Abstract

The establishment of a national system of federal-state manufacturing modernization centers to serve small- and medium-sized manufacturing systems is a major experiment in the US. The system’s rapid expansion has made it impossible to begin with an experimental design that would systematically permit tests of the variables and relationships likely to affect the long-term economic effectiveness of these centers or their political and financial stability. Relatedly, many existing evaluation studies and journalistic narratives are marred by serious analytical and empirical flaws. It is not too late to attempt to improve practice in future evaluations. Several areas of expanded evaluation are described, including theories of industrial restructuring, regional technology infrastructure, public management, financing, evaluation design, evaluability assessment, firm learning, and measurement of benefits and costs. In each case, evaluation is presented as a means of simultaneously yielding information relevant to formative and summative program-level decisions and to hypothesis testing.

1. Introduction

This article is written from a two-fold vantage point: first, from the recognition that significant improvements are both needed and feasible (albeit at a cost) in the evaluation of federal and state manufacturing modernization programs; and second, from experiences in the difficulties of implementing formal evaluation designs, especially those employing experimental and quasi-experimental formats. From these vantage points we (1) propose an expanded framework for evaluating manufacturing modernization programs; (2) present a summary assessment and analysis of the current state of knowledge; and (3) outline our perspective on directions of future research.

2. The purposes of evaluation: an expanded perspective

Why conduct evaluations? The purposes of program evaluation, according to Goldenberg, are to learn about a program’s operations and effects, to control the behavior of those responsible for program implementation, and to influence the
responses of outsiders in the program’s political environment (Goldenberg, 1983, p. 515). The first element contains the traditional categories of formative and summative evaluations; the second, elements of program accountability and project milestones; and the third, both the substantive documentation of program accomplishments and the control of political symbols ².

We propose two additional interrelated purposes: to test the hypotheses contained within the emerging national commitment to manufacturing modernization centers, and to use evaluation-based findings on the workings and impacts of these centers to test related hypotheses about structural changes in American manufacturing, processes of technological innovation, regional economic growth, and firm learning. Justification for these additional purposes arises in light of substantial increases in the number and scale of manufacturing modernization programs. From approximately 42 programs in 28 states with total expenditures from all sources of $83 million (Clarke and Dobson, 1991), the number of programs has grown rapidly following enactment of the Omnibus Trade and Competitiveness Act (which established the Manufacturing Technology Centers (MTC) program), and the implementation of the Technology Reinvestment Program in 1994 (which provided an initial infusion of approximately $200 million in matching funds for technology deployment). In effect, the United States is engaged in a major experiment to establish a national system of federal-state manufacturing modernization centers that will be accessible to most of the nation’s small- and medium-sized manufacturing firms.

The rapidity of the system’s expansion has made it impossible to begin with an experimental design that would systematically permit tests of the variables and relationships likely to determine the effectiveness and efficiency of center operations or their longer term political and financial stability. These variables include core delivery strategies, range of services, characteristics of host institutions, interorganizational relationships between host institutions and other service delivery providers, industry and firm characteristics within a service delivery area, existence of alternative (possibly competitive) private-sector and public-sector service providers, and funding strategies. Each of these variables has been identified in one form or another in existing reviews of the MTC and related programs as a significant determinant of center effectiveness. However, the degree of evidence varies for these potential causes, many of which are confounded with others. It is not readily apparent how these variables interact to determine project outcomes, much less program design.

It is likely to be too late to incorporate systematic variation in these variables in the construction of the emerging national system, but it is not too late to attempt to incorporate attention to these factors in subsequent evaluations. At issue is the need not simply to know that projects (centers) have succeeded or failed, but to know why. If we do not know why programs have failed or succeeded, we run the risk of attributing success or failure incorrectly. This could lead to incorrect conclusions about generalizable best-practices, to mistaken program-level decisions about the future shape of the system, and correspondingly, to a likely ineffective or inefficient use of program resources.

This type of risk may be illustrated by manufacturing modernization programs that are concerned with low usage rates for existing ‘best-practice’ technologies and slow adoption rates for new technologies by small- and medium-sized manufacturing firms. Evidence clearly attests to this pattern (Office of Technology Assessment, 1990), but what precisely is its economic import? Several factors that determine the rate and extent of usage of existing or new production techniques have been identified (e.g. Mansfield, 1968; Rosenberg, 1972; Tornatzky et al., 1983). Among these factors are the following: (a) diffusion occurs over a period of time; (b) rates and extent of adoption depend on the characteristics of technologies as well as of firms; and (c) a firm's

² See, for example, Sapolsky's (1972) description of the cosmetic rather than substantive use of Program Evaluation and Review Techniques (PERT) in the development of the Polaris submarine.
decision either to wait (lag) or to reject use of a technology adopted by others may be economically rational behavior.

Consider the relationship of these propositions to the issue of identifying the impacts of industrial modernization programs. Many, if not all, of the specific technical improvements being highlighted by modernization programs (adoption of CAD/CAM, compliance with ISO 9000 standards, total quality management) are part of the general ethos of making things better; they are, so to speak, in the air. These innovations are being promoted by many sources. As in conventional diffusion dynamics, their adoption is made more attractive over time by reductions in technical and economic risks, and more necessary by competitive pressures. Conversely, adoption of these specific innovations by non-adopters may be made less attractive over time by the appearance of yet other functionally equivalent, economically more profitable innovations. Given this diffusion milieu, a firm's acceptance or rejection of the services of a modernization center may or may not be economically efficient. Given this context, by what standards do we gauge the effectiveness of a program's activities?

To raise the level of analysis another notch, modernization programs have come to learn that their objectives are to change firm behaviors, not simply production techniques. This broader understanding has been accompanied by a new problem, however. Much of the generic needs assessment of small- and medium-sized manufacturing firms that has served to build political support for manufacturing modernization programs relates to the symptoms of poor competitive behavior, not to an understanding of internal firm decisionmaking processes. In effect, centers work only with that relatively small percentage of firms defined as willing and able to accept assistance. As noted by the Manufacturing Technology Centers Third Year Review Panel (Manufacturing Technology Centers – Broad Programmatic Issues, 1992, p. 9).

The Manufacturing Technology Centers must cope with the fact that on the order of 70% of small- and medium-sized manufacturing companies are reported to be unable, unwilling, or unprepared to adopt the technologies and practices that would enhance their competitiveness. Of the remaining 30%, approximately two-thirds are willing but unable to adopt new technology; and one-third are both willing and able to do whatever it takes to enhance their competitiveness using advanced technology.

Nevertheless, the gap between the size of the population used to justify modernization programs and the population(s) actually served is too great to be acceptable in the long run. MTCs may still be able, in the aggregate, to produce noticeable improvements in the performance of selected segments of American manufacturing.

Distillation of best-practice modernization programs notwithstanding, few studies have focused on the decisionmaking processes within firms that lead them to be receptive or indifferent to the overtures of manufacturing modernization services or related offers of improvement. More to the point, little of the information needed to change this situation is likely to be generated by conventional evaluations, unless programs were designed with these very issues in mind. For example, how do variations in service delivery organization and strategies affect the responses of firms to the solutions offered to them?

3. The current state of practice

Time series of the number of evaluations of federal and state technology development and manufacturing modernization programs do not exist. There is little doubt, however, that the pace of activity has quickened. In 1994, the National Science Foundation (NSF) initiated evaluations of its Engineering Research Centers (ERC) and Science Technology Centers (STC) programs, while the National Institute of Standards and Technology (NIST) accelerated development of a sizeable intramural and extramural evaluation effort for the rapidly expanding MTC program. New York State has recently commissioned evaluations of both its Centers for Advanced Technology program (SRI International, 1992) and its
Industrial Extension Service program (Nexus Associates, Inc., 1993). These activities follow a series of external assessments of Ohio's Thomas Edison program (National Research Council, 1990; Mt. Auburn Associates, 1992), internal studies and surveys of user satisfaction for Pennsylvania's Ben Franklin Partnership Program and Industrial Resource Centers program (Ben Franklin Partnership Program, Strategic Investment Committee, n.d.; Pennsylvania Department of Commerce, 1993), and executive and legislative branch scrutiny of a number of other state programs. Much of this activity relates to university–industry–government programs designed to foster technological innovation. However, most of the same political, methodological, and empirical issues surrounding evaluation design and research of these programs also relate to manufacturing modernization centers. Thus, this is an appropriate time to take stock of the current state of practice in the evaluation of manufacturing modernization/industrial extension programs.

First, as detailed in earlier assessments, existing evaluation studies and related journalistic narratives are marred by serious analytical and empirical flaws (Feller, 1988; Shapira et al., 1993a). These flaws are related to casual treatment of numerous methodological concerns, such as a tendency to measure “outcomes that are only incidental and that do not represent truly targeted problems” (Mohr, 1988, p. 19). They are also linked to the casual use of analytical concepts and empirical evidence, particularly those concepts related to causal inference and to estimating program benefits and costs. These flaws are at times based on faddish or contested theories of state economic development (such as the role of small firms as sources of net job creation), which are used to gauge a program’s strategy and activities (Harrison, 1994). Most evaluations of industrial extension programs are process-oriented. Their intention is to provide program managers and center directors with information about the means for improving performance. Such evaluations, however, contain few measures of effectiveness or efficiency. Whatever their objective, most existing evaluation studies incorporate few if any safeguards against standard threats to internal or external validity. They typically lack baseline measures and even more frequently lack comparison or control groups. This limitation makes it difficult to determine whether the changes attributed to the intervention are due instead to other factors, such as trends in place prior to the intervention, cyclical changes in economywide or industrywide variables, or other historical events that also led to changes in firms that did not receive experimental treatment. Bias in selecting firms to highlight program ‘success’ is manifestly evident in many evaluations, especially in journalistic accounts. Vignettes on single firms represent only a small percentage of the experiences of firms that receive services from industrial extension programs (and an even smaller percentage of the number of firms that could or should have received services). Some studies continue to advance propositions found to be counterproductive and subsequently scrapped by existing programs. Given these and other flaws, existing studies may be imperfect if not misleading guides to policymaking, program design, and program operations. Collectively, they bring to mind this adage: “It ain’t what we don’t know that bothers me so much; it’s all the things we do know that ain’t so.”

Osborne and Gaebler, in *Reinventing Government*, for example, cite Pennsylvania's Ben Franklin Partnership Program's formula for allocating state funds among competing advanced technology centers as a “key innovation ... [that] forced each center to embrace the mission defined by the state — commercial development of technological innovation — and to push for the results the state wanted private sector investment and job creation in Pennsylvania” (1992, pp. 278–279). In fact, the formula was found to contain counterproductive incentives and has been scrapped (Ben Franklin Partnership Program, Strategic Investment Committee, n.d.).

The state-of-the-art has retrogressed in some respects. A.D. Little's evaluation of the State Technical Services (STS) program in 1969 dealt with a program that was a precursor to contemporary industrial modernization programs. This evaluation addressed many of the problems in estimating economic outcomes that beset efforts to evaluate current programs. The fact that the A.D. Little report's generally favorable assessment of the STS program was followed by the program's demise should serve as a telling reminder in all exegesis on evaluation design — do not confuse the intrinsic quality of a study with the wisdom of subsequent policy decisions (Roesner, 1989).
Second, stock-taking may overcome the difficulty experienced by researchers in benefiting from the state-of-the-art because its contents are unknown or inaccessible. Evaluation studies of federal and state technology development and deployment programs constitute a fugitive literature (Shapira et al., 1993b). Each evaluation study has essentially been a one-time, idiosyncratic event. Cumulative learning and improvement (in which the methodological approaches of one study contribute to improved techniques in subsequent studies) have essentially been absent from these studies. The lack of comparability among studies also has limited their utility as ‘benchmarks’, however restricted or tentative.

Third, to state the obvious, the well-known difficulties of evaluating technology development and manufacturing modernization programs will not be lessened unless they are tackled aggressively. These difficulties include: (a) operational ambiguities in key program objectives, such as economic competitiveness; (b) multiple, not necessarily consistent, program objectives (e.g. productivity increases and/or job retention) that readily lead to selective reporting of impacts; (c) lengthy, often circuitous processes among the research, development, and dissemination activities supported by these programs that make it difficult to link program inputs and outputs; (d) the difficulty in implementing rigorous research designs in many settings; (e) the societally and spatially dispersed character of benefits that involve participants other than those directly involved in a program's activities; (f) the proprietary nature of data needed to measure final outcomes; and (g) the lack of accessibility within sites and comparability across sites of management information systems (Watkins and Wills, 1986; Government-University-Industry Research Roundtable, 1987; Feller, 1988).

It would be easy to attribute the flawed quality of existing studies solely to inherent conceptual difficulties or difficulties in obtaining data; however, it would be wrong to do so. Evaluations in many program areas deal concurrently with similar issues relating to the specification of objectives, construct validity, research design, and empirical estimation. Yet in many of these areas (substance abuse prevention, manpower training, psychotherapy, medicine, education) a substantial (albeit not necessarily convergent) evaluation research literature, tradition, and practitioner community exist. In the case of state technology development and industrial extension programs, however, at least until recently the ability to repeat the litany of formidable barriers has served as a convenient rationale for not conducting evaluations or as an umbrella justification for limited expectations or egregious shortcomings.

Recently, improvements in the quality of evaluation studies as well as a more modest set of claims for the stability or generalizability of findings are indeed evident. ‘Lessons’ drawn from both national surveys and case studies of the experiences of state-supported modernization programs offer a workable set of ‘best-practice’ project guidelines (Shapira, 1990; Clarke and Dobson, 1991), as well as an informative guide to the set of issues that should be addressed in national evaluations. The recent evaluations of New York’s Centers for Advanced Technology and Industrial Extension Service programs have made extensive use of economic analysis to specify program benefits and costs (Feller and Anderson, 1994). The Nexus evaluation of the Industrial Technology Extension Service (ITES) program used firm-based employment data organized by SIC code to construct comparison groups for the firms serviced by ITES. (N: See article by Oldsman, this issue.) 5 The Midwest Manufacturing Technology Center (MMTC) explicitly introduced a comparison group into its internal impact assessments. (MMTC: See article by Luria and Wiarda, this issue.) The MMTC is attempting to

---

5 The importance of comparison groups in evaluation programs is illustrated by the following scenario posed in the Nexus study: Firms that received ITES assistance may have grown more rapidly (in sales, employment, etc.) than the average for all New York firms over some period of time, but if the ITES-serviced firms were in industrial sectors that had above-average growth rates, it would not follow that the program, otherwise seemingly successful, was producing incremental change. Such a finding also would raise questions concerning whether programs were producing net benefits or simply astutely picking winners with which to work.
measure whether its customers move away from their baselines faster than non-customer panel members on measures related to the substance of their MMTC engagements.

Several evaluation initiatives also have been undertaken by NIST's Manufacturing Extension Partnership (MEP) program. MEP, in cooperation with the Manufacturing Technology Centers, has taken actions to develop a mutually agreed-upon set of common performance measures that will permit subsequent comparative analysis. NIST and the Manufacturing Modernization Forum have sponsored preparation of a guide to evaluation that addresses many of the design issues raised in this paper (Nexus Associates, Inc., 1994a,b). A manufacturing modernization evaluation community encompassing NIST, MTCs, state participants, and multiple suppliers of evaluation expertise is taking shape, largely through activities sponsored by NIST (Shapira et al., 1994). The development of such a community, called by Campbell (1984) a "contentious community of scholars", is a key ingredient for checks on quality and for further advancement in practice.

Such improvements suggest that the quality of evaluation studies to date has been constrained by demand as much as by technology. Demand is used here both in the general sense (the evaluation literature speaks of the different purposes of evaluations), and the specific sense (economists speak of the quantity of a (quality-adjusted) good that will be purchased at a given price). Simply put, until recently, many governors and program managers had been interested in (or at least content with) demonstrating the value of their entrepreneurial initiatives without opening their programs to independent, third-party scrutiny, or having to invest in the resolution of data or technical issues.

What then determines the demand for high-quality evaluation studies? A political product life cycle seems to enter here. In the formative, early years of new programs, aspirations, promises, and expectations, salted with selected anecdotes, typically carry the day. As time passes, political sponsors leave office and legislatively mandated reviews take hold. Successor administrations and program managers are left with due bills to account for a program's cumulative impacts and to justify continued support. Evaluation is a means of responding to these demands. The rigor of the evaluation, in turn, appears to depend on the answer to another question: how good an answer is sufficient to meet the needs of the final consumer of the study's findings?

State programs are clearly facing more exacting scrutiny after approximately a decade of support. Federal managers of manufacturing modernization programs appear never to have had a lengthy first-stage honeymoon. The MTCs program began with legislatively mandated reviews after 3 years; this had to be accomplished before the second allotment of funding was disbursed. Similarly, external reviews have been required of the original three MTCs as they now seek extended federal funding beyond 6 years.

The accelerated pace of interest in systematic, third-party evaluation of technology development and deployment programs, and the willingness of agencies to both provide resources for the conduct of evaluations and demand higher standards of performance from their contractors, reflect a new systemic emphasis on accountability in all public-sector programs in a national period of constrained public expenditures and skepticism about the efficiency of public-sector institutions. At the federal level, it both reflects and anticipates the requirements upon agencies of the Government Performance and Results Act of 1993 to develop performance indicators and annual performance plans. Finally, at the national level, it

Indeed, the common characteristic of the evaluation agenda presented below is that it would entail studies that cost more, take longer, and involve more complex designs than are found in existing evaluation studies. In addition, much of the benefits would be in the form of improved methodology and data sets. Goods that generate positive externalities tend to be undersupplied because individual actors consider only private, not social, benefits. The potential benefits to the relevant policy or evaluation communities of more valid findings about a specific program's impacts may not be sufficient to induce an agency to invest additional funds to generate a rigorous design if a less rigorous design serves (satisfices) to provide the documentation required by legislative or executive authorization and/or appropriations committees.
anticipates NIST's heightened need following the 1994 mid-term elections to document accomplishments in programs that are redolent of industrial policy even while the MTC program is still cresting behind the support of an incumbent administration.

While reinforcing in many respects, the building pressure for evaluation at both federal and state government levels is not without its own complexities and latent tensions. The 'partnership' structure of MTCs involves a mix of federal and state funding that creates overlapping but not identifiable sources of demand for evaluation. As noted, state governments have been slow to date to evaluate manufacturing modernization centers, possibly relying on continued federal funding as an imprimatur of acceptable levels of MTC performance. But even if state governments begin to keep pace with NIST's interest in evaluating MTCs, differences could yet exist in the agendas (and possibly) the criteria used by them and federal agencies. If past and present practice is a guide, state evaluations of manufacturing modernization programs are likely to center on a few impact measures, and to ignore the set of variables (e.g. organizational structure, service delivery strategy) described above that are more broadly hypothesized to affect the performance of MTCs. Few states can be expected to view the benefits of incremental information garnered from complex evaluation designs as worth the cost.

Indeed, the decentralized, polymorphous, longitudinal character of the MTC program requires a national evaluation perspective. But any national or cross-site evaluations conducted may detect differences in the effectiveness or efficiency of program strategies that call into question the design or operation of a state program. In a related manner, state and federal government evaluation schemes might differ in the inclusion or weight accorded specific criteria. The impact measure/performance indicator 'bottom line' to a state might be job creation, whereas to a federal decisionmaker it might be a favorable benefit–cost ratio: the two measures are not the same. Given different evaluation questions, different findings would not be surprising. How and by whom, then, are these findings to be used? Would a state continue to fund an MTC whose operations had been found to be ineffective or inefficient according to a NIST-funded study? These are questions of intergovernmental relationships and national and state politics, not of evaluation design.

4. Directions for expanded evaluation

Several areas of expanded evaluation are immediately identifiable. In each case, we view evaluation as a means of simultaneously yielding information relevant to formative and summative program-level decisions and to hypothesis testing.

4.1. Industrial restructuring

On what basis will America's manufacturing activities be organized, and what will be the role of small- and medium-sized firms in this evolving system? Competing perspectives exist on the relative importance to be attached to innovations at the point of production, to flexible specialization, and to interfirm production networks (Piore and Sabel, 1984; Kenney and Florida, 1993). To what extent should the design of the MTC system or the operations of specific MTCs be shaped by the debate? For example, what purposeful stand should MTCs take in establishing priorities among clients, industries, or services? Should they continue, as at present, to largely serve those firms that seek (or finally accept) their services, or should they concentrate (and redouble) their efforts to serve 'foundation' firms, leaving other firms to adjust on their own to market forces? Given planned or natural variation in center strategies, evaluation would help gauge the effectiveness of pure or mixed strategies.

4.2. Regional technology infrastructure

The emerging federal-state system of MTCs, manufacturing outreach centers (MOCs), and state-supported industrial extension services has a spatial as well as a sectoral dimension; it is designed to improve the economic performance of regions as well as of firms. What is the nature of the potentially reciprocal relationship between
establishment of a manufacturing modernization center and other elements in a region's technology infrastructure (Feller, 1994)? Do regions with MTCs experience higher rates of economic growth than those without centers? Do regions without centers develop indigenous substitutes for federally funded centers?

4.3. Public management

A diverse set of organizations (universities, state agencies, not-for-profit organizations) serve as the host institutions for MTCs. A sizeable percentage of centers either operate directly as brokers (referring firms to other service providers), or rely on brokerage activities for a part of their services (Wyckoff and Tornatzky, 1988; Clarke and Dobson, 1991; Nexus Associates, Inc., 1994a, b). To what extent is the performance of centers affected by their degree of integration with other organizations that provide related modernization services, and with the host region? Are there systematic differences in the effectiveness of different sets of host institutions? For example, the early experiences of the Southeast Manufacturing Technology Center (which was managed through the University of South Carolina) and the Northeast Manufacturing Technology Center (which was managed through Rensselaer Polytechnic Institute) suggest that research universities may not be the most effective organizational home. In fact, management of each MTC has been shifted to other organizations. Yet are these two experiences determinative judgments of all universities, including those with established programs in technology transfer? Relatedly, given the importance of interorganizational linkages in the provision of modernization services, what are the most effective methods for coordinating activities between and among host institutions, cooperating organizations, and firms?

4.4. Finances

The original MTC legislation required the stepping down of federal support over a 6-year period, although recent changes in NIST's appropriations now permit extended federal funding upon favorable review. Federal expectations are still that MTCs will reduce their dependence on public-sector funds (which now constitute approximately 80–90% of their revenues) through increased use of fee-for-service. Not all MTCs, however, share the same view on the use of fee-for-service, some holding to a 'public service of public good' philosophy more akin to that historically associated with cooperative extension. Given these different views, what types of funding formulae should be negotiated between and among the federal government, state governments, firms, and other sources for the long-term support of centers? In the near term, what effects do current funding ratios among public and private sources of revenue have on center effectiveness? What is the elasticity of demand for center services given the imposition of fees-for-services?

4.5. Evaluation design

What is (are) the most effective method(s) for conducting evaluations of manufacturing modernization centers? The starting point for this assessment was quasi-experimental design, primarily because so many studies ignore the threats to internal and external validity, and because quasi-experimental design seems to be the most plausible and perhaps strongly feasible framework to minimize such threats. But other evaluation techniques could be used more effectively than they are at present. For example, randomized experiments in which a center or set of centers systematically varies the nature of services offered to clients are possible, at least in principle. Or, firms seeking services might be randomly assigned to different limits on the number of days of service they can receive, to assess the impact of service 'dosage'. While these or other randomized evaluations may not be easy to implement, experience in many other areas attests to their feasibility (Boruch and Wothke, 1995).

In addition, case studies are currently an underused and at times misused technique. Yin, for example, has contended that "the case study is the method of choice when the phenomenon under study is not readily distinguishable from its context" (1993, p. 3). This description would seem to apply to the diverse, decentralized character of industrial modernization programs. In fact, much
of what appears to be the findings of expert panels, as in the third-year and sixth-year reviews of the MTC program, are casually structured case histories or non-systematic surveys of non-systematically selected groups of firms.

4.6. Evaluability assessment

Increased requirements and expectations concerning methodological rigor are likely to quickly jar against the reality that many of the organizations that are new hosts of MTCs and MOCs likely have limited experiences with evaluation, as well as initially undefined and changing target populations and services. Evaluability assessment, a set of procedures designed to establish readiness for more systematic, summative evaluation, may be an important evaluation step in such cases. Evaluability assessment examines the structure of a program, examines whether some plausible sequence links the services delivered and the outcomes expected, and considers readiness for evaluation (Wholey, 1979; Rutman, 1980). The use of evaluability assessments also may reduce the likelihood that formal evaluations are conducted too early in the life history of a center or set of centers. However, evaluability assessment is not always an easy or welcome offer of assistance (Mark et al., 1994). It can involve criticism of existing site or agency evaluation procedures and the requirement that sites cooperate in providing records, data sets, and access to clients to determine the feasibility of fulfilling data requirements. These issues are magnified when efforts are made to conduct cross-site evaluations of independent and heterogeneous projects.

4.7. Firm learning

How is knowledge about firm learning to be incorporated into modernization programs? Firms learn as they interact with their environment acquiring information, acting on it, and embedding new practices into routine operations and organizational memory. To the extent that modernization and technical assistance programs approach firm problems primarily from the perspective of technological deficiencies rather than of organizational ineffectiveness, they run the risk of offering solutions to the wrong problem or of offering solutions in ways that do not correspond to firms' perceptions of their needs. Clearly, the existing situation in which approximately 90% of firms are not responsive to the activities of MTCs is not politically sustainable. Rather than being a summary judgment on firms' receptivity, this lack of responsiveness may reflect the limitations of dominant paradigms of what firms need and how to approach them.

4.8. Measurement of economic benefits (and costs)

How feasible are benefit-cost and related economic techniques for measuring effectiveness and efficiency in evaluating manufacturing modernization programs? How applicable are the estimation techniques used to compute private and social rates of return to technological innovations, and academic research to modernization programs (Mansfield et al., 1977; Mansfield, 1991)? What types of existing data sets might be used more creatively in constructing proxies for firm-based reports of benefits? Relatedly, how much may be asked of the firms receiving center services with regard to providing data? While initial answers to these questions may be tentative and imperfect, they nevertheless may pave the way for improved evaluation techniques. Without improvements in the measurement of economic impacts, attention will remain focused on intermediate indicators (e.g. scrap rates, adoption of new or existing technologies, sales) that are imperfectly linked to final outcome measures on the profitability and competitiveness of firms. Programs will be unable to answer increasingly more demanding questions on 'bottom line' impacts 7.

---

7 Pennsylvania's Ben Franklin Partnership Program has conducted an extensive review of the input and output characteristics of its centers and awards, devoting special attention to commercialization as an outcome measure. Although Ben Franklin centers report data on the number of products being sold in the market or new processes being implemented internally by the sponsoring firm, state officials recognize that the indicators lack economic significance because they may include product innovations for which only one unit is sold or process innovations that lead to few if any economic benefits (Ben Franklin Partnership Program, Strategic Investment Committee, n.d.).
Without better understanding of the microeconomics of firm behavior, the observed impacts of center activities may not be interpreted correctly. Certainly, the time for providing more than surveys of client satisfaction is rapidly approaching.

5. Conclusions

In our assessments of the impacts of technology development and manufacturing modernization programs, we have frequently encountered difficulties in meeting the methodological standards described and implied above. Thus, we are diffident about presenting these proposals as the only directions for the course of future evaluation endeavors. Instead, we offer a pragmatic approach to improvements in current practice. Although it is essential to aspire to and heed textbook maxims for the conduct of evaluations, craft and practice can be improved by learning-by-doing as well as by technical advances that shift outward the technological frontier of evaluation methodology. The standard for future action is not a single flawless study that satisfies all strictures, but rather a succession of studies that critique and improve upon each other as they collectively advance toward norms of formal evaluation methodology. Echoing Cronbach’s (1982) call for a “fleet of studies” rather than a single evaluation study, the emphasis should be on a coherent and comprehensive evaluation program as much as it is on the quality of single studies. As recent work suggests, advances in the quality of studies are possible, provided of course that elected officials and agency heads who control evaluation budgets and the program officials who monitor access to data believe that seeking them is worthwhile.

References

Ben Franklin Partnership Program, Strategic Investment Committee (n.d.), An Analysis of the Research and Development Component of the Ben Franklin Partnership Program (Harrisburg, PA).


Mt. Auburn Associates, 1992, An Evaluation of Ohio’s Thomas Edison Technology Centers, Final Report to the Ohio Department of Development (Columbus, Ohio).


Rosenberg, N., 1972, Factors affecting the diffusion of technology, Explorations in Economic History 1 (Fall), 3–33.


Shapira, P., J. Youtie and J.D. Roessner, 1993a, Review of literature related to the evaluation of industrial modernization programs, Paper prepared for the workshop on Evaluation of Industrial Modernization Programs: Developing Best Practice (Georgia Institute of Technology, Atlanta, GA).

Shapira, P., J. Youtie and J.D. Roessner, 1993b, Current practices in the evaluation of industrial modernization programs, Paper prepared for the workshop on Evaluation of Industrial Modernization Programs: Developing Best Practice (Georgia Institute of Technology, Atlanta, GA).


Yin, R., 1993, Applications of Case Study Research (Sage, Newbury Park, CA).