the figure.

As previously emphasized the academic and professional literatures typically focus on selected elements of the conceptual model in Figure III-1. However, because scholars generally do consider more than a single facet of this conceptual model of firm performance, it is not possible to uniquely map each study in the literature to a single element in Figure III-1. In fact, it is difficult to segment the more recent and relatively more complete analyses uniquely into the broad categorizations of organizational culture or product attributes illustrated in Figure III-2. However, a useful way to organize this literature review is

to describe alternative conceptual, theoretical models of investments in quality and then proceed to the empirical studies of such investments. Studies in the literature fall broadly into those that focus primarily on organizational culture and those that focus primarily on product attributes. As such, the literature divides into four areas:

- theoretical models of quality investment that emphasize investments in organizational culture
- theoretical models of quality investment that emphasize investments in product attributes
- empirical models of quality investment that focus on investments in organizational culture
- empirical models of quality investment that focus on investments in product attributes

A detailed review of the literature in each of these areas is in the Appendix to this report.

Figure III-1
Conceptual Model of Firm Performance

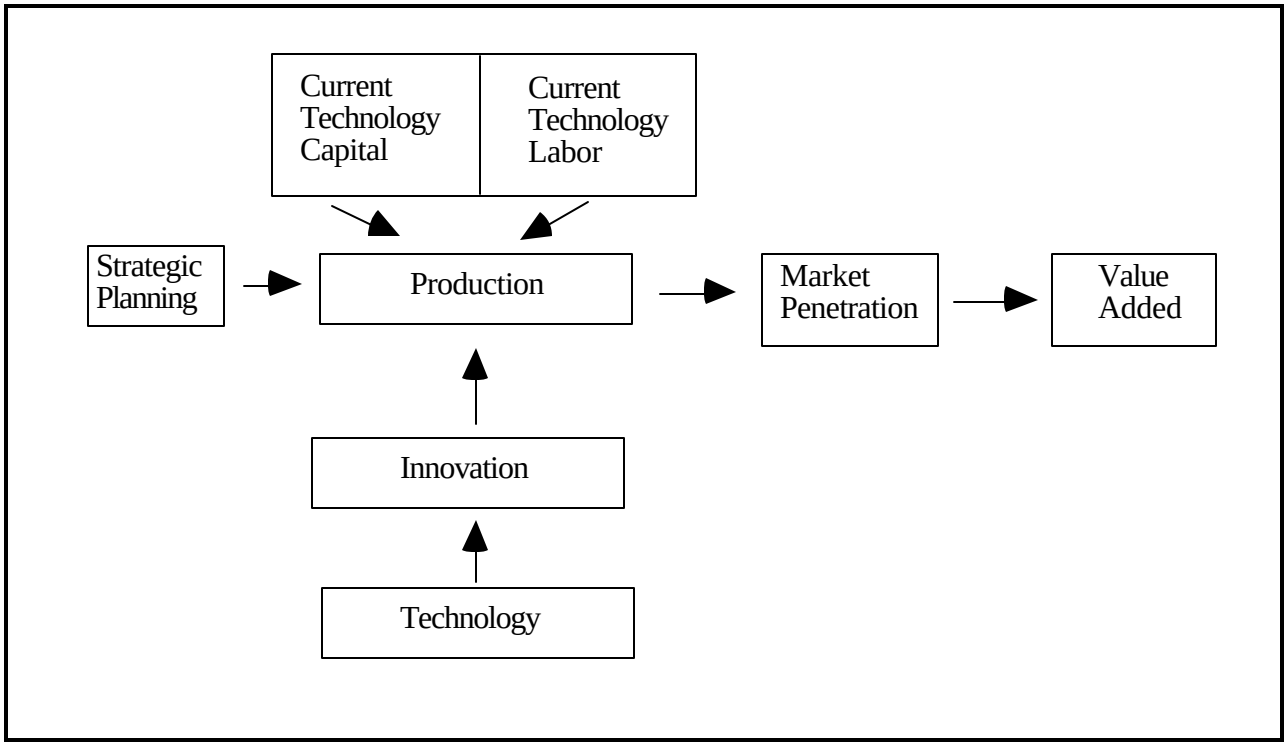
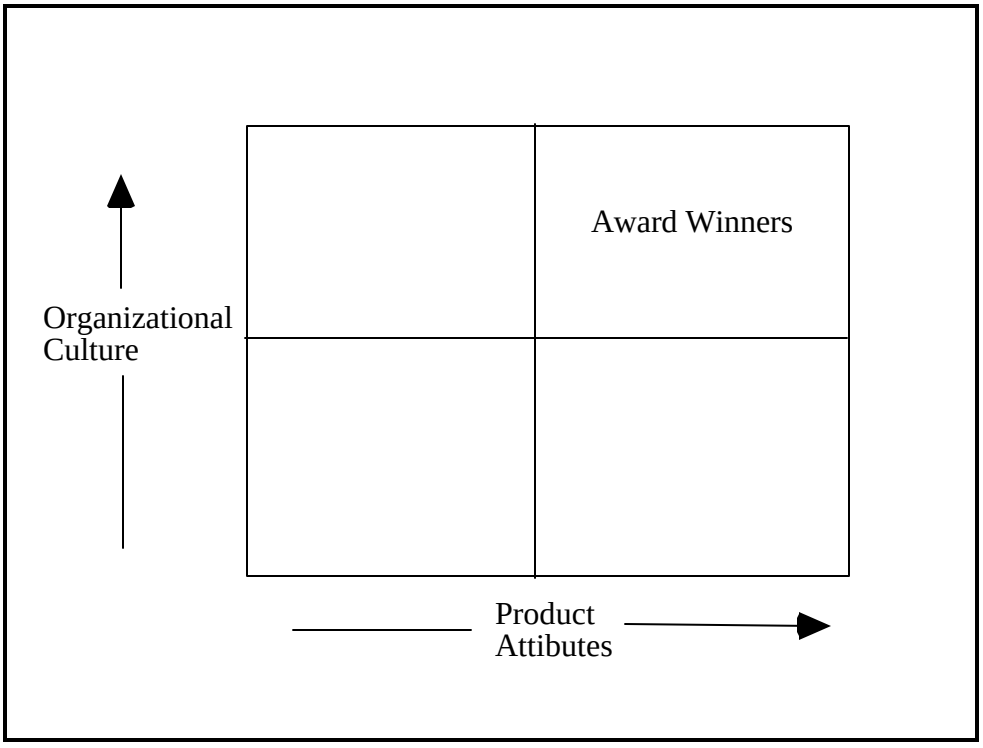


Figure III-2
Investments in Performance Excellence



IV

Empirical Models of Investments in Performance Excellence

A. Attributes Specified in the Baldrige Criteria for Performance Excellence

The Baldrige Criteria listed in Section II define the quality attributes recognized as the hallmarks of Baldrige Award winners and more generally of organizations achieving quality in their overall performance. The criteria imply empirically measurable proxies for the attributes to be used in formal tests linking different types of investments in quality to the performance of organizations. The Appendix to this report discusses some of the formal tests in the literature about quality management.

B. Relationship between Attributes and Firm Performance

The key fact potentially linking performance excellence attributes embodied in the Baldrige Criteria to the characterization of organization performance is that, absent a single accepted theory of Total Quality Management, firms have used the Baldrige Criteria to develop and guide their own TQM policies. As reported by Black and Porter (1996, p. 2):

> Sunday and Liberty [1992] suggest that large numbers of organisations use the criteria in this way. This is supported by the fact that despite the thousands of requests for copies of the criteria booklet, only a few hundred organizations actually apply for the award [Heaphy, 1992; Sunday and Liberty, 1992]. The circulation of copies of the guidelines has also increased dramatically since the award's inception …, making the Baldrige Award the best established framework for TQM practice. More recently, the European Quality Award assessment model … introduced in 1992, is available to organisations wishing to use the criteria for self-assessment of their TQM practice.

Similarly, Juran (1995) and George and Weimerskirch (1994), as observed in the literature review, emphasize the acceptance and use of the Baldrige Criteria to inform quality management practices by large numbers of organizations. Further, organizations in the rapidly evolving service sector are increasingly using the Baldrige Criteria (Blodgett, 1999). Moreover, the key hypothesis about the impact of the Baldrige National Quality Program is linked to the foregoing key fact—namely, the Program is hypothesized to have improved performance throughout the U.S. economy by emphasizing quality management and by providing a coherent and comprehensive set of criteria that organizations can adapt and use in creating their own quality management programs.

V

Applications to the Baldrige National Quality Program

A. Trends in Applications

Figure V-1 illustrates that the number of applications to the Baldrige National Quality Program decreased during much of the 1990s. The increase in applicants for 1999 and 2000 is because the Award's scope was broadened to include awards for educational organizations and health care providers. The number of applications from the traditional areas in 1999 and 2000 were 27 and 30, respectively. Blodgett (1999, p. 74) reports:

> In 1999, there were 52 Baldrige applicants, of which almost half were from the two new categories—education and health care. Among the total, there were 16 educational organizations, nine health care organizations, 11 service companies, 12 small businesses and four manufacturers.

B. Award Winners

The companies that have received the Malcolm Baldrige National Quality Award through 2000 are listed in Table V-1. Through 2000, 43 awards were announced.

Figure V-1
Applications to the Baldrige National Quality Program, 1988-2000

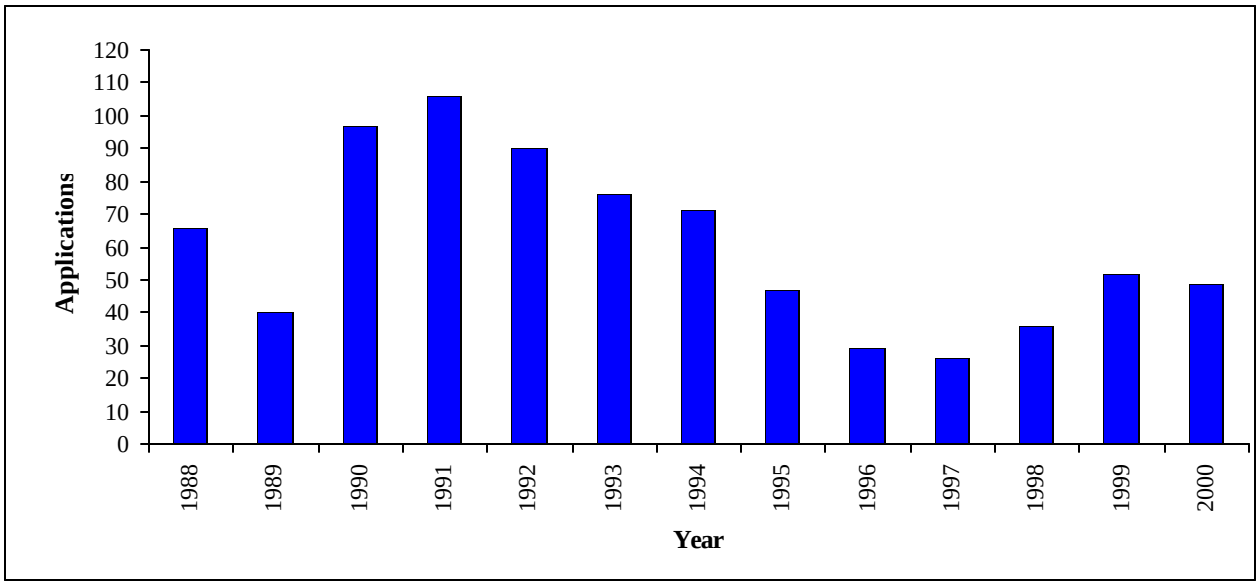


Table V-1
Malcolm Baldrige National Quality Award Winners, 1988-2000

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 & 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.
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: http://www.quality.nist.gov.

VI

Economic Evaluation of the Baldrige National Quality Program

A. Overview of the Evaluation Methodology

A.1. Systematic approaches to program evaluation

Fundamental to any evaluation of a federal program, research program or otherwise, is that the program is accountable to the public. For research programs, such accountability refers to being able to document and evaluate research performance using metrics that are meaningful to the institutions' stakeholders—the public, including taxpayers.

The Government Performance and Results Act of 1993 (GPRA) is clear that public institutions' research programs will identify outputs and quantify the economic benefits (outcomes) associated with such outputs. Some public agencies skirt the issue by arguing that the research they do or that they fund is peer reviewed, and thus it is sound; and if the research is sound, it must be socially valuable.

Many will embrace the importance of having research peer reviewed both at the pre-funding stage as well as upon completion. However, the peer review process does not address in any precise or reliable way whether or not the research is socially valuable from an economic standpoint. It is not so much that a formal analysis of social (aggregate) economic rates of return is officially out of bounds for the peer review process; rather, such an analysis is simply not a part of the review process as it currently exists. Other public agencies are attempting to be more exact in their approach to meeting the GPRA requirement to quantify outcomes. However, the hurdle that is difficult to clear for most public agencies is how to quantify benefits in a methodologically sound and defensible way.

Link and Scott have developed through ongoing evaluations of federal research programs two approaches to the *economic* evaluation of publicly-funded research. These approaches are somewhat at odds with traditional evaluation methods, apart from peer review, but traditional evaluation methods are limited in a GPRA world that is performance accountable since the question asked by these traditional methods is less appropriate than the ones that Link and Scott advocate.

When evaluating publicly-funded and publicly-performed research, the Link and Scott approach is based on a counterfactual evaluation method;[6] when evaluating publicly-funded privately-performed research programs, the Link and Scott approach is based on a spillover evaluation method.[7] Both of

[6] The genesis of this approach is in Link (1996a), and recent applications are in Link (1996b) and Link and Scott (1998).

[7] The genesis of this approach is in Link and Scott (2001) and Scott (1998), and recent applications are in Audretsch, Link, and Scott (2001).

these evaluation methods are set forth as background for this evaluation of the Baldrige National Quality Program.

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.[8] 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**.**

A.2. 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 benefits and costs 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. 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 in 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.

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

The following two evaluation methods are more appropriate for publicly-funded research than the traditional economic approaches. The first method is for such research also being publicly performed and the second for the research being privately performed.

A.3. 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 that question reveals benefits of the public's investments—namely, the private sector's costs avoided because of the public's investments. Additional benefits are the benefits from the public sector's investments that industry would be unable or unwilling to duplicate.[9] With those benefits of costs avoided plus the value of benefits that industry could not have replicated—obtained in practice through extensive interviews with administrators, federal research scientists, and those in the private sector who would have to duplicate the research in the absence of public performance—counterfactual rates of return and benefit-to-cost ratios can be calculated. Those metrics answer the fundamental evaluation question: Are the public investments a more efficient way of generating the technology than private sector investments would have been?

The answer to that fundamental question aligns with the public accountability issues implicit in GPRA, and 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 question is yes; the counterfactual method tests that hypothesis.

A.4. The spillover evaluation method

There are important projects where economic performance can be improved with public funding of privately-performed research. Public funding is needed when socially valuable projects would not be undertaken without it. If the expected private rate of return from a research project falls short of the

[9] 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 are surplus-creating innovations more generally) stream of returns to the R&D investments is valued as benefits. In that special case, the approached used herein 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 the respondents have reported, and further to be conservative in the estimates of the benefits of the Baldrige Program, it is assumed throughout that the private sector could—for the additional costs identified in the survey—have replicated the results of the Program.

required rate, then the private sector firm will not invest in the project. Nonetheless, if the benefits of the research spill over to consumers and to firms other than the ones investing in the research, the social rate of return may exceed the appropriate required rate. It would then be socially valuable to have the investments made, but since the private investor will not make them the public sector should. By providing some public funding and thereby reducing the investment amount needed from the private firm or firms doing the research, the expected private rate of return can be increased above the required rate. Thus, because of this subsidy, the private firm is willing to perform the research that is socially desirable because much of its output spills over to other firms and sectors in the economy.

The question asked in the spillover method is one that facilitates an economic understanding of whether the public sector should be underwriting a portion of private-sector firms' research. What proportion of the total profit stream generated by the private firm's R&D and innovation does the private firm expect to capture? Hence, what proportion is not appropriated but is instead captured by other firms that imitate the innovation or use knowledge generated by the R&D to produce competing products for the social good?

The part of the stream of expected profits captured by the innovator is its private return, while the entire stream is the lower bound on the social rate of return. In essence, the method weighs the private return, estimated through extensive interviews with firms receiving public support about their expectations of future patterns of events and future abilities to appropriate R&D-based knowledge, against private investments. The social rate of return weights the social returns against the social investments.

The application of the spillovers model to the evaluation of public funding / private performance of research is appropriate since the output of the research is only partially appropriable by the private firm with the rest spilling over to society. The extent of the spillover of such knowledge with public good characteristics determines whether or not the public sector should fund or partially fund the research.

A.5. Methodology applicable to an evaluation of the Baldrige National Quality Program

In a broad sense, the Baldrige National Quality Program is a measurement-and-standards infrastructure research and development 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 application of the criteria for performance excellence are evaluated. In this 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, the counterfactual evaluation method is directly applicable 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.

As discussed below, organizational benefit and private assessment-cost data were collected through surveys to the U.S. organizational members of the American Society for Quality, and then extrapolated to the aggregate economy. Cost data describing the combination of public and private funds used to administer the Program 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.

B. American Society for Quality [10]

The American Society for Quality (ASQ) agreed to a request from the management of the Baldrige National Quality Program Office to distribute a survey administered by the Program Office to 875 U.S. private-sector companies and public-sector organizations (hereafter, members).[11]

ASQ's 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:

- 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 February 16, 1946.

C. Social Costs of Operating the Baldrige National Quality Program

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.

[10] See http://www.asq.org.

[11] There are in addition to these U.S. organizational members, over 200 international organizational members plus over 120,000 individual members.

The public source of funds for the Baldrige National Quality Program is an annual allocation from the NIST budget. Column (2) of Table VI-1 shows the Program's annual allocations from NIST by fiscal year beginning with its first year of operation, 1988.

In addition to the pubic 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 VI-1 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.[12]

Industry also supports the Program through volunteer examiners during the application and evaluation process. In column (5) of Table VI-1 are the total hours of examiner time devoted to training, application review, and site visits.

Column (6) of Table VI-1 reports the estimated Program costs (in constant 2000 dollars), by year. The present value of these costs, brought forward at the real social rate of return of 7 percent to account for the social opportunity costs of these funds following the guidelines of OMB (1992), is $118,617,000.

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

D. Social Benefits Associated with the Baldrige National Quality Program

D.1. Evidence of social benefits associated with the Baldrige National Quality Program

As previously reviewed, a number of studies, including some conducted within the Baldrige National Quality Program Office, conclude that the market's valuation of the financial and managerial performance of companies that have received the Baldrige Award is greater than for comparable companies that have not.

Also, with reference to the management literature on quality, George and Weimerskirch (1994, pp. 5-6) champion the Baldrige criteria as the leading model of total quality management (TQM) with the following observations:

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

[12] The Foundation reimburses between 60% and 70% of examiner travel costs and the remainder is paid by the examiner's company or organization.

- 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. [NIST now estimates that by 2001 the number of copies distributed is nearly two million, with the number of copies made being at least as many, including an indeterminate but large number of copies downloaded from the Baldrige Program's web site.]
- More than half the states in the country now have state quality award programs based on the Baldrige criteria. [NIST now reports that by 2001 the number of states with such programs exceeds forty.]
- Several countries including Argentina, Australia, Brazil, Canada, and India are developing or have implemented quality award programs based on the Baldrige criteria. [NIST reports that by 2001 the number of countries with such programs exceeds fifty.]
- 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 U.S. 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) 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 observations about aspects of the social benefits associated with the Baldrige National Quality Program, as shown in Table VI-2 the Baldrige Criteria have been adopted by states as a foundation or benchmark for their own quality award programs, thus signifying one dimension of benefits. In fact, this evidence supports, in part, the manner in which the counterfactual evaluation method is applied herein.

D.2. Estimation of 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 in detail below; here is a brief overview. Benefit data were collected by survey from a sample of the U.S. organizational members of ASQ. These benefit data

were extrapolated first to the ASQ U.S. organizational membership as a whole, and then to the U.S. 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 constant 2000 dollars).

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 constant 2000 dollars).

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

Step 1 quantifies the probability that an ASQ member who received a survey would respond to the survey. As noted above, the ASQ agreed to distribute to its 875 U.S. organizational members a survey administered by the Baldrige National Program Office. Sixty-five organizations returned completed or partially completed survey instruments.[13]

The average probability of response is 65 returned surveys out of 875 sent surveys, or a 7.43 percent response rate. However, for the statistical analysis that follows, an estimated probability of response for each of the 875 members is needed as a control variable used in Step 2—to avoid biased estimates that might result if without the variable the error in the model for the probability of self-assessment would be correlated with the probability of response.

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

(1) Prob (response) = F (2-digit SIC industry variables)

where the dependent variable used to estimate equation (1) equals 1 if the member returned a completed or partially completed survey and 0 otherwise, and where the 2-digit Standard Industrial Category (SIC) industry variable categories are as described in the note to Table VI-3.[14] Equation (1)

[13] ASQ sent an electronic reminder to each survey recipient approximately three weeks after the initial mailing. No member-specific information is reported herein to ensure confidentiality.

[14] ASQ provided the 2-digit industry for approximately 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.

posits that the probability of a member responding can be predicted based on the industry in which that member produces.

The probit results from equation (1) are in Table VI-3.

For each of the 875 surveyed members, equation (1) produces a predicted value for the probit index, z, for the probability of response.[15,16]

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 for the estimation of net benefits because benefits are realized only when a self-assessment is performed. First, a probability of self-assessment model is estimated and second, a prediction of the probability of self-assessment for each ASQ member is calculated in Step 3 below.

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:

(2) Prob (self-assessment) = F (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 company 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 company applied for the Malcolm Baldrige National Quality Award? If yes, in what year(s)?

[15] 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 nonresponse perfectly, and the 16 observations were dropped from the sample used to estimate the model and assigned a probability of response of zero.

[16] Based on equation (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 $F'(z)dz / (1-F(z))$.

Has your company applied for a state quality award? If yes, in what year(s)?

and where the competitiveness variables noted in equation (2) above are defined in terms of a member's Likert responses (7 = strongly agree to 1 = strongly disagree) with the following two survey statements:[17]

The possibility or threat of new competition is significant. (*comp*)

Our customers have a significant ability to bargain on the price of our primary products. (*barg*)

and where the relevant control variables (to ensure against any response bias in the estimates) are based on estimates of the probability of response (*probres*) to the survey from equation (1).

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

The probit results from equation (2) are in Table VI-4.[18,19]

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 equation (2) presented in Table VI-4. Step 3 builds on steps 1 and 2 in that the first step provided a probability of response used to estimate well (using a subset of ASQ members) the model of the probability of self-assessment in the second step. That model in turn is now used in the third step to estimate a probability of assessment for all of the ASQ members.

As noted, with reference to the estimation of equation (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 *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 as reviewed above. 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 for example, the possibility or threat of new competition.

[18] When the hazard rate is included in equation (2) in place of the probability of response, the estimated probit model performed almost identically to the model reported in Table VI-4. 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.

[19] The model in equation (2) is estimated with 65 observations, however the 2-digit industry variables, *dtrcomut* and *dfire*, are dropped along with the 5 observations where they equal 1 because they predict assessment perfectly. Thus, the results in Table VI-4 are based on 60 observations.

With reference to equation (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 VI-4.[20]

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 company 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 above with reference to the counterfactual evaluation method, members' responses to this statement represent credible time estimates of the benefits (i.e., the costs avoided reported in constant 2000 dollars) 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 company has been an award applicant, what was the total economic cost (fully burdened) to your company to obtain, understand, collect relevant information, and comply with the Baldrige Criteria or state application requirements?
> $_____ per year during the year(s) _____.

and,

[20] The mean value of the probit index (n=810) = -0.7041409, corresponding to a probability of assessment = 0.2602325. In the calculations below, 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 equation (2)—recall from the preceding footnote that those two categories are perfect predictors of assessment—and hence n=810. In the calculations below, 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 equation (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 below.

> If your company 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 company to perform the self-assessment?
> $_____ per year during the year(s) _____.

Thus, for each of the 14 responding members, a time series of real (in constant 2000 dollars) costs incurred to make the Baldrige Criteria operational is also developed.[21]

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-.03)/(1+.03), where k is each member's reported hurdle rate and where the prevailing annual rate of price inflation over the reported time intervals is estimated at 3 percent.[22] 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 percent 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:[23]

(3) $NPV_{2000} = F$ (2-digit industry variables, size variables)

where member size was provided by ASQ for 874 of the 875 members. The least-squares results from equation (3) are in Table VI-5.[24]

The estimated coefficients in equation (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.

---

[21] Such costs are often referred to as pull costs. See Link and Scott (1998).

[22] Regarding the hurdle rate, each member was asked to respond to the following statement:

> 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)?
> _____ percent.

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 since that yields the rate of return r in constant dollars: X(1+r)=Y/(1+a). Since 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.

[23] The mean value of NPV_{2000} (n=14) = $17.7 million.

[24] 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.

The predicted values from equation (2) are point estimates for the probability of each member of ASQ conducting a self-assessment. The predicted values from equation (3) are 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. By using the standard errors of the predictions from equations (2) and (3), there will be explicit recognition of the uncertainty in the estimates from 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 equation (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.[25]

Second, regarding the predicted value, conditional on self-assessment, of the net present value of benefits associated with the Baldrige Program from equation (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 the 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 equation (2) and the lower-bound of the net present value of benefits from equation (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 percent 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 values 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 percent probability of being greater than the one used.

The sum of the lower-bound derived value of net benefits for ASQ members is $2.17 billion.[26]

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

[25] 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.

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

billion. When compared to 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-to-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, then total social benefits can be forecast using the following formula:

(4) Economy Value = (value for ASQ) /
(proportion taken by the ASQ members in the 50 represented industrial sectors)

where the latter value is calculated to be 0.0880285.[27]

Thus, under this assumption, the conservative present value of social benefits is $24.65 billion.[28] When compared to the present value of the total social cost associated with the Program of $119 million, the ratio of ASQ benefits to social costs is 207-to-1.[29]

E. 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 net private benefit to ASQ members as a result of the Baldrige National Quality Program is $2.17 billion (rounded in constant 2000 dollars). And, the conservative estimate of the present value of aggregate economy-wide net social benefits associated with the Program through 2000 is $24.65 billion (again, rounded in constant 2000 dollars). As also explained above, the present value of the social cost to operate the Program through 2000 is $119 million (rounded in constant 2000 dollars). From an evaluation perspective, these values yield benefit-to-cost ratios between 18.2-to-1 and 207-to-1.

---

[27] The size data for industrial sectors were assembled using information in U.S. 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.

[28] $2.17 billion / 0.088025 = $24.65 billion.

[29] 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 8 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 to total social costs of $119 million the resulting benefit to cost ratio is 173-to-1.

Table VI-1
Baldrige National Quality Program Operating Costs

(1) Fiscal Year	(2) NIST Allocations ($)	(3) Foundation Allocations ($)	(4) Company Reimbursed Examiner Expenses ($)	(5) Examiner Time (hours)	(6) Total Operating Costs (constant 2000 dollars)
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 constant 2000 dollars 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). Regarding the increased budget for 2000, recall that the number of applicants to the Baldrige Program increased in 1999 and 2000 because the Award's scope was broadened to include awards for educational organizations and health care providers.

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 constant 2000 dollars 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 constant 2000 dollars 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 year of examiner time is $125,000 (in constant 2000 dollars based on 2000 hours per year). The estimated value of examiner time is included in column (6) without additional adjustment.

Table VI-2
Application of Baldrige Criteria to State and Local Quality Award Programs

Year	Number of States with Award Programs Tied to Baldrige Criteria	Number of Service and Manufacturing Organizations that Applied for State and Local Quality Awards
1991	8	111
1992	12	144
1993	19	357
1994	29	428
1995	37	574
1996	42	804
1997	43	974
1998	44	830

Source: Blodgett (1999), NIST (1998).

Table VI-3
Probit Results for Probability of Response to the Survey (n=859)
(asymptotic t-statistics in parentheses)

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
Chi-squared (10)	19.94**

Notes:
***significant at 0.10 level **significant at 0.05 level *significant at 0.01 level
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.
dnonmin =1 for the agriculture, forestry, fisheries, minerals, and construction industries, and 0 otherwise; includes SICs < 20.
dchempet = 1 for chemicals, petroleum, and rubber, and miscellaneous plastics, and 0 otherwise; includes SICs 28, 29, and 30.
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.
dmats = 1 for the remaining manufacturing SICs, and 0 otherwise; includes SICs 20 through 27 and SICs 31 through 34; observations with *dmats* = 1 are in the intercept.
dtrcomut = 1 for transportation, communications, and utilities, and 0 otherwise; includes all 2-digit SICs greater than 39 and less than 50.
dwholret = 1 for wholesaling and retailing, and 0 otherwise; includes all 2-digit SICs greater than 49 and less than 60.
dfire =1 for finance, insurance, and real estate, and 0 otherwise; includes all 2-digit SICs greater than 59 and less than 70.

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 SIC 80.
dbusser = 1 for business services, and 0 otherwise; includes SIC 73.
dhealth = 1 for health services, and 0 otherwise; includes SIC 80.
dpubadm = 1 for public administration, and 0 otherwise; includes 2-digit SICs greater than 89 and less than 100.

Table VI-4
Probit Results for Probability of Self-Assessment (n=60)
(asymptotic t-statistics in parentheses)

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
Chi-squared (5)	9.11***

Notes:
***significant at 0.10 level **significant at 0.05 level *significant at 0.01 level
There are 65 observations available to estimate the model in equation (2); however, the 2-digit industry variables, *dtrcomut* and *dfire*, are dropped along with the 5 observations where they equal 1 because they predict assessment perfectly. Thus, the results above are based on 60 observations.

Table VI-5
Least-Squares Results for Net Present Value of Benefits (n=14)
(t-statistics in parentheses)

Variable	Estimated Coefficient
size	-83844.49** (-2.48)
$size^2$	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

Notes:
***significant at 0.10 level **significant at 0.05 level *significant at 0.01 level
The explanatory member-size variable is measured in millions of dollars, while the dependent variable for value is measured in dollars.

VII

Summary

This report summarizes the findings from an economic evaluation of Baldrige National Quality Program. Specifically, the net private benefits associated with the Baldrige National Quality Program to the economy as a whole are conservatively estimated to be $24.65 billion. When compared to the social costs associated with the Program of $119 million, it is clear that, from an evaluation perspective, the Baldrige National Quality Program is socially beneficial as summarized by a benefit-to-cost ratio of 207-to-1.

Regarding the generalization from the ASQ membership to the entire economy, thus producing a social benefit-to-cost ratio of 207-to-1, only 11 of the 875 ASQ members have received to date the Malcolm Baldrige National Quality Award.[30] Clearly, based on the requests for Baldrige application materials and criteria and 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, generalizing about the net social benefits of the Program from ASQ members to the economy as a whole may underestimate the true social benefits associated with the Program. In that case, the social benefit-to-cost ratio of 207-to-1 understates, even beyond the conservative estimation procedure used in this study, the true benefits of the Program.

[30] One of the 11 Baldrige Award winners was among the 14 providing the detailed data about the net present value of benefits that was used to estimate the model reported in Table VI-5. Accounting for the presence of the Award winner in that sample does not affect the estimates obtained with the model. The model reported in Table VI-5 was reestimated with the addition of a qualitative variable equal to 1 for the observation of the Award winner and zero otherwise. The estimates of the parameters of the model and their significance were virtually unchanged (the estimated parameters and their t-statistics were essentially the same), and the coefficient on the qualitative variable indicating the Award winner was not significantly different from zero. The t-statistic for the coefficient was just 0.13 (the probability of a greater absolute value for the statistic given the null hypothesis is 0.896); the R^2 increased from the 0.513 reported for the model shown in Table VI-5 to just 0.514.

References

Aghion, Philippe, and Howitt, Peter, "A Model of Growth Through Creative Destruction," *Econometrica*, vol. 60 (1992), pp. 323-351.

Ahire, Sanjay L., Golhar, Damodar Y., and Waller, Matthew A., "Development and Validation of TQM Implementation Constructs," *Decision Sciences*, vol. 27, no. 1 (Winter 1996), pp. 23-56.

Anderson, J. C., Rungtunsanatham, M., and Schroeder, R. G., "A Theory of Quality Management underlying the Deming Management Method," *The Academy of Management Review*, vol. 19, no. 3 (1994), pp. 472-509.

AT&T, *A History of Quality Control and Assurance at AT&T, 1920-1970*. A Series of 14 video cassettes available from AT&T Customer Information Center (1989).

Audretsch, David B., Albert N. Link, and John T. Scott, Public/Private Technology Partnerships: Evaluating SBIR-Supported Research," *Research Policy*, forthcoming.

Baldwin, William L., and Scott, John T., *Market Structure and Technological Change* (Chur: Harwood Academic Publishers, 1987).

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, vol. 63, no. 4 (July 1995), pp. 841-890.

Black, Sandra E., and Lynch, Lisa M., "Human Capital Investments and Productivity," *American Economic Review* (May 1996a).

Black, Sandra E., and Lynch, Lisa M., "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity," manuscript, Harvard University and U.S. Department of Labor, May 1996b.

Black, Simon A., and Porter, Leslie J., "Identification of the Critical Factors of TQM," *Decisions Sciences* , vol. 27, no. 1 (Winter 1996), pp. 1-21.

Blodgett, Nancy, "Service Organizations Increasingly Adopt Baldrige Model," *Quality Progress*, December 1999, pp. 74-76.

Bresnahan, Timothy F., "Departures from Marginal-cost Pricing in the American Automobile Industry," *Journal of Econometrics*, Vol. 17 (1981), pp. 201-227.

Bresnahan, Timothy F., "Competition and Collusion in the American Automobile Oligopoly: The 1995 Price War," *Journal of Industrial Economics*, Vol. 35, 1987, pp. 457-482.

Bush, Vannevar, *Science—the Endless Frontier*. Washington, DC: U.S. Government Printing Office, 1945.

Caballero, Ricardo J., and Jaffe, Adam B., "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," *NBER Macroeconomics Annual 1993*, edited by Olivier Jean Blanchard and Stanley Fischer (Cambridge, Massachusetts: MIT Press, 1993), pp. 15-74.

Cohen, W. M., and Levin, R. C., "Empirical Studies of Innovation and Market Structure," *Handbook of Industrial Organization*, Vol. II, edited by R. Schmalensee and R. D. Willig (Amsterdam: North-Holland, 1989).

Committee on Science, Engineering, and Public Policy, *Evaluating Federal Research Programs*, Washington, DC: National Academy Press, 1999.

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

Council on Competitiveness, "Quality is Key for Tomorrow's Workplace," *Challenges*, vol. 9, no. 7 (October/November 1996), pp. 1, 4.

Council of Economic Advisers, *Economic Report of the President*, January 2001, Washington, D.C.: U.S. Government Printing Office, 2001.

Feenstra, Robert C., and Levinsohn, James A., "Estimating Markups and Market Conduct with Multidimensional Product Attributes," *Review of Economic Studies*, Vol. 62, No. 1 (January 1995), pp. 19-52.

Garvin, D. A., "Competing on the Eight Dimensions of Quality," *Harvard Business Review*, vol. 65, no. 6 (1987), pp. 101-109.

George, Stephen, *The Baldrige Quality System: The Do-It-Yourself Way to Transform Your Business* (New York: John Wiley & Sons, Inc., 1992).

George, Stephen, and Weimerskirch, Arnold, *Total Quality Management: Strategies and Techniques Proven at Today's Most Successful Companies* (New York: John Wiley & Sons, Inc., 1994).

Goldberg, Pinelopi K., "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, vol. 63, no. 4 (July 1995), pp. 891-951.

Grant, E. L., "Shewhart Medal Address," *Industrial Quality Control* (July 1953).

Grant, E. L., "Statistical Quality Control in the World War II Years," *Quality Progress* (December 1991), pp. 31-36.

Griliches, Zvi, "Research Costs and Social Returns: Hybrid Corn and Related Innovations," *Journal of Political Economy*, vol. 66 (1958), pp. 419-431.

Grossman, Gene M., and Helpman, Elhanan, "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, vol. 58 (1991a), pp. 43-61.

Grossman, Gene M., and Helpman, Elhanan, "Quality Ladders and Product Cycles," *Quarterly Journal of Economics*, vol. 106 (May 1991b), pp. 557-586.

Heaphy, M. S., "Inside the Baldrige Award Guidelines, Category 5: Management of Process Quality, *Quality Progress*, vol. 25, no. 10 (1992), pp. 74-79.

Helton, B. Ray, "The Baldie Play," *Quality Progress*, vol. 28, no. 2 (February 1995), pp. 43-45.

Hendricks, Kevin B., and Singhal, Vinod R., "Does Implementing an Effective TQM Program Actually Improve Operating Performance: Empirical Evidence from Firms that Have Won Quality Awards," *Management Science*, vol. 44, no. 9 (1997), pp. 1258-1274.

Hendricks, Kevin B., and Singhal, Vinod R., "The Long-Run Stock Price Performance of Firms with Effective TQM Programs as Proxied by Quality Award Winners," *Management Science*, vol 47, no. 3 (2000), pp. 359-368.

Hendricks, Kevin B., and Singhal, Vinod R., " Firm Characteristics, Total Quality Management, and Financial Performance, *Journal of Operations Management*, vol. 19 (2001), pp. 269-285.

Hertz, Harry, "Letter to the Editor," Manufacturing News, vol. 7, no. 19 (October 30, 2000), p. 2.

Huselid, Mark A., and Becker, Brian E., "High Performance Work Systems and Firm Performance: Cross-Sectional versus Panel Results," *Industrial Relations*, (1996), in press.

Ichniowski, Casey, Shaw, Kathryn, and Prennushi , Giovanna, "The Effects of Human Resource Management Practices on Productivity," NBER Working Paper No. 5333 (November 1995).

Juran, J. M., "The Taylor System and Quality Control," a series of articles in *Quality Progress*, May 1973 through December 1973 (listed under "Management Interface").

Juran, J. M., "World War II and the Quality Movement," *Quality Progress* (December 1991), pp. 19-24.

Juran, J. M., editor, *A History of Managing for Quality: The Evolution, Trends, and Future Directions of Managing for Quality* (Milwaukee, Wisconsin: ASQC Quality Press, 1995).

Konczak, Lee J., "Review" of Lawler, Mohrman, and Ledford (1995), *Personnel Psychology* (April 1996), pp. 495-499.

Lau, R. S. M., "A Survey of Competitiveness and Manufacturing Practices," *South Dakota Business Review*, vol. LIV, No. IV (June 1996), pp. 1, 4-9.

Lawler, Edward E., III, Mohrman, Susan Albers, and Ledford, Gerald E., Jr., *Creating High Performance Organizations: Practices and Results of Employee Involvement and Total Quality Management in Fortune 1000 Companies* (San Francisco: Jossey-Bass, 1995).

Link, Albert. N., *Technological Change and Productivity Growth* (Chur: Harwood Academic Publishers, 1987).

Link, Albert N., "Economic Impact Assessment Guidelines for Conducting and Interpreting Assessment Studies," NIST Planning Report 96-1, 1996a.

Link, Albert N., *Evaluating Public Sector Research and Development*, New York: Praeger Publishers, 1996b.

Link, Albert N., and Tassey, Gregory, *Strategies for Technology-based Competition: Meeting the New Global Challenge* (Lexington, Massachusetts: Lexington Books, 1987).

Link, Albert. N., Quick, Perry., and Tassey, Gregory., "Investments in Product Quality: A Descriptive Study of the US Optical Fiber Industry," *Omega: International Journal of Management Science*, vol. 19, no. 5 (1991), pp. 471-474.

Link, Albert N. and John T. Scott, *Public Accountability: Evaluating Technology-Based Institutions*, Norwell, Mass.: Kluwer Academic Publishers, 1998.

Link, Albert N., and John T. Scott, "Estimates of the Social Returns to Small Business Innovation Research Projects," in *The Small Business Innovation Research Program* (edited by Charles W. Wessner), Washington, DC: National Academy Press, 2000, pp. 275-290.

Link, Albert N. and John T. Scott, "Public/Private Partnerships: Stimulating Competition in a Dynamic Market," *International Journal of Industrial Organization*, vol. 19 (April 2001), pp. 763-794.

Mansfield, Edwin, John Rapoport, Anthony Romeo, Samuel Wagner, and George Beardsley, "Social and Private Rates of Return from Industrial Innovations," *Quarterly Journal of Economics*, 91 (1977), pp. 221-240.

Martin, Stephen, *Advanced Industrial Economics* (Cambridge, Massachusetts: Blackwell Publishers, 1993).

Martin, Stephen, "Empirical Research in Industrial Economics: The Automobile Industry," manuscript, Centre for Industrial Economics, Institute of Economics, University of Copenhagen, July 1996.

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

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

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

NIST, Malcolm Baldrige National Quality Award: 2001 Criteria for Performance Excellence (Gaithersburg, Maryland: National Institute of Standards and Technology, 2001), contents also available at http://www.quality.nist.gov/

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

Office of Management and Budget, "Circular No. A-94: *Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs*" (Washington, DC: Office of Management and Budget, 1992).

Porter, Michael E., *Competitive Advantage: Creating and Sustaining Superior Performance* (New York: The Free Press, 1985).

Porter, Michael E., *Competitive Strategy: Techniques for Analyzing Industries and Competitors* (New York: The Free Press, 1980).

Reimann, Curt W., and Hertz, Harry S., "The Malcolm Baldrige National Quality Award and ISO 9000 Registration: Understanding Their Many Important Differences," *ASTM Standardization News*, (November 1993), pp. 42-53.

Reinganum, Jennifer F., "The Timing of Innovation: Research, Development, and Diffusion," *Handbook of Industrial Organization*, Vol. I, edited by R. Schmalensee and R. D. Willig (Amsterdam: North-Holland, 1989).

Scherer, F. M., "The Welfare Economics of Product Variety: An Application to the Ready-To-Eat Cereals Industry," *Journal of Industrial Economics*, vol. 28 (December 1979), pp. 113-134.

Scherer, F. M., and Ross, David, *Industrial Market Structure and Economic Performance*, 3rd edition (Boston: Houghton Mifflin, 1990).

Scott, John T., "Financing and Leveraging Public/Private Partnerships: The Hurdle-Lowering Auction," *STI (Science, Technology, Industry) Review*, vol. 23 (1998), pp. 67-84.

Scott, John T., "The Service Sector's Acquisition and Development of Information Technology: Infrastructure and Productivity," *Journal of Technology Transfer*, vol. 24 (1999), pp. 37-54.

Solow, Robert M., "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, vol. 39 (August 1957), pp. 312-320.

Sunday, J. L., and Liberty, L., "Benchmarking and the Baldrige Award," *Quality Progress*, vol. 25, no. 9 (1992), p. 75.

Tassey, Gregory, *The Economics of R&D Policy* (Westport, Connecticut: Quorum Books, 1997).

Tassey, Gregory, "Lessons Learned about the Methodology of Economic Impact Studies: The NIST Experience," *Evaluation and Program Planning*, vol. 22 (1999), pp. 113-119.

Tirole, Jean, *The Theory of Industrial Organization* (Cambridge, Massachusetts: MIT Press, 1988).

Townsend, Pat, and Gebhardt, Joan, "The Importance of the Baldrige to the US Economy," *Journal for Quality and Participation*, (July/August 1996), pp. 6-13.

Trajtenberg, Manuel, *Economic Analysis of Product Innovation: The Case of CT Scanners* (Cambridge, Massachusetts: Harvard University Press, 1990).

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

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

Wisner, J. D., and Eakins, S. G., "A Performance Assessment of the U.S. Baldrige Quality Award Winners," *International Journal of Quality and Reliability Management*, vol. 11, no. 2 (1994), pp. 8-25.

Wareham, Ralph E., and Stratton, Brad, "Standards, Sampling and Schooling," *Quality Progress* (December 1991), pp. 38-42.

Working, Holbrook, "Statistical Quality Control in War Production," *Journal of the American Statistical Association*, (December 1945), pp. 425-447.

Appendix

Detailed Review of the Academic and Professional Literature Related to Investments in Quality

Investments in Quality and Firm Performance

Theoretical models of quality investment that emphasize investments in organizational culture

In the quality management literature, the theoretical methodology for modeling private investments in quality is, as explained by Black and Porter (1996), distilled from case studies and other evidence based on experience about efforts to achieve quality, and the collected wisdom therefrom is gathered together as the package of concepts and prescriptions called Total Quality Management (TQM). Much of this work can be characterized as focused on investments in organizational culture. Even when firm strategies include techniques for improving product attributes, such as techniques of statistical quality control, the emphasis remains on an organizational culture within which such efforts will work (as opposed to having various isolated segments of the organization work on a product and then "throw it over the wall" to wait for the quality department to catch the bad products at the end of the process) (Juran, 1995, p. 561).

Black and Porter (1996) categorize, in the context of the underlying literature, various managerial theories that are included under the TQM banner. TQM encompasses benchmarking, customer feedback, leadership, planning, process management, quality cost measurement, quality policies, quality systems, statistical process control, supplier management, teamwork, training, and zero defects. They (1996, p. 2) note, within the context of the numerous prescriptions for managers trying to guide their companies toward performance excellence by setting out the steps to follow, that "no single model has yet established itself as a basis for Total Quality Management theory."[31]

Ahire, Golhar, and Waller (1996) also review the quality management literature and also observe that the prescriptions for performance management are offered piecemeal—there is a plethora of various strategies offered. They (1996, p. 23) conclude that "contemporary quality management (QM) literature … lacks scientifically developed and tested constructs that represent an integrative QM philosophy;" thus, they are essentially in agreement with Black and Porter in that they (Ahire, Golhar and Waller 1996, p. 24) believe that much remains to be done in unifying approaches to performance management:

[31] This is not a surprising conclusion given the conceptual model of firm performance in Figure III-1. Therein it is clear that quality enters the model in many places and interacts with many elements of the model. Thus, any step-by-step prescription is obviously incomplete.

> The QM theory is far from being fully developed. Anderson, Rungtunsanatham, and Schroeder [1994] make the only known effort of synthesizing a theory of quality management. They assess the impact of Deming's management method on a firm's organizational behavior and practice of quality management. However, … this work suffers from a lack of systematic scale development, content validity, and empirical validation. Hence, it falls short on overall generalizability of results.

Black and Porter (1996) emphasize that because there is no single accepted theory of TQM, many firms have relied on quality award criteria to fashion their own TQM management practices. Such quality awards include the Deming Prize in Japan, the Malcolm Baldrige National Quality Award, and the European Quality Award. Designed to honor achievement of firms with successful quality management practices, the awards have certainly served as the basis for companies' own TQM policies. Indeed, George (1992) emphasizes the Baldrige Program's essence as a "Do-It-Yourself Assessment," and he also emphasizes that firms can use that assessment process to improve the performance of their overall operations.

Reimann and Hertz (1993, p. 52) explain, in contrast to ISO 9000 registration (which is a much narrower program in substance and procedure that is focused on conformity of specified operations to documented requirements to ensure buyers of the specified conformity), the primary purpose of the Baldrige Program is educational—to share learning about how firms can ensure overall operational excellence and to encourage the development and use of that learning by "… 1) promoting awareness of quality as an important element in competitiveness, 2) recognizing companies for successful quality strategies, and 3) fostering sharing of lessons learned."

The relevant literature supports the following interpretation: the criteria used to judge the most prominent quality award—the Baldrige Award—encompass the most comprehensive theory of performance management. That interpretation is expressed in the following proposition:

Companies that endeavor to implement the business practices embodied in the non-results categories of the Malcolm Baldrige National Quality Award criteria will be more likely to exhibit performance excellence.

The Baldrige Criteria, then, define a theory, perhaps the preeminent theory, of quality management. That point is put more emphatically by George (1992, p. 267), who concludes:

> The Malcolm Baldrige National Quality Award and the companies that have won it offer a beautiful melody, a do-it-yourself kit for transforming your organization. All that is left is the reaching.

A useful bridge from the theory of quality management to the empirical work describing and exploring its effects is provided by George and Weimerskirch (1994). They use the Baldrige Criteria as a model—a "model of integration" that provides a firm with a way of understanding and organizing its activities to be

more efficient and effective. For that reason George and Weimerskirch are classified here, in the theory section of the quality management literature. But, they also belong (below) in the empirical section of the quality management literature because they use the experiences of 53 leading U.S. companies to illustrate the model of total quality management that is defined by the Baldrige Criteria. They define a holistic theory of action with the Baldrige Criteria, and then they illustrate the theory with examples from leading companies. George and Weimerskirch (1994, p. 5) support the Baldrige Criteria as the leading model of TQM:

> We believe the Baldrige [C]riteria define the new management model because they provide the best guide to understanding, assessing, controlling, and improving your organization.

Theoretical models of quality investment that emphasize investments in product attributes

Economists have not directly addressed the issue of management policies to promote quality. They have studied investments that can generate higher quality and more efficient firm performance—investments in R&D and in product differentiation. The economists' niche within the conceptual literature focuses on investments in product attributes. Further, the related economic theory typically addresses questions other than management methods to improve quality, and instead focus on how characteristics of markets (such as the degree of competition in a market) and the broad characteristics of firms (such as the size or diversification of a firm) affect R&D investments which will ultimately lower costs or improve products.

Baldwin and Scott (1987), Reinganum (1989), Scherer and Ross (1990), and Martin (1993) provide reviews of the earlier theoretical economic literature studying innovative investments. There are two prominent strands of recent work in the economics literature that address the issue of quality; neither examines the issues of quality management addressed in the quality management literature. One strand of the recent economics literature is essentially a macroeconomic, international open economy literature grounded in the microeconomics of innovation. This is the "quality ladder" literature discussed briefly in Section II of this report. The other strand is a microeconomic literature that addresses the effects of differing qualities of a product in a market (local, domestic, or international). This second literature is refereed to herein as the "product-quality differentiation" literature. Both strands of this economic literature conclude that firms' investments in quality have important implications for the relative competitiveness and growth of entire nations, and of course the firms of those nations, as well as important implications for pricing and competitive position of companies within industries. This literature does not, however, speak to issues about which quality management practices to employ in which situation, or of how to effectively manage investments in quality or quality-related practices.

The theoretical portion of the "quality ladder" literature is in Grossman and Helpman (1991a, 1991b), Aghion and Howitt (1992) and papers cited therein, as well as in Caballero and Jaffe (1993)—which leads to the major empirical application of the "quality ladder" or "endogenous growth" literature as discussed in the following section. The recent theoretical literature brings together ideas from the

microeconomic literature about Schumpeterian competition reviewed by Baldwin and Scott (1987) and Reinganum (1989) with the macroeconomic literature about the theory of growth evolving from the work of Solow (1957). These models have been used to address broad, aggregate international economics issues—such as appropriate trade policies and intellectual property policies in the presence of spillovers of knowledge—about the rates of growth in industrialized countries pioneering innovations and in countries whose entrepreneurs imitate those innovations.

Caballero and Jaffe (1993, pp. 16-17) describe the "quality ladder" theory in the context of their version of the model which they develop to frame their empirical study:

> Our aim … is to create a framework for incorporating the microeconomics of creative destruction and knowledge spillovers into a model of growth, and to do so in such a way that we can begin to measure them and untangle the forces that determine their intensity and impact on growth. … We develop a model in the spirit of Grossman and Helpman … and Aghion and Howitt … that gives a simple relationship for the effect of new products on the value of existing ones. At any given time, the economy consists of a continuum of monopolistically competitive goods indexed by their quality, $q \in [0, N_t]$. The newest goods are always the best, i.e., the process of research advances the frontier by increasing N_t. Because of the quality ranking implicit in this process, constant marginal cost producers see their profits—relative to those of the (new) leader—decline over time. The rate of decline depends (positively) on the degree of substitutability between new and old goods and on the pace at which new goods are introduced. This captures the endogenous process of creative destruction described earlier and … yields intuitive equations relating the rate of growth in a firm's value relative to that of the industry to the firm's number of new ideas relative to the industry average. By relating the concept of new ideas to that of new patents, it is possible to use these equations to gauge the empirical magnitude of creative destruction.

The messages within the "quality ladder," endogenous growth literature are clearly of direct interest to managers. What could be more fundamental than the story of how innovation improves the quality of goods and thereby erodes the market position of laggards who do not keep pace? Equally clearly, however, the "quality ladder" literature is not generating results directly pertinent to the quality management literature in the sense that the results inform understanding of how economies evolve with different conditions for innovation and imitation and in different policy regimes rather than offer direct guidance about how to manage to managers trying to improve the performance of their companies, via innovation or in other ways. However, the results in the economics literature do point to controls needed in studies of managerial performance and can inform managerial decisions as Porter (1980, 1985) has shown.

The "product-quality differentiation" literature has always been squarely footed on empirical ground, and the empirical work is well motivated by theory. The roots of the work date to Bresnahan's (1981) important analysis of the automobile industry. Automobiles differ in many ways, and the demand side of the framework collapses those differences into a single dimension—quality. From the perspective of the

managerial literature, such an abstraction eliminates much of importance since ensuring the various aspects of quality would be expected to require different managerial practices. And the framework assumes that consumers are uniformly distributed—in terms of their preferred product types—over the continuous interval reflecting the different qualities of automobiles. As Martin (1996) observes, this assumption violates reality and eliminates an important problem for managers of new product developments—namely the decision of what sorts of varieties of the product should be produced to appeal to the discontinuous agglomerations of customers who, herd-like, prefer a particular variety. In any case, in the theoretical framework, the demand for a particular variety depends on its own price and quality and on the prices and qualities of the products are "close"—immediately to the left or right on the interval measuring quality. On the cost side, the firm's total costs increase with not only the amount of the good produced, but with its quality. Although surely that is a sensible abstraction for the purposes of the economics literature, the managerial literature needs the additional details about the costs of ensuring the various aspects of quality that are eliminated in the simple abstraction of the product-quality model.

Martin (1996) observes that the theoretical model assumes that the qualities of the varieties of products produced by each firm are given; then pricing decisions are made. Clearly from a managerial standpoint, the assumption is restrictive, since rivalry over varieties and qualities would be a key part of an actual competitive situation. However, the model is interesting for the purposes that it was developed—namely, the framework shows the pricing and allocative efficiency aspects of the Nash equilibrium emerging for the firms producing in a market with product-quality differentiation. The principal results (Martin, 1996) are that the socially optimal number of varieties may be fewer than the number actually produced (because individual firms introduce varieties based on the effect on their individual profits, ignoring the effects of their own products on the profits that other firms earn from competing varieties. Further, in addition to the usual welfare loss from higher prices caused by market power, there is in addition to the conventional deadweight loss (resulting because some consumers who would have bought the good if price had equaled marginal costs do not purchase anything given the higher price) the welfare loss caused by "quality downshifting" as consumers buy lower quality products rather than the higher quality ones they would have purchased if market power had not increased price. The purpose of the theory is to inform empirical estimation of the pricing effects in a market, and the findings will be summarized in the next section. Bresnahan (1987), among other things, uses the basic model just described to explore the effects of cooperation among a market's firms, where cooperation takes the form of joint profit maximization. Again, the theory provides the framework for hypothesis testing that is described in the next section.

The product-quality differentiation literature has blossomed into the so-called "new hedonics" literature. As a prominent example, in the theoretical model of Feenstra and Levinsohn (1995) the quality of products has a multidimensional character; with each product is associated a vector of characteristics. Again, the "quality," here the product's characteristics, is taken as given. Thus, as Martin (1996, p. 14) observes, just the second stage of a two-stage game is being modeled; the first stage in which rivalrous firms decide on the characteristics of their products is not modeled. With greater complexity of the product space, the model requires even more restrictive assumptions than Bresnahan's model. Once again, the model assumes that the distribution over the characteristics space of customers most

preferred types of products is uniform, with the associated loss of realism noted earlier. Note that the abstractions may be a problem for applications to the practice of quality management without reducing the model's usefulness for the purposes intended by the authors, so the comments about realism are intended only to draw the distinctions of importance for this review of the literatures about quality in the context of evaluating the Baldrige National Quality Program. Again, costs depend on "quality," but here that means that costs are a function of product characteristics. The equilibrium price structure with price-setting firms implies that the price of a model of automobile increases as the distance increases, in product characteristics space, between it and rival models. Feenstra and Levinsohn explore the nature of the equilibrium prices given various assumptions about whether firms set prices or quantities or some firms do one thing while others do another. As Martin (1996) points out, these possibilities are interesting in actual fact, because for example the use of voluntary export restraints implies that the foreign suppliers are setting quantities while the domestic suppliers could set prices. The basic messages of such theory are surely of interest to managers—some approaches to pricing or quantity decisions yield higher profits than others. However, these questions are rather far removed from the issues of managing for quality addressed by the Malcolm Baldrige National Quality Award. Other work in this area includes Berry, Levinsohn, Pakes (1995) and Goldberg (1995).

Trajtenberg (1990) places the "new hedonics" literature squarely in the context of innovations with his study of CT scanners that among other things relates citation-weighted patent counts to the willingness of consumers to pay for scanners with particular characteristics and hence to "quality" measured as consumer welfare. Such pioneering work is clearly of interest to managers pursuing success by applying the Baldrige Criteria because customer satisfaction is at the heart of the performance of a firm that successfully achieves the overall operating performance exemplified by the Baldrige Award. Yet, the work is of indirect importance for the quality management literature, which is of course focused on the methods by which managers can marshal the R&D investments and product designs to increase customer satisfaction.

Earlier theoretical literature reviewed by Tirole (1988, pp. 95-131) explores the economics of markets in which products are produced in a variety of qualities and at times consumers do not have complete information about the quality of products—knowledge about quality is acquired before or after purchase, or in some cases never. Economists' interests are primarily in understanding how the markets for such products work as well as how well they work— considering, for example, whether the number of varieties produced is optimal or how moral hazard and adverse selection affect the warranty system that emerges for a product for which customers do not have complete information, or how repeat purchases provide a way that customers can monitor quality, or when the difficulties consumers face in evaluating quality justify government intervention in the market process. The usefulness of such literature for quality management practices is, of course, indirect. For example, understanding how markets with incomplete information work could help a company design a product information and warranty system that gave it a competitive advantage over its rivals. Or, understanding the product selection problem, could allow a firm to position its products in the product characteristics space in a way that gave it a sustainable competitive advantage (Scherer, 1979). These are important issues, but they are not directly the issues faced in the quality management literature.

In all, the economics literature on product-quality differentiation offers insights about price-cost margins and competition that is of interest to managers; however, in the context of quality management issues of concern to those working with the Baldrige Criteria, the economics literature differs from the quality management literature in two important ways. First, at best the literature addresses just product quality, which is just one aspect of the holistic, total quality management set of issues. Second, the literature does not address the issues of which management tools can improve performance, but rather predicts pricing performance conditional on a variety of assumptions about firms, customers, and markets.

Empirical models of quality investment that focus on investments in organizational culture

There is a large and important body of empirical work that has developed information about the use of quality management and has tested the hypothesis that the adoption of such management practices improves performance. Much of the work in this area uses the Baldrige National Criteria to define total quality management and to test for the importance of its various categories in use and in results. In the quality management literature, empirical analyses of quality management practices began to develop rapidly, clearly stimulated by the newly awakened interests in quality management during the late 1980s and 1990s. The work largely emphasizes investments in organizational culture as a way to improve quality.

For example, Link, Quick, and Tassey (1991) compare the firms in the U.S. optical fiber industry with their foreign competitors by surveying the firms about the proportions of their operations, capital, R&D, and overhead budgets that are devoted to quality assurance and about the functional objectives of their quality budgets. Thus, the Link, Quick, and Tassey study looks at the whole firm and implicitly measures investments in both organizational culture and product attributes. However, the focus is on gaining understanding of the successful organization's orientation toward quality, and like most of the literature that is focused directly on management for quality, the study is then developing understanding of the orientation, in a broad sense the culture, of successful companies in their pursuit of quality. The U.S. firms were the dominant firms in the industry in the early 1990s, and the empirical methodology of the authors allows inferences about the link from quality investments to international competitiveness. The U.S. firms invested more—three to four times more—of their operations, capital, R&D, and overhead budgets in quality assurance, and within their quality budgets themselves they focused more attention, than their less successful foreign rivals did, on improving manufacturability and improving product performance.

Some studies explore the effects of using the Baldrige Criteria to implement quality management policies. The U.S. General Accounting Office (1991) found that firm performance was improved by the quality efforts of the highest scoring Baldrige Award applicants in 1988 and 1989. As reviewed by Black and Porter (1996, p. 2):

> The evidence from this small sample suggested that the organizations achieved improved employee relations, better quality, lower costs, greater customer satisfaction, improved market share and improved profitability. Common features appearing in these high-

> scoring organizations were customer focus, management leadership in quality values, employee involvement, an open corporate culture, fact-based decision making and partnerships with suppliers. This report offers sound evidence supporting the relevance of implementing and maintaining TQM as defined by the Baldrige assessment framework.

Other studies (Hendricks and Singhal, 1997, 2000, 2001); Helton, 1995; Wisner and Eakins, 1994) show that the Baldrige Award winners have enjoyed strong financial performances, implying that quality improvement programs leads to increases in market values for the firms that invest in quality improvements. Helton considered a hypothetical portfolio of $1,000 invested in the stock of each of the 11 publicly owned Baldrige Award winners through 1993, with the investment in each being made on the day the winner was announced by the U.S. Department of Commerce. By September 1, 1994, the cumulated investment of $11,000 (accumulated over the time from November 14, 1988 when $1,000 would have been invested in Motorola and Westinghouse to December 14, 1993 when $1,000 would have been invested in Eastman Chemical) would have resulted in a portfolio worth $21,887. Some of the Award winners were divisions of the publicly traded companies whose stock prices were tracked, so the actual performance of a fund invested in just the winning parts of the parent companies would presumably have performed even better. Alternatively, one could form the portfolio by investing $1,000 in each whole company winner and for winning subsidiaries investing $1,000 times the percent of the whole company employment taken by the winning subsidiary. NIST actually did this experiment and others (NIST, Feb. 5, 1996) and for each MBNQA portfolio constructed compared the results with portfolios formed simultaneously by investing identical dollar amounts at the identical times in the Standard & Poor's (S&P) 500. As of August 1, 1995:

> The 14 publicly-traded winners outperformed the S&P 500 by over 4 to 1, achieving a 248.7% return compared to a 58.5% return for the S&P 500. The 5 whole company winners outperformed the S&P 500 by greater than 5 to 1, achieving a 279.8% return compared to a 55.7% return for the S&P 500.

The experiment was repeated for the 41 publicly-traded applicants that received site visits as a part of the Award evaluation process. These applicants also outperformed the S&P 500, in their case by greater than two to one. Updating the study yet again through December 2, 1996, similar results were found (NIST, 1997). When updated through 1999, the stocks of publicly-traded U.S. companies that have received the Award outperformed the S&P 500 nearly 5 to 1 (NIST 2000).

The Hendricks and Singhal studies were broader in their scope. Based on a sample of nearly 600 companies that won national, state, or local quality awards, and a control sample of firms that did not win such awards, Hendricks and Singhal concluded that award winners out performed the control group's performance, where performance was quantified as the growth in stock price, employment, sales, operating income, total assets, and return on sales. Hendricks and Singhal also showed that small award-winning firms out performed large award-winning firms in these dimensions.

The Council on Competitiveness (1995) report on the Baldrige National Quality Program reaches conclusions consistent with the belief that firms use the Baldrige Award Criteria to form their own performance management programs and consistent with the hypothesis that such programs have significantly improved the competitiveness and performance of U.S. companies. The Council (1995) observes:

> The Baldrige National Quality Award and its state and local offshoots have been key to the effort to strengthen U.S. competitiveness. The annual government investment . . . in this program is leveraged by . . . private sector contributions. The impact of the Baldrige Award on the competitiveness of U.S. industry and the dividends it pays to the U.S. economy far exceed these investments.

Black and Porter (1996) try to identify the factors that are critical to the success of quality management and thereby try to improve the usefulness of the Baldrige Criteria as a self-assessment framework for firms using them to implement their own quality management programs. Black and Porter (1996) extracted 32 items from the Baldrige Criteria for the nonresults categories and added another seven items to cover issues that they believed were not adequately covered in the framework. The resulting 39 items were then listed in a questionnaire sent to a sample of European managers. Black and Porter then used factor analysis to identify ten critical factors of TQM. Black and Porter (1996, pp. 20-21) developed the following descriptive labels for the ten factors: (i) people and customer management, (ii) supplier partnerships, (iii) communication of improvement information, (iv) customer satisfaction orientation, (v) external interface management, (vi) strategic quality management, (vii) teamwork structures for improvement, (viii) operational quality planning, (ix) quality improvement measurement systems, and (x) corporate quality culture. Of course, the descriptive labels are attempts to summarize the import of the several questionnaire items associated with each factor, and interested readers should consult Black and Porter directly for the list of items appearing in each of their critical factors for TQM.

Like Black and Porter (1996), Ahire, Golhar, and Waller (1996) conclude that the large number of quality management strategies need to be integrated into a holistic, well-integrated strategy. Because agreement on such a holistic quality management policy does not exist, they first analyze the quality management literature to identify 12 key constructs of integrated quality management strategies. They survey manufacturing firms to test the constructs empirically and validate their usefulness, comparing their own formulation with other comprehensive approaches to TQM. Ahire, Golhar, and Waller (1996) provide an excellent review of several of earlier studies (apart from the important Black and Porter study which appeared simultaneously with their own and which is discussed above) that attempt to identify the key constructs of a holistic approach to quality management. As they observe, the earlier studies complement their own, but the Black and Porter (1996) and Ahire, Golhar, and Waller (1996) studies draw on a larger body of literature to develop their quality management constructs to be tested.

The twelve constructs developed from the quality management literature by Ahire, Golhar, and Waller (1996) bear a strong family resemblance to the ideas encompassed in the Baldrige Criteria. Indeed, after studying the details of the seven criteria (NIST, 1997) the twelve constructs can readily be assigned to various parts of the seven criteria. Indeed, the criteria are frequently among the sources

cited in support of each of the Ahire, Golhar, and Waller constructs, and Ahire, Golhar, and Waller conclude their presentation of their twelve constructs by observing (p. 34): "These constructs span the entire range of activities deemed critical by the Malcolm Baldrige Award." That may be an overstatement; certainly the Baldrige Criteria are more complex in their comprehensive details and amenability to comprehensive evaluation of a firm's holistic or total quality management tailored to the firm's own situation and needs—especially when the firm using the criteria to create its own quality management policies has actually applied for the Baldrige Award and undergoes the entire process—described by Reimann and Hertz (1993, p. 44) and NIST (1997)—of site visits by teams of examiners and the comprehensive reports providing feedback after a panel of judges reviews the site visit reports. However, the Ahire, Golhar, and Waller constructs can be fit into the framework implicit within the Baldrige Criteria, even though the converse would—because of the comprehensiveness of the MBNQA criteria and interactive process—probably not be completely convincing. The twelve constructs are:

1. Top Management Commitment: Ahire, Golhar, and Waller (p. 27) consider and cite fifteen sources from the quality management literature to summarize with six items—one, for example, being management's allocation of adequate resources to quality improvements—the commitment of top management to quality.
2. Customer Focus: Ahire, Golhar, and Waller (pp. 27-28) use ten sources to develop four items—one, for example, being the availability of customer complaint information to managers—to capture the customer focus of a company's quality management.
3. Supplier Quality Management: Ahire, Golhar, and Waller (p. 28) develop a six-item scale—including, for example, the consideration of the supplier's delivery performance—to represent the effectiveness of management of supplier quality from seven sources.
4. Design Quality Management: Ahire, Golhar, and Waller (p. 29) use eleven sources to develop six items—for example, one of the items is emphasis on the design team's marketing experience—to evaluate a company's management of design quality.
5. Benchmarking: Citing just two sources, but observing that benchmarking, the use in quality management of analysis of best practices—products and processes—of leading competitors, has been extensively discussed in the quality management literature, Ahire, Golhar, and Waller (pp. 29-30) develop five items—including, for example, the emphasis the company places on benchmarking competitors products and processes—to measure a company's use of benchmarking.
6. Statistical Process Control (SPC) Usage: Seven sources from the quality management literature are cited by Ahire, Golhar, and Waller (p. 30) in support of their four-item scale to assess a company's use of SPC. One of the items, for example, is the extent of production employees' knowledge of SPC tools.
7. Internal Quality Information Usage: To measure the effectiveness of the use of information internally, Ahire, Golhar, and Waller (pp. 30-31) develop six items, citing five sources in the quality management literature to support their choice. One item, for example, is the availability for managers of data about the cost of quality.
8. Employee Empowerment: Four sources in the quality management literature are cited by Ahire, Golhar, and Waller (p. 31) to support their choice of five items to measure the extent of employee empowerment for a company. For example, one item is the extent to which workers are encouraged to find and fix problems.

9. Employee Involvement: Ahire, Golhar, and Waller (pp. 31-32) cite twelve sources to support their choice of eight items to represent the extent of a company's strategy for employee involvement. For example, the availability of profit-sharing programs is one of the items.
10. Employee Training: The quality management literature emphasizes the importance of training employees in quality management, and Ahire, Golhar, and Waller (p. 32) cite eleven sources to support their choice of five items—for example, one of the items is the availability of resources for training—to measure a company's strategy for training employees.
11. Product Quality: Ahire, Golhar, and Waller (pp. 32-33) review Garvin's (1987) rather comprehensive definition of product quality as encompassing performance, features, conformance, reliability, durability, serviceability, aesthetics, and perceived quality. Because their survey of quality management strategies focused solely on the automotive components manufacturing industry (SIC 3714), not all of Garvin's quality attributes were relevant, and Ahire, Golhar, and Waller use just four of his eight items and added two more of their own, citing four sources from the QM literature in addition to Garvin to support their choices.
12. Supplier Performance: Ahire, Golhar, and Waller (p. 34) measure supplier performance with six items, supporting their choices with references to five sources in the quality management literature. For example, among the items is the willingness of suppliers to improve quality.

Ahire, Golhar, and Waller empirically validate their twelve quality management constructs with a survey of manufacturers of automotive components, having the respondents evaluate the importance of the various items for a company's quality management by using a 7-point Likert scale to assess each item. For one example, with the fifth construct, benchmarking, the respondent evaluated five items, one of which was "We are engaged in extensive benchmarking of competitors' products that are similar to our primary product," by using a 7-point scale (1 = strongly disagree, 4 = neutral, 7 = strongly agree). Ahire, Golhar, and Waller used confirmatory factor analysis to refine and validate their construct scales and then estimated correlations among the twelve constructs that they had developed from their review of the quality management literature. Correlations among the constructs are all positive, which Ahire, Golhar, and Waller (p. 41) believe ". . . supports the notion that the quality management strategies should be implemented holistically rather than piecemeal. Many of these constructs exhibit synergy with one another."

Among other suggestions for future research, Ahire, Golhar, and Waller (p. 47) observe: ". . . the constructs developed here could be used in subsequent empirical research on integrated quality management strategies to develop and test causal models of quality management implementation effectiveness." Lau (1996, p. 6) also suggests the use of survey data, about the importance of various aspects of quality management, to test hypotheses about how various human and technical factors affect quality and about how quality management affects a firm's profitability and quality.

Lau (1996) has developed descriptive data about manufacturing firms in computer and electronics industries (SIC 357 and 367) and proposes to use the data for hypothesis testing about the effectiveness of various aspects of quality management. These data thus provide information about manufacturers in different industries other than the Ahire, Golhar, and Waller sample, and Lau's 7-point Likert scale survey covered the importance of a somewhat different set of factors that potentially affect

quality. Lau has developed information about the respondents' assessment of their firm's competitive environment (six items, including for example the possibility or threat of new competition), factors affecting competitiveness (nine items, including for example innovative designs), workforce (ten items, including for example how much discretion line workers have over the pace of their work), organization (ten items, including for example the clarity of the division of responsibilities in the company), manufacturing technologies (ten items, including for example the extent to which TQM has been implemented), flexibility (ten items, including for example the company's ability to develop or modify new product designs), product quality (ten items—again largely following Garvin, 1987), and comparative performance (ten items, including for example profitability).

Lawler, Mohrman, and Ledford (1995) report that in multiple regression analyses financial performance measures such as the return on sales, investment, and equity were significantly related to the extent that employee involvement (EI) practices and TQM were used by the companies surveyed, although Konczak (1996) in reviewing the book criticizes the fact that the technical details of the regression analysis to support the report were not actually provided. Lawler, Mohrman, and Ledford provide the third in a series of reports for a project sponsored by the Association for Quality and Participation. The series documents the EI and TQM practices in Fortune 1000 companies. The 1995 work describes the survey results for 1993, but also includes data collected in the first two surveys in 1987 and 1990. Konczak (1996, pp. 497-498) provides a concise overview of the findings from the surveys: all of the EI practices—information sharing, knowledge development, reward systems, and power sharing—appear to be important; about five-sixths of the respondents reported an increase in the use of TQM practices since the 1990 survey; most respondents believe that EI and TQM practices improve performance; and—although Konczak (1996, p.498) observes that the information is less convincing because detailed statistical support is not shown—the practices appear to improve performance somewhat, explaining a statistically significant but small amount of the variance in performance. Konczak (p. 498), continuing with the useful overview of the work, believes that the LML data characterizing the types of companies and their business environments "confirm what most readers have probably learned from their own experiences (e.g., EI and TQM are likely to be adopted when an organization faces tough competitive pressure)." The Lawler, Mohrman, and Ledford monograph, which Konczak (p. 498) considers "the best source currently available that provides an inside view (i.e., the senior management point of view) of employee involvement and total quality management practices in large organizations," concludes with a look at the future of quality management practices and finds that an increase in investments in quality management is planned by most firms.

Like the quality management literature, the economics literature has recently begun to provide some evidence to support the importance of quality management for firm performance. One theme that emerges is the importance of deploying systems of quality management practices—coherent policies that combine many individual practices into a holistic focus on improving quality performance. That message from the economics literature's empirical contributions to the study of management decisions to improve quality is in accord with the findings in the quality management literature itself.

Ichniowski, Shaw, and Prennushi (1995 at http://www.nber.org) emphasize the need for packaging complementary quality management tools to improve human resource policy:

> Increasingly, firms are considering the adoption of new work practices, such as problem-solving teams, enhanced communication with workers, employment security, flexibility in job assignments, training workers for multiple jobs, and greater reliance on incentive pay. This paper provides empirical evidence to address the question: do these human resource management practices improve worker productivity? For this study, we constructed our own data base through personal site visits to 26 steel plants which contained one specific steelmaking process, and collected longitudinal data with precise measures on productivity, work practices, and the technology in these production lines. The empirical results consistently support the following conclusion: the adoption of a coherent system of these new work practices, including work teams, flexible job assignments, employment security, training in multiple jobs, and extensive reliance on incentive pay, produces substantially higher levels of productivity than do more traditional approaches involving narrow job definitions, strict work rules, and hourly pay with close supervision. In contrast, adopting individual work practice innovations in isolation has no effect on productivity. We interpret this evidence as support for recent theoretical models which stress the importance of complementarities among a firm's work practices.

Ichniowski, Shaw, and Prennushi (1995) focus on a single steelmaking process. A state of the art study using data across "a large nationally representative sample of manufacturing businesses"—the Educational Quality of the Workforce National Employers Survey (EQW-NES)—is provided by Black and Lynch (1996b) who link the EQW-NES with the Bureau of the Census' Longitudinal Research Database (LRD) and extend the productivity analysis that economists have often used to study the effects of capital investments and R&D to explore the effects of workplace practices, human capital, and information technology on establishment productivity. "More specifically, … [Black and Lynch] examine how workplace practices, human capital investments, and the diffusion of information technology explain the unobserved employer 'fixed effect' in standard production functions that do not control for these types of factors" (Black and Lynch, 1996b, pp. 2-3). Again, the importance of coherent, holistic quality management efforts is supported in the sense that the evidence suggests an effective quality management policy is built up from attention to meaningful, well-integrated practices throughout the workplace. Black and Lynch find that workplace practices do affect productivity, and the key is in how the practices are implemented. For example, they find that it is not enough to simply adopt a TQM system—that alone has an insignificant affect on productivity in their estimations. However, productivity is significantly increased when a larger proportion of a plant's workers are involved in decision making, a workplace practice that an effective TQM system incorporates through regular meetings involving the plant's workers in the decision making process. Black and Lynch find that investing in the human capital, especially education, of workers and diffusing the usage of computers among non-managerial employees increases labor productivity significantly.

Thus, the economics literature provides a direct way to model empirically firms' investments in quality—firm or establishment productivity is explored in the framework of the economics literature's production functions augmented with measures of workplace practices and investments in human capital. Black

and Lynch provide a review of the earlier efforts in that direction and then provide an improved analysis using the Educational Quality of the Workforce National Employers Survey matched with the Bureau of the Census Longitudinal Research Database to provide evidence in both cross section and panel data. Especially noteworthy among the earlier works reviewed by Black and Lynch (1996b) are Huselid and Becker (1996) and Black and Lynch (1996a). Huselid and Becker find that a firm's market value is higher by about $16,000 per employee if the Huselid/Becker index of a firm's human resource systems is one standard deviation higher. Black and Lynch (1996a), as evaluated in their own 1996b (pp. 8-9) review of the literature:

> … examine directly the impact of education and training on establishment productivity in both the manufacturing and non-manufacturing sectors … [and] conclude that education raises productivity approximately 5-13 percent, depending on the sector. In addition we find that the impact of training investments by employers differed depending on their nature, timing and location. More specifically, we find that formal training outside working hours has a positive effect on productivity in manufacturing businesses, while computer training raises the productivity of non-manufacturing establishments considerably.

Empirical models of quality investment that focus on investments in product attributes

Less direct evidence about investments in quality abounds in the economics literature. The distinction, though, is that the studies reviewed above are directly concerned with measuring the effects of the various managerial policies that make up TQM, whereas the vast amount of empirical work in the economics literature that is indirectly related to quality investments (i) does not consider managerial policies explicitly and (ii) does not focus on the measurement of the effects on performance of variance in such quality management policies across firms. These studies are largely focused on investments in product attributes, construed broadly to include R&D to improve the production process for producing the product as well as R&D to develop the product itself.

Perhaps most prominent among such indirect studies of investments in quality are the studies of investments in R&D and studies of the effects of such investments on productivity. The reviews of Baldwin and Scott (1987), Cohen and Levin (1989), and Link (1987) provide overviews of the earlier economics literature on R&D and productivity; clearly the focus in the economics literature until the recent studies in the wake of the new interests in quality management beginning in the late 1980s has not been on measuring the effects of differing managerial approaches on the effectiveness of R&D, but rather on how the structure of markets (e.g., the degree of competition in a market) and the broad characteristics of firms (e.g., the size of a firm) affected R&D investment and on how such investment affected the productivity of firms and industries and even entire economies. Scherer and Ross (1990) and Martin (1993) are good sources of overviews of the focus of economists' concerns as they have conducted empirical work on R&D investments and more generally on investments in products and processes—including investments in product differentiation other than R&D investments—that firms make to improve their performance. Of course, the recent quality management literature reviewed in

this report cites, and has built upon, studies that combine the insights in the economics literature with a focus on management issues. Among the first studies to exploit the complementarities between the economics and managerial literatures and to combine economists' knowledge about R&D investments with recommendations for managerial policies increasing international competitiveness of firms was Link and Tassey (1987). There is, then, underlying the recent quality management literature, a literature that has brought the knowledge found in the economics literature to bear on managerial issues. Link and Tassey (1987) review much of that literature, and the recent quality management literature reviewed above reviews much of it, too. This review turns now to an update on such indirect studies of quality investments in the recent economics literature.

The recent empirical economics literature that indirectly explores quality investments has focused on the "quality ladder" model and on product differentiation in the context of international competitiveness—focuses pertinent to the present report, although, nonetheless the work is indirect in the sense that unlike the recent literature reviewed and advanced by Black and Lynch (1996b), the studies are not evaluating the effects of the various types of TQM policies.

Caballero and Jaffe (1993) develop and estimate a model in the spirit of the "quality ladder" theory discussed briefly in the theoretical literature reviewed above. They are able to develop several important facts about the diffusion of ideas, the obsolescence of knowledge, the amount of knowledge embodied in patents at various times over the twentieth century, the spillovers of knowledge generated by patents at various times, and the relative size of the public knowledge stock at various times during the century. Regarding an especially noteworthy fact in the context of the present report, Caballero and Jaffe (1993, p. 17) observe:

> [W]e use market value and patents data on 567 large U.S. firms. The data are annual for the period 1965-1981, and the firms are assigned to 21 technological sectors. We estimate 21 sectoral panels and find that, on average (over time and sectors), creative destruction is about 4% per year. That is, in an average sector in an average year a firm that does not invent sees its value relative to that of the industry erode by about 4%. This number varies widely across sectors; drugs have the largest average creative destruction, with about 25% per year.

Knowing such facts is obviously important for managers; the knowledge conditions the importance of managing to achieve innovations. Clearly though, the knowledge is not knowledge about *how* to manage successfully, so again, this part of the economics literature about quality investments is of indirect importance to managers but directly important to evaluation of the Malcolm Baldrige National Quality Award. What such economic literature about quality investments provides to the evaluation is the knowledge that one must control for differences across firms and industries in performance that would occur holding constant the managerial efforts. Thus, managers pursuing similar strategies with similar vigor might have different success maintaining or growing their firms' market values simply because the rate of creative destruction for some industries is greater than for others.

Bresnahan (1981) estimates for the automobile industry the product-quality differentiation model described in the preceding section. He uses automobile characteristics such as the number of cylinders, horsepower, and miles per gallon to proxy for product quality, and he finds that price-cost margins are much smaller for lower-quality vehicles than for the higher-quality vehicles. He also finds the quality downshifting phenomenon (described in the preceding theory section) to be important. Bresnahan (1987) uses the model to compare the evidence for mid-1950s pricing in the automobile industry with the predictions of the model given alternative assumptions about the cooperative versus noncooperative behavior of automobile manufacturers. Feenstra and Levinsohn (1995) provide estimates for their model that was described in the preceding section; they use a sample including new car models with substantial sales in 1987 in the United States. The estimation allows, for example, an inference about whether the various automobile firms set price or quantity. As discussed in the preceding section reviewing theory, these product-quality differentiation models are important and yield important results, but they are not directly important for the quality management issues that are the focus of the MBNQA. Perhaps with time these models, and the other models in the "new hedonics" literature that were discussed in the preceding section, will evolve, allowing the introduction of the realism necessary to inform managerial issues more directly. However, the essential point, at the present stage of the evolution of the economics and quality management literatures, is that the two literatures are addressing fundamentally rather different issues.

As noted earlier at the end of the review of the theoretical literature, the economics literature on product-quality differentiation offers insights that are surely of more than passing interest to managers—insights about price-cost margins and competition. The usefulness of the economics literature to those seeking to verify the importance of various quality management policies or to those seeking to develop such policies is in providing an understanding of the conditions affecting performance apart from management. Note too, that the product-quality differentiation economics literature addresses just product quality, which is just one aspect of the holistic, total quality management set of issues. More to the point, as noted earlier the literature is not about which managerial tools can improve performance, but rather predicts pricing performance conditional on a variety of assumptions about firms, customers, and markets. As such, it can of course inform managerial decisions (Porter, 1980; 1985).