
9

EVALUATION APPROACH FOR THE ARKANSAS PILOT RURAL ENTERPRISE CENTER

**David A. Swanson,
John Opitz, Clayton
Franklin, Sandra
Miller, and Floyd
Fenix**

David A. Swanson is Senior Demographic Specialist at the Arkansas Institute For Economic Advancement, College of Business Administration, University of Arkansas at Little Rock, Arkansas. John Opitz works at the Arkansas Institute For Economic Advancement, Clayton Franklin works at the Economic Development Center, Henderson State University, Arkadelphia, Arkansas. Sandra Miller directs the Arkansas Rural Enterprise Center, WinRock International, Morrilton, Arkansas. Floyd Fenix runs WoodWorks, Inc. in Malvern, Arkansas.

In this chapter, we provide an overview of the technology deployment project (Technology Deployment Through Manufacturing Networks or TDTMN) funded by NIST in Arkansas under the grant titled the "Arkansas Pilot Rural Enterprise Center and describe the general strategy we have adopted for the evaluation of the TDTMN project. The following section first provides the overview along with background material. The final section contains a description of the general evaluation strategy we have adopted.

Project Overview And Background

The Arkansas Pilot Rural Enterprise Center is designed to test and evaluate the TDTMN project, which is a demand-driven, manufacturing-network approach to technology deployment to small, largely rural firms in two industries, metalworking and woodworking. Together, these two industries account for 80 percent of Arkansas' defense contracts to small manufacturers and provide nearly 17 percent of the state's durable goods manufacturing jobs.

The TDTMN project is designed to offer cost-sharing incentives for technology deployment. These incentives are constructed on the existing relationships and activities of two manufacturing networks already in existence in Arkansas: the Arkansas Metalworking Connection; and the Arkansas Wood Manufacturers Association. The participants will evaluate the effects of the cost-sharing incentives along with an array of supportive programs thought in advance to have an effect on technology deployment.

Why use a manufacturing network for technology transfer? It has to do with how we defined our target population, which is based primarily on where investment decisions are actually made. Given our target population, the use of a manufacturing network as an instrument of technology transfer has to do with the size and dispersion of the manufacturing units employing the target population in Arkansas.

The Target Population

In the first quarter of 1993 there were 4,145 manufacturing employment units in Arkansas (Opitz, 1994). These units employed about 238,000 during the first quarter of 1993 - an average of 57 workers per unit. Over 720 of these manufacturing units have headquarters out-of-state. These units employ an average of 214 workers. Thus, out-of-state firms employ about 155,000 (65 percent) of the state's manufacturing workers.

Most of the out-of-state units are primarily production processes. Few, if any, of their engineering, marketing, product development functions, and staff are located within Arkansas. Years of economic development experience has found that investment and technology transfer decisions affecting Arkansas are primarily made at out-of-state locations. There is no consistent, efficient way to reach the out-of-state firms through an Arkansas-based technology deployment center. Therefore, we are not addressing out-of-state managed units in our technology deployment proposal. Technology transfer for these firms is probably better addressed through the states where they are headquartered (e.g., California, 38 firms; Connecticut, 26 firms; Illinois, 72 firms; Michigan, 26 firms; Missouri, 72 firms; New York, 72 firms; Ohio, 47 firms; Pennsylvania, 24 firms; and Texas, 74 firms).

Considering, then, where investment decisions are made, a technology transfer program based in Arkansas is left with small firms: the 3,425 manufacturing units headquartered in Arkansas employ about 83,000 - an average of 24 workers per unit. Moreover, these firms are widely dispersed (Opitz, 1986).

Research has also found that a sizable portion of the in-state managed firms is "foods and kindred products." This is an industry where the state and federal governments have already funded a great deal of research and development, including technology deployment. Removing these firms from the target population leaves manufacturing units headquartered in Arkansas that employ just under 50,000. Because they are not desirable at this time as targets of technology transfer, the Arkansas Pilot Center does not seek to affect logging operations,

sawmills, and newspaper printing. This reduces the target population by about 16,000. Thus, there are about 34,000 employees of units with in-state headquarters that belong to the target population. About 85 percent (29,000) of this target population can be found in the 590 metal-working and 700 wood-product manufacturing units headquartered in Arkansas. This is the population targeted by the Arkansas Pilot Center through the TDTMN Project.

Size, Dispersion, and Established Networks

Given the target population, why use "networks" as an instrument for technology transfer? Primarily, it has to do with the size and dispersion of the nearly 1,300 metal-working and wood-product manufacturing units employing the targeted population of 29,000. With an average of just over 22 workers per unit, these 1,300 manufacturing units are widely distributed over the 90,000 square miles of rural landscape that makes up the state of Arkansas. (Opitz, 1994). Of the 1,300 or so units employing the targeted population, 210 are currently members of either The MetalWorking Connection, Inc. (MCI) or the Arkansas Wood Manufacturers' Association (AWMA). The University of Arkansas at Little Rock and Henderson State University are the network brokers for MCI. WinRock International's Arkansas Rural Enterprise Center is the network broker for AWMA. In surveys conducted by the brokers through the two networks, technology transfer assistance has been designated a priority area by firms in the two networks.

The Technology Deployment Strategy

A manufacturer's decision to invest in appropriate advanced technology depends primarily on six conditions:

- (1) an awareness of the technology;
- (2) understanding of the suitability of a given technology for a given manufacturing production system;
- (3) adequate sales and sales margins to generate a return of the technology investment;
- (4) access to capital to purchase and operationalize the technology;

-
- (5) a workforce prepared to use the technology; and
 - (6) a confidence on the part of owners/management capabilities so that the potential benefits of the technology can be realized.

The two networks, MCI and AWMA, provide a base from which to address each of the six conditions as a set. The two networks are already working together to learn from one another's experiences. For example, WinRock has established a for-profit sales company to assist AWMA in the expansion of sales of Arkansas-made wood products. MCI developed a detailed process capabilities inventory to promote the sales of metal products made in Arkansas and would like to use it in conjunction with a for-profit sales company patterned on the one WinRock developed for AWMA. AWMA, in turn, plans to develop a youth apprenticeship program for its members patterned on one pioneered by MCI. Both networks have initiated the development of a revolving loan fund to provide short-term financing. The TDTMN activities of the Arkansas Pilot Rural Enterprise Center are designed to complement the actions of the two networks.

TDTMN Activities Proposed For The Arkansas Pilot Rural Enterprise Center

There are six major TDTMN activities proposed for the Center, divided equally between the metal-working and wood-products manufacturing units and channeled through the respective networks and network brokers. A full description is found in the grant proposal and related documents (Miller, Franklin, Opitz, and Fenix, 1993)

The six TDTMN activities are:

- (1) Individual and Group Counseling by technical specialists operating through the respective networks primarily supported by the Center
- (2) Technical Training and On-site Consultation by highly specialized private and university-based consultants on a cost-sharing basis

- (3) Competitive analysis by teams within each network, with the manufacturing unit paying 20 percent of the cost
- (4) Specialized Consulting Services on a cost-sharing basis, patterned after a project established by the New York Department of Economic Development in 1988.
- (5) In-Plant Technology Demonstrations on a cost-sharing basis
- (6) Project Evaluation, Documentation, and Dissemination of information gained through the evaluation processes.

Major Objectives Of The Arkansas Pilot Rural Enterprise Center's TDTMN Project

The pilot center has four major objectives through its TDTMN activities:

- (1) to evaluate the ability of the two manufacturing networks and their brokers to act as an institutional structure promoting industrial competitiveness
- (2) to determine if differences in the organizational structure of the two networks affect the ability of their respective manufacturing members to adopt new technologies
- (3) conduct a comparative case study evaluating the impact of each network's methods
- (4) assess public policy implications of a demand-driven, entrepreneurial approach to technology deployment.

The Proposed Evaluation Strategy For The TDTMN Project

Following Rossi and Freeman (1985), we consider "evaluation" to be defined as "...the systematic collection, analysis, and interpretation of information to answer questions about the efficiency and effectiveness of program implement-action and impact, using the principles of social science research." Within the context of the TDTMN Project, we acknowledge that an evaluation of both process and outcome (Blalock, 1990: 19) is feasible

but with limited resources for evaluation we focus on outcome.

In evaluating a given program, we ideally want to identify and precisely measure casual relationships. Clear identification and precise measurement call for an "experimental" approach, in accordance with scientific principles. However, as has been noted by others (Oldsman, 1994) the "gold standard" represented by a rigorous experimental approach may not be feasible for a new program like the one we propose. Moreover, we argue that it is not cost-effective to try to use rigorous experimental controls in evaluating a program that affects human subjects, takes place in a non-clinical setting, and is based on voluntary participation and response. The use of an experimental design in such a situation creates the potential for a high "participation and response burden," which is not in accordance with the objectives of either evaluation research or the Paperwork Reduction Act of 1980 (P.L. 96-511). The use of an experimental design may also lead to conflicts with P.L. 93-346, regarding the protection of human subjects. Thus, following Oldsman (1994) we propose the use of the "strong alternative" to the "gold standard" of a rigorous experimental design: a "quasi-experimental" design. This type of design can produce cost-effective, reliable, and useful information that minimizes response burden and ethical issues. Care must be taken, however, to reduce the bias that can arise from a number of different sources in a quasi-experimental research design (Johnson and Stromsdorfer, 1990; and Simpson, 1986).

"Bias" in this context is simply the net effect of those factors not accounted for in an evaluation design. By coupling scientific and statistical principles with a high level of understanding of the program and its participants, much of the potential bias can be controlled, if not eliminated. The plan we propose, is, in fact, designed to meet this objective.

In using a quasi-experimental design rather than a purely experimental one, the trade-off, as noted by Blalock (1990: 33) is that quasi-experimental designs are generally weaker in controlling bias than experimental designs, but are

usually more realistic and have findings that are more generalizable.

There are two quasi-experimental evaluation approaches possible to follow in this program: (1) Net Impact Evaluation; and (2) Gross Outcome Evaluation. We will follow the line associated with Net Impact Evaluation.

The key issue in determining net impact is the selection and measurement of an appropriate "control" group that can be compared with the "experimental" group. This type of evaluation requires a rigorously experimental design and is limited in the number and kinds of questions it can answer (Johnson and Stromsdorfer, 1990). However, we believe that with some relaxation of the rigorous requirements, it can provide a cost-effective assessment of the contribution of the TDTMN program while keeping the participation and response burden low for those concerns being assessed.

The Proposed Net Impact Evaluation Component

There are four major issues to be addressed in the design of the Net Impact Evaluation of TDTMN:

- (1) Developing A Testable Research Hypothesis. Under the general evaluation, the specific research hypothesis we have tentatively developed for this project is: "Manufacturers who participate in networks are more likely to adopt improved technologies than manufacturers who do not participate in networks."
- (2) Sample Design. A sample of TDTMN participants must be assembled so that the results of the evaluation can be generalized. A valid "control" group must be assembled so that the impact of TDTMN can be distinguished from the impact of other factors that also can affect performance. The critical issue here is how to convince those not participating in the TDTMN Program to provide data. One possibility is to offer incentives such as remuneration. Because of the nature of the program and the fact that participation by any one firm is not likely to

be "continuous," the definition of "group" must be thought of very broadly. That is, for any given "intervention," a firm that is a participant may not have been a participant in an earlier intervention. Thus, the concept of "control" and "experimental" will cover specific interventions as informed by other "participation," which will be part of the set of variables used in the evaluation process. The size of the "TDTMN" group and the control group making up a given intervention/participation spell should be such that specific program impacts can be measured with a maximum of precision, given resource and response burden constraints. There are approximately 1,200 firms making up the "target" population.

- (3) Data Collection. "Pre-program" and "post-program" data for both groups must be obtained, processed, and analyzed. The terms are, as noted, oversimplifications in that firms may enter and withdraw from participation depending on the service. Consequently, specific interventions will need to be tracked and "spells" of participation recorded along with information on the type of participation, services rendered, and outcomes. Much of this will be done through the computerized tracking system, which will be expanded to cover each member of the target population over the course of the project. Data from administrative records must be merged with primary data collected directly from the expanded tracking system.
- (4) Analysis. The "two" groups must be examined in detail to determine comparability and to identify any adjustments that may be needed to correct for selection bias. We propose to distinguish between two groups, the "experimental" and "control" by a measure of "participation" at the "pre-project" stage of the project. Specifically, we plan to identify those who qualify as "participants" as constituting the experimental group and those who qualify as "non-participants" as constituting the control group. Keep in mind, these membership in

one group or the other is done by self-selection, in essence, not randomization (Hence, our use of the term "quasi-experimental"). Potential measures of participation at the pre-project stage include: payment of dues (in the network); attendance in workshops (sponsored by the network); participation in surveys (sponsored by the network); service on boards or committees (of the network); requests for technical assistance (from the network); participation in network events

It may be the case that a composite measure of participation will be designed using some or all of the preceding. At the outset of the project, a baseline survey will be conducted. Some of the data from this survey will be used for the network tracking system. For purposes of testing the research hypothesis, however, it will provide a description of technical capabilities (behavioral measures) as well as attitudinal measures relating to the research hypothesis (i.e., likelihood of adopting improved technologies) The survey will be conducted over both "participants" and "non-participants" as identified above.

We will use multivariate statistical procedures to analyze the data. The dependent variable may be a single variable representing one of the behavioral or attitudinal items designed to measure "likelihood of adopting improved technologies" or be a composite indicator representing some or all of both types of measures. If the dependent variable is continuous (i.e. measured at the interval or ratio level), we will use OLS multiple regression or closely related OLS variants, with an eye toward controlling design effects and bias, as will be noted. We will assess the effects of each individual independent variable on a given dependent variable through the interpretive framework of multiple regression using standard criteria such as R^2 , partial regression coefficients, both standardized and unstandardized. Since the data conceptually represent a "sample," we will use inferential statistics to address random error.

If the dependent variable is not continuous (i.e., discrete), we will use the appropriate OLS variants (i.e., Discriminant Analysis, Log-Linear). As in the OLS regression, the analysis will include the use of inferential statistical procedures to address random error (Swanson, Schiller, and McDade, 1992). Since the "sample" is not strictly random, we will also use established procedures and criteria to assess and correct for bias (Swanson, 1986; Swanson, Schiller, and McDade, 1992).

With the multivariate approach we will be able to address the overall effect of the TDTMN program in accordance with our research hypothesis while simultaneously controlling for the effects of other independent variables within the multivariate framework (Swanson, Schiller, and McDade, 1992).

References

- Blalock, A. (ed.) 1990 *Evaluating Social Programs at the State and Local Level*. W.E. Upjohn Institute For Employer Research: Kalamazoo, Michigan.
- Johnson, B. and E. Stromsdorfer 1990 "Evaluating Net Program Impact." pp. 43-132 in A. Blalock (Ed.) *Evaluating Social Programs at The State and Local Level*. W.E. Upjohn Institute For Employer Research: Kalamazoo, Michigan.
- Miller, S., C. Franklin, J. Opitz, and F. Fenix 1993 "Technology Deployment Through Manufacturing Networks." Proposal submitted to the National Institute of Standards and Technology.
- Oldsman, E. 1994 "Evaluation as a Management Tool" Presentation made to the Manufacturing Extension Partnership. 10 March.
- Opitz, J. 1994 "Why Manufacturing Networks For Technology Transfer?" Unpublished Research Memorandum, Arkansas Institute For Economic Advancement, University of Arkansas at Little Rock.
- Opitz, J. 1986 *Planning Arkansas's Economic Future: The New Manufacturers*. Arkansas Institute For Economic Advancement, University of Arkansas at Little Rock.
- Rossi, P. and H. Freeman 1985 *Evaluation: A Systematic Approach. 4th Edition*. Prentice-Hall: New York.
- Shapira, P., J. Youtie, and J.D. Roessner 1994 *Evaluating Industrial Modernization Programs: Issues, Practices, and Strategies*. School of Public Policy and Economic Development Institute, Georgia Institute of Technology. Atlanta, GA.
- Simpson, C. 1986 "Evaluating Gross Program Outcomes" pp. 131-228 in A. Blalock (Ed.) *Evaluating Social Programs at The State and Local Level*. W.E. Upjohn Institute For Employer Research: Kalamazoo, Michigan.
- Swanson, D. 1986 "Missing Survey Data in End-Use Energy Models: An Overlooked Problem." *The Energy Journal* 7 (3): 149-157).
- Swanson, D. J. Schiller and K. McDade 1992 *Tacoma-Pierce County Employment and Training Consortium Job Retention Research Project*. Center For Social Research. Pacific Lutheran University: Tacoma, Washington.
- U.S. Office of Management and Budget 1983 *Contracting For Surveys. Statistical Policy Working Paper, No. 9*. Regulatory and Statistical Analysis Division, Office of Information and Regulatory Affairs, Office of Management and Budget: Washington, D.C.

Evaluations Under Development or in Progress: Discussion

Pounds. No one likes to be evaluated by people who have not done or won't do what they're doing. Evaluation is seen as a hammer. But at the program level, evaluation is extremely important for internal and external effectiveness. It is important for me to have data such as the number of visits it takes to close a project or the cycle time it takes to do a project. We have a lot of dedicated people who want to make a difference, so they are open to program evaluation that provides this type of information.

We need to have data collection methods that are not condescending. We're not the only factor that causes change. There are other business methods that are just as valuable and important as modernization program interaction with the customer. And data collection should be nonbureaucratic. I would love to see that we are collecting some of the same data across programs, but that each program could have a customized portion as well.

Who is the customer for the evaluation--the firm or the sponsor? I straddle both sides. I'd like to see a definite mission statement to include what competitiveness is. I'd also like to see us seek some sort of return-on-investment criteria. If the interactions we have with firms help them stay viable and retain jobs, what measures can we use as return-on-investment? What are these companies putting back into the economy? Sorting out how much an impact our program has is difficult. And for multiple projects, how do we separate the impact of each project?

I hope that collecting data won't result in institutionalizing what we do. We need to have the ability to grow, evolve, and continue to change. If this is a program to lead industry, then it should be evolving.

Sabel. One theme is the "geekiness" or negative image of evaluation. Evaluation is usually done *post hoc*, imposed by legislatures on

programs. Evaluation is to government programs as inspection is to firms. That is what I don't like about evaluation. What is terrific about Eric and Dan's evaluation, however, is that they are self-reflective evaluations that produce compelling results of better ways to design the program.

Shapira. We've received feedback from some program field managers that because they are delivering services they should be mainly responsible for evaluation and making program changes to improve quality. My response is: let's first make the system explicit, find out and describe how the program works through the evaluation process. When we've established some rules based on this description, then we can decentralize responsibility for quality efforts back to the people delivering services.

Sabel. Because evaluation focuses on institutional processes, it will reveal weaknesses. The natural response from those delivering services to evaluators' saying, "we'll make rules and then we'll give them back to you" is, "I don't believe your results." Yet, it should be natural for the organization to evaluate.

Oldsman. Evaluation is a threatening activity. People have a vested interest in the outcome. Evaluation forces people to challenge their assumptions about what the program is doing. Results that suggest that people have to move from their assumptions are threatening.

Heller. The issue is breaking down the barriers in the trenches, and those are real barriers. Without the barriers torn down, the data will be questionable. The test of whether you can take a formative approach and use it for summative results, is whether your intent is formative. It is important to go back to the field and ask them whether this evaluation approach is useful.

Dugger. Fear can be reduced immensely by implementing an in-check system with field agents. Because we spent so much time with field agents, they came up with excellent data collection ideas. It is important to do the in-check now with field staff first.

Jones. Who is the client for the evaluation is an important question. People in programs want to do their best. But program staff see evaluation as a

marketing tool for legislator funding, whether to continue an award or not. Evaluation should be seen as a continuous performance tool with positive incentives for learning and making changes. Instead of using evaluation as a marketing tool, we should use it as a tool for high performance improvement like companies do.

Mark. NIST could say we'll use evaluation to look at different configurations of service delivery and modes rather than to award funding. However, deliverer satisfaction isn't adequate as an impact. My sense is that legislative staff members won't be satisfied with something as non-rigorous as satisfaction-centered evaluations.

Mendelowitz. By using a single evaluation tool for both program improvement purposes and as a basis for funding decisions, you create a conflict of interest on the part of those being evaluated. Accurate data is needed to evaluate program performance and to develop improvements. However, program staff will have a disincentive to report data accurately, if the information may be used to adversely affect their future program funding.

Swanson. Even though instruments have multiple personalities, and they don't do their jobs perfectly, over the long run they can still improve performance. When I first started teaching, forms began to be handed out to students to evaluate my performance. Even though the forms were used for several different purposes, the result was that I improved as a teacher. Students became more satisfied with my teaching, and some of these students even got jobs!

Kaufman. There is confusion about why we are evaluating. Is our purpose to improve the program? Or is it to decide the future of the centers? I would want to know as an MEP, what will NIST do with its measures? What will the measures be used for?

Shapira. To the question of what NIST will do with this, NIST has been very supportive regarding improving the program. NIST has identified and assisted centers with problems, and has taken an active role in reshaping them. It is important to encourage NIST to continue to do this.

Mendelowitz. NIST has an impressive track record of working closely with its centers. But the program is making a transition from a pilot to a national program. In many expanding programs, hands-on management inevitably is replaced by oversight from a distance, which relies on performance indicators instead of direct personal knowledge. The trusting relationship centers have with NIST may change when performance indicators and evaluation, with all their potential imperfections, become the primary management tool of a national program.

Redman. People have so many questions about what NIST wants. Until NIST provides some direction, you feel you are operating in a great vacuum. There is a disconnection between what's going on with NIST and what's going on with the centers. It is becoming clear that the process is no longer organic--you don't feel you are part of the national level. Your decisions are taken with uncertainty. I would like suggestions of what we can do at the national level to make you feel like a partner. That is the challenge of moving from a small to a large program. The outcomes will be meritorious if the partnership process is worked out. As to the customer issue--there is no one customer. So we have to ask, how do we design an evaluation that meets the needs of multiple customers?

Russell. NIST must be able to justify return on investment, and our evaluation system should embrace that need. Clearly program managers should be armed with information about where we are and are not doing a good job. NIST can help by providing a national communications system to share problems and solutions. This would make problems easier to overcome.