The research program of the Center for Economic Studies (CES) produces a wide range of theoretical and empirical economic analyses that serve to improve the statistical programs of the U.S. Bureau of the Census. Many of these analyses take the form of CES research papers. The papers are intended to make the results of CES research available to economists and other interested parties in order to encourage discussion and obtain suggestions for revision before publication. The papers are unofficial and have not undergone the review accorded official Census Bureau publications. The opinions and conclusions expressed in the papers are those of the authors and do not necessarily represent those of the U.S. Bureau of the Census. Republication in whole or part must be cleared with the authors.

USING MATCHED CLIENT AND CENSUS DATA TO EVALUATE THE PERFORMANCE OF THE MANUFACTURING EXTENSION PARTNERSHIP

By

Ron S. Jarmin

Center for Economic Studies,
U.S. Bureau of the Census

CES 95-7 April 1995

All papers are screened to ensure that they do not disclose confidential information. Persons who wish to obtain a copy of the paper, submit comments about the paper, or obtain general information about the series should contact Sang V. Nguyen,

Editor, Discussion Papers, Economic Planning and Coordination, Center for Economic Studies, Room 1587, FB 3, Bureau of the Census, Washington, DC 20233-6101, (301-763-2065) or INTERNET address snguyen@info.census.gov.

Abstract

This paper proposes a framework for evaluating the Manufacturing Extension Partnership (MEP). The MEP is administered by the National Institute of Standards and Technology (NIST) as part of its effort to improve the global competitiveness of U.S. manufacturing industries. As the name implies, the MEP is modelled after agricultural extension. Rather than farmers the MEP's target population is small and medium sized manufacturers, generally those with less than 500 employees.

The MEP currently supports 44 manufacturing extension centers around the country. These centers provide technical and business assistance for manufacturers much as county extension agents do for farmers.

The goal of evaluation is to see if MEP engagements lead to positive outcomes from the view of important MEP stakeholders (e.g., MEP clients, MEP centers, NIST, state and local governments and Congress). These outcomes are discussed in McGuckin and Redman (1995) and include: Process Outcomes (e.g., adoption of a new technology by a client); Intermediate Outcomes (e.g., reduction in the clients defect rate); Business Outcomes (e.g., survival and profits) and Policy Outcomes (increases in employment,wages and/or exports).

The evaluation framework described in this paper has two components. The first component is an evaluation dataset which contains measures of many of the program outcomes listed above for both MEP clients and a representative control group of non-clients. This dataset will be constructed by linking MEP client records with plant level Census data housed at the Center for Economic Studies of the Census Bureau. The Census data provides measures of several outcome and control variables which are comparable across both plants and time. The Census data include observations for all manufacturing plants in the U.S. from which representative control groups can be constructed. The MEP client records provide data on the type and intensity of extension engagements. Linking these rich sources of information yields a comprehensive and powerful dataset for MEP evaluation.

The second component is an evaluation methodology which exploits this rich dataset to make statistical inferences about the impact of MEP services, while carefully controlling for other influences. By using this methodology, we can address many of the shortcomings which plagued previous attempts to evaluate extension services.

In addition to evaluation, the dataset described in this paper may be used to profile the characteristics of MEP clients and compare them to non-clients. The Census data contain the complete universe of manufacturing establishments in the U.S.

These comparisons can, therefore, be very thorough and be performed along a number of interesting dimensions (e.g., compare clients to non-clients in a Center's service area, compare across regions, industries and so on).

Keywords: evaluation, microdata, Manufacturing Extension Partnership

I like would to thank Brad Jensen, Shelia Martin, Bob McGuckin, Ariel Pakes and John Redman for constructive comments on earlier drafts of this paper.

I. INTRODUCTION

The goal of this paper is to assess the feasibility of using the data resources held at the Center for Economic Studies (CES) of the U.S. Bureau of the Census to evaluate the effectiveness of the Manufacturing Extension Partnership (MEP). The MEP is administered by the National Institute of Standards and Technology (NIST) as part of their effort to improve the competitiveness of U.S. manufacturing industries. As the name implies, the MEP is modelled after agricultural extension. Rather than farmers, the target population for MEP extension and outreach is small and medium sized manufacturers.

The MEP operates 44 manufacturing extension centers around the country which provide technical and business assistance to small and medium sized manufacturers much as county extension agents do for farmers. This assistance often consists of providing "off the shelf" solutions to technical problems. However, MEP centers can also channel more recent innovations generated in government and university laboratories to smaller U.S. manufacturing concerns which may not have access to such information. The idea is that MEP services will help these firms become more productive and compete more effectively in the international marketplace.

In order to maximize the effectiveness of the program, it is crucial that MEP stakeholders (e.g., MEP clients, MEP centers, NIST, state and local governments and Congress) have detailed

information about its current performance and that a reliable evaluation framework be in place to analyze its future performance. This paper addresses the issues involved in developing an evaluation framework.

The first step towards developing this framework is to perform, in the near term, a pilot evaluation project. This project would employ currently existing data resources to develop an evaluation framework for the MEP and assess the feasibility of using this framework once the MEP completely up and running.

There are four areas in which this pilot project will provide useful information. First, it will help us learn how to efficiently and accurately evaluate the performance of the MEP. This includes knowing what the relevant evaluation models are and assuring that the appropriate data are available to test them. Second, the pilot project will provide preliminary evidence on the effectiveness of the services currently provided by the MEP. This is accomplished by comparing the performance of MEP client firms to non-clients and by comparing clients receiving different types and intensities of MEP services. Third, it will give policy makers some of the information they will need to increase the effectiveness of the MEP. Finally, the pilot project will permit us to construct limited competitive profiles of MEP clients and suggest how a complete profile of all MEP clients should be constructed.

This paper is organized as follows. First, in section II, we briefly review previous attempts to evaluate agricultural extension programs. Many of the problems encountered in these studies are things we will need to address in evaluating the MEP. Second, in section III, we discuss the first component of our evaluation framework, the evaluation dataset to be constructed by linking MEP client records to plant level Census data. Third, in section IV, we outline the second component of the evaluation framework, the regression based methodology with which the evaluation dataset will be analyzed. We provide a brief summary of the evaluation framework in section V. Finally, we give recommendations on how we should proceed in section VI.

II. LESSONS LEARNED FROM AGRICULTURAL EXTENSION

Few efforts have attempted to rigorously evaluate manufacturing extension. It is, therefore, instructive to first review methods used in past studies to assess the effectiveness of agricultural extension programs. Although significant differences exist between agricultural and manufacturing extension[1], both programs have generically similar objectives (i.e., improve farm/manufacturing performance through outreach and education), and share many of the same evaluation issues. In

[1] See Feller (1993) and Shapira (1990) for discussions about the differences between agricultural and manufacturing extension. See True (1969) for a history of agricultural extension in the U.S.

evaluating either agricultural or manufacturing extension, the goal is to assess whether extension services have any impact on client performance.

The agricultural economics and economic development literatures contain many studies which seek to measure the impact of agricultural extension. Birkhaeuser, Evenson and Feder (1991) review this literature.

In their review, Birkhaeuser, Evenson and Feder (hereafter, BEF) find that researchers typically employ regression analysis to examine the relationship between farm performance and the receipt of extension services. Most such studies find that extension has significant and positive impacts on knowledge diffusion, technology adoption, productivity and profits. BEF note that most studies stop short of claiming that agricultural is beneficial from the view of society as a whole. However, several studies suggest that rates of return to agricultural extension can be very large.

Important for MEP evaluation, BEF point out that the existing studies of agricultural extension are subject to a number of qualifications concerning data and methodology. First, most studies lacked a proper "control group" of similar farmers not receiving extension services against which to compare the performance of those receiving extension services. Use of a control group is important because it permits an estimate of what

might have occurred in the absence of a program, in this case the extension service.[2]

The members of a "good" control group would be as similar to those receiving services as possible. In the agricultural extension context, an evaluator might first consider how closely selected characteristics of farms operated by those not receiving services corresponded to those of farms operated by service recipients. The most important characteristics would be those which most directly influence farm performance, such as type of crops grown, soil quality, size and location.

Second, many studies may have biased estimates of the impact of extension services. This can occur if farmers with some characteristic, such as ability, that is not observable by the evaluator "select" themselves into the class of farmers receiving extension. It could very well be the case that farmers with more ability are the ones most likely to seek out additional information through extension. Biased estimation may also occur

---

[2] For example, a farmer receiving extension services may have increased crop yields two percent a year prior to receiving those services. However, after the farmer began receiving extension services, this rate abruptly jumped from two to four percent. In the absence of other information, we might conclude that extension services made a major contribution to improved farm performance. However, comparing our farmer's performance against a control group of similar farmers, we might find that members of this control group <u>also</u> experienced significant performance improvements. Further investigation might reveal that much better weather was primarily responsible for this improvement. Consequently, the effect of receiving extension services appears considerably less than what we might have concluded in the absence of this comparison.

if extension agents "select" high ability farmers to receive the bulk of their services. In either case, an evaluator can mistakingly credit extension with the superior performance of the high ability farmers. This is because the evaluator can't control for the unobserved characteristics that determine whether farmers receive extension services. To get unbiased estimates of the impact of extension services, the evaluator must account for the "selection bias". To do so requires that the evaluator understand the process by which individual farmers become extension clients.

A third problem common to these studies is that they often fail to take into account the type of services received (e.g., training in silage storage techniques or in the choice of seed varieties) and the intensity with which these services are provided (e.g., number of field agent days of service or cost)[3]. This makes it impossible to know how greatly individual extension

[3] For example, an evaluator might find that farmers receiving services performed much better than a control group of similar farmers receiving no services. However, further analysis, utilizing data on the type and level of services provided, might reveal that most of this better performance was attributable to a small subset of farmers to whom extension agents had provided intensive training in new cultivation techniques, and that, in fact, the performance of other farmers receiving different and/or less intensive services differed little from farmers in the control group. Such an outcome would suggest that the simple receipt of extension services did not foster better performance, but rather the receipt of service of a specific type and/or sufficient intensity level.

services vary in their effect, and leaves the analysis vulnerable to mistaken interpretation.

Finally, these studies also fail to account for the influence of other non-extension programs and secondary information flows. If clients and non-clients differ systematically in their access to non-extension services (these could be offered, for example, by seed companies and other farm vendors), then estimates of the impact of extension may be biased. Also, these studies do not allow for the benefits of extension services to "spillover" from clients to non-clients. For example, it is likely that the knowledge of a new cultivation method flows easily from a client farmer to his non-client neighbors.

In summary, most studies of agricultural extension have found evidence that these programs provided substantial benefits. However, they generally have suffered from four major methodological problems: 1) lack of a control group; 2) selection bias, 3) a failure to incorporate information about the characteristics of the services provided and 4) failure to control for the influence of non-extension services and secondary information flows.

III. PRIMARY DATA SOURCES

Ideally, an evaluation framework for the MEP should explicitly take problems discussed above into account. To do this we require a dataset that has firm/plant specific measures

of performance (e.g., employment growth, survival, productivity, etc.), measures of MEP services (e.g., participation, number of engagements, etc.), and measures of items other than MEP services that influence performance (e.g., plant characteristics such and size, location, etc.) for a sample of both clients and non-clients. In this section, we discuss how to construct a dataset which meets many of these requirements.

This dataset would draw on two primary sources of data: 1) confidential databases on the activities of individual manufacturing plants housed in the Center for Economic Studies (CES) of the U.S. Bureau of the Census, and 2) a small number of data items from Center client records on each plant receiving MEP services. Below, we describe these data in more detail. The discussion subsequently describes how we might use these data to address MEP program evaluation issues.

A. U.S. Bureau of the Census Manufacturing Databases

The CES has microdata[4] files that should prove very useful for MEP evaluation. These data will serve two purposes for MEP evaluation. First, they will allow us to construct non-client control groups and to examine the performance and characteristics of clients before they interacted with the MEP. Second, these

[4] All analysis performed with Census microdata is subject to disclosure analysis.

data provide a number of performance (e.g., productivity, wages, etc.) and control (size, capital intensity, etc.) measures not available in the MEP clients records. Importantly, these data are comparable across both plants and time. The two datasets that are most pertinent to this project are the Standard Statistical Establishment List (SSEL) and the Longitudinal Research Database (LRD).[5]

1. Standard Statistical Establishment List (SSEL)

The SSEL is a list of the names, addresses and other information, for all plants covered by the Federal Insurance Contributions Act (FICA). This list is updated annually and is what the Census Bureau uses to create mailing lists for its economic censuses and surveys. Some of the information contained in the SSEL is provided by the Internal Revenue Service (IRS) and the Social Security Administration (SSA). This information is combined with information form the Census Bureau's Company Organization Survey (COS) and previous SSELs to create the current SSEL.[6]

[5] Other datasets that might be of use include the Survey of Manufacturing Technology (SMT), the Enterprise Statistics and the Research and Development Survey.

[6] Appendix A contains a brief discussion of the availability of the SSEL. Also, table A.1 lists some of the variables available in the SSEL which are of interest for MEP evaluation.

In relation to MEP evaluation, the SSEL is useful for several reasons. First, the SSEL is the only database that contains information on **all** manufacturing plants for every year of interest to the MEP. This makes the SSEL an invaluable tool for assessing the representativeness of other datasets, such as Performance Benchmarking Service's Benchmarking Panel Dataset, used in MEP evaluation,.

Second, the information on names, addresses and other plant identifiers will be useful for matching the Census data to the MEP client data. This matching procedure is discussed in more detail below.

Third, the SSEL contains some basic data on employment and payroll. For many small plants, this is the only data on performance available for non-Census years that is consistent across plants.

Finally, another important use of the SSEL for the MEP evaluation project is survival analysis. The SSEL contains variables which show when a plant is sold or closed.[7] Thus, even for small plants, the SSEL provides a reliable way to measure survival. The SSEL can also be used to track firm[8] (as

[7] When a plant is closed, it is not immediately dropped from the SSEL. It remains in the SSEL for approximately 3 years until processors are sufficiently convinced that the plant is actually closed.

[8] Tracking firm survival is more problematic because the firm identifiers in the SSEL can change due to mergers and divestitures not associated with firm death. To get around this

distinguished from plant) survival.

2. The Longitudinal Research Database (LRD)

The second source of information available from the Census Bureau is the LRD. Census constructed the LRD by linking the plant level data collected for the Census of Manufacturing (CM) and the Annual Survey of Manufactures (ASM).

The LRD[9] contains detailed, time series production data for manufacturing plants. The LRD can be used to track, across time, productivity, wages, employment and other measures of interest to the MEP. These measures can be tracked at a number different levels and dimensions of aggregation. For example, we can track employment growth at the plant, firm and industry levels. Another example would be to track wage changes at the state, region and national levels.

During Census years (i.e., those ending in 2 or 7), the LRD contains data for all of the approximately 350,000 manufacturing plants in the U.S. However, most plants with less than 20 employees (approximately 150,000) are "administrative record" cases. These establishments do not receive a detailed Census

problem, McGuckin and Nguyen (forthcoming) have developed procedures for identifying ownership changes in Census data.

[9] Appendix A discusses the availability of the LRD and Table A.2 lists the data items contained in the LRD.

form.[10] Instead, to reduce the reporting burden on small plants, the Census Bureau compiles data for these plants from records maintained by the IRS and the SSA .

Therefore, the only reliable data items for these administrative record cases are equivalent to the data contained in the SSEL (i.e, employment, payroll, SIC, location, etc.). The Census Bureau estimates ("imputes") values for the remaining variables. For most industries, however, the amount of activity accounted for by administrative record plants is very small.[11]

In non-Census years, the LRD contains data for the approximately 55,000 establishments covered in the Annual Survey of Manufactures (ASM). Plants with more than 250 employees are always included in the ASM. These "certainty" cases account for roughly 1/5 of the plants surveyed in the ASM. The LRD contains very complete, high quality time series data for these larger plants.

The ASM also contains data for a sample of smaller plants. The Census Bureau randomly selects these smaller plants within size and industry categories. The sample is designed to be representative of the population of manufacturing establishments

---

[10] All establishments with more than 20 employees are mailed a Census form. Based on industry and other factors, some smaller establishments also receive a form.

[11] An exception is apparel.

in the U.S. Census redraws this sample every five years.[12] When the ASM sample is redrawn, non-certainty cases from the previous sample are excluded.

The LRD, therefore, has only limited time series information for plants with fewer than 250 employees. Since the target population for MEP services is small and medium sized manufacturers, this feature of the LRD weakens its utility for evaluation purposes. However, the ASM's sample of over 40,000 plants with fewer than 250 employees should include a significant number of MEP client plants as well as a representative sample of non-client plants.

B. MEP Client Data

Now, consider the data to be provided by the MEP centers. Once collected, these data would be sent to CES where they will be matched to the Census data and analyzed. To protect the confidentiality of MEP client data, the staff at CES will apply the same disclosure procedures as they do for Census data.

1. Ideal Client Evaluation Data

Ideally, the MEP centers would provide plant or firm level data on three types of measures. Table 1 list some of the

[12] The ASM was last redrawn in 1989 and will be again in 1994.

variables that the MEP centers should supply for evaluation purposes.

First, we need detailed, systematic and uniform information on the quantity and type of engagements[13] for each client plant or firm. Examples of variables which measure this type of information are in the top portion of Table 1. These variables are all counts of engagements possessing certain characteristics. That is, for each variable (i.e., row in the upper part of Table 1) the MEP centers simply add up the number of engagements meeting the specified requirement they had with a given client, over the course of a calendar year. As an example, for the variable, "# of engagements by type," the Center would report for each client the number of engagements in each of the 14 "substance categories" listed.

---

[13] The term "engagement" is used here to denote an association between a MEP center and a client that includes at least one of the following: a **formal assessment (FA)**, a **technical assistance project (TAP)** and/or a **referred technical assistance project (RTAP)**. The terms in bold are service delivery types defined in The MEP Quarterly Report Users Guide. Any **initial meetings (IMs)** or **informal engagements (IEs)** that lead to an engagement should be included as part of that engagement. An engagement might start, for example, with an **IM** and be completed upon the client receiving the last deliverable from a **TAP**. An engagement, therefore, will most likely span a significant time period. Also, note that a client may have more than one engagement ongoing at any time. Finally, an engagement can include multiple **TAP**s and/or **RTAP**s as long as each individual "**A**ssistance **P**roject" is part of a larger effort to achieve a common and well defined goal. This discussion points to the need for NIST and the Centers to agree on definitions before we proceed with the pilot evaluation.

Second, we would like the MEP centers to provide measures of engagement intensity. These variables are listed in the middle portion of Table 1. These refer to either inputs committed by the Center[14] or fees received by the Center for its services. These variables are intended to measure the "dosage" of MEP services the client receives from engagements of each type.

The final type of information that is needed for evaluation is client identifiers. These are the variables CES analysts will use to match the client data to the Census data. Examples of this type of information are listed in the lower portion of Table 1. The process of matching records from the two sources is discussed in more detail below.

2. Current Status of MEP Client Data

Some preliminary investigations suggest that MEP client records are lacking in several dimensions.[15] Record keeping is not uniform across the centers. Some centers keep detailed records by client while others do not. Also, variables are not always defined in the same way across centers. The most uniform

[14] Some have argued that client inputs be included when measuring engagement intensity or "dosage." We believe that doing so is problematic due to the potential for serious measurement error. Accounting practices will differ significantly across clients leading to differences in the way they report their "investments" and other expenditures which result from an MEP engagement.

[15] This discussion is based on conversations with Eric Oldsman of Nexus Associates.

source of information, across the centers, is The MEP Quarterly Report which does not contain client specific information. Work is underway to catalog the data that are currently available from client records kept by the centers. When this work is completed, we will have much better understanding of the type of evaluation questions that can be answered and what additional data (e.g., Census data) is required to address them.

C. Data Matching

The dataset required by the evaluation framework discussed below is a panel containing both client firms/plants and a control group of non-client firms/plants. This dataset should contain variables which measure performance, MEP services and a number of characteristics which can be used as control variables. It will permit analysts to test a number of hypotheses concerning the effectiveness of the MEP. To construct this dataset, we must be able to identify MEP clients in the LRD or SSEL so that the MEP clients records can be matched to the Census files. An important part of this pilot evaluation project is to find the most efficient method of matching MEP client records to the SSEL and LRD.

This is done by matching variables which are common across the two data sources. Variables of this type which are available in the SSEL and MEP client records include company names,

addresses and so on. Computer matches done on this type of data are not completely reliable and must be verified by hand. Therefore, this method is time consuming and costly. However, since the current number of MEP client records to be matched for the pilot project is relatively small, this process should not be too cumbersome. But as the number of centers and clients to be included in the evaluation grow, it will become more important to have a more efficient matching method in place.

One way to improve the efficiency of the matching process is for the MEP centers to include the Employer Identification Number (EIN)[16] in each client record. This enables analysts at CES to locate the MEP clients in the SSEL or LRD quickly. Since more than one plant can have the same EIN, some additional information such as state or SIC is necessary to complete the match. Employing the EIN is simply an efficient way of greatly reducing the number of records that qualify as a potential match and must be searched over.

D. Data Related Implications for Center Participation in the Pilot Evaluation Project

Many MEP Centers are quite new. These most likely don't have many completed engagements which can be evaluated with

[16] The EIN is the taxpayer identification number assigned by the IRS.

currently available Census data. The most comprehensive post-engagement performance data are contained in the 1992 CM. These data are useful only for clients served by the end of 1991. Since we are interested in testing the feasibility of tracking client performance following engagement with MEP, the best candidate centers would be the original seven MTC's and any more recent TRP awardees already operating several years prior to TRP award.

E. Summary of Data Issues

To sum up this section, the MEP evaluation project requires a panel of both client firms/plants (treatment group) and a representative control group of non-clients. Such a dataset can be constructed by matching MEP client records to Census data contained in the LRD, SSEL or other files. Conditional on successfully matching the records from the two sources, this would be to most comprehensive dataset available for MEP evaluation. The pilot evaluation project will examine the feasibility of matching the client data to the LRD/SSEL and then assess the quality of the matched dataset to determine whether it is useful for evaluation.

IV. METHODOLOGY

A goal of the pilot evaluation project is to test the feasibility of using the dataset described above to determine whether the performance of client plants is systematically related to the receipt of MEP services. Based on the evaluation literature reviewed earlier, an evaluation of the MEP should incorporate an appropriate control group, address the issue of selection bias, and include a way of testing the relative effect of different service packages received by clients. This section describes a methodological framework based on regression analysis that can address these issues.[17]

A. Regression Framework

The most appropriate way to test hypotheses concerning the impact MEP services on client performance is to employ one or more variants of a general regression framework. This methodology has been employed to evaluate agricultural extension programs, and it has seen use in countless studies of plant and firm performance in the general economics literature.

Using regression analysis, the evaluation analyst tries to explain the variation in a "dependent" variable of interest across a sample of firms with one or more independent variable(s). For example, one might wish to explain changes in

[17] Appendix B contains a slightly more technical discussion of the regression framework.

productivity, over a particular time period, for plants in a particular MEP Center's service area. From the dataset described above, the analyst would draw observations on plant productivity for each plant located in Center's service area during the specified time period. She would also draw observations for the variables which she hypothesizes explain the observed changes in productivity. The list of "explanatory" or "independent" variables might include each plant's industry, its size, whether it is a branch plant and whether it has participated in the MEP, over the time period in question. The list of explanatory variables should include all variables which influence productivity for the sample of plants.

For evaluation, the analyst is most interested in testing whether MEP participation has an affect on productivity. In this case, the participation variable can be referred to as the "treatment" variable. The other explanatory variables are called "control" variables. The hypothesis the evaluation analyst wishes to test is; does MEP participation (the "treatment") have a positive impact on productivity after "controlling" for other factors which also influence productivity.

The regression procedure provides two things which allow the analyst to test this hypothesis. First, the procedure provides a parameter estimate for each explanatory variable included in the regression. These measure the magnitude of the relationship between productivity (the "dependent" variable) and each of the

explanatory variables while holding the influence all other explanatory variables constant. If the dependent variable were measured by the natural logarithm of productivity, the parameter estimate for the MEP participation variable[18] would measure the mean percent difference in productivity between clients and non-clients.

The second thing the regression procedure provides is a measure of statistical significance for each of the parameter estimates. If the participation parameter is positive and significant, the analyst "accepts" the hypothesis that participation in the MEP improves plant productivity. If not, the analyst "rejects" the hypothesis.

The evaluation dataset described above, will allow analysts to perform many separate regression analyses in order to test a wide variety of evaluation hypotheses. Table 2 lists some of the more important variables which can be included in these analyses. The table breaks these variables into dependent variables (i.e., those which measure performance or outcomes), treatment variables and control variables.

B. How This Framework Addresses Problems Faced in Previous Evaluation Studies

[18] Participation would be measured by a "dummy" variable that takes a value of one if a plant participates and 0 otherwise.

This regression methodology applied to the evaluation dataset described in the last section is capable of addressing the four evaluation issues discussed in section II. First, the methodology incorporates the need for a control group. By both including non-client plants and (control) variables, other than MEP services, which affect performance, the regression model is able test whether MEP services have a positive impact on plant performance.

Second, an evaluation analyst can test for the existence of selection bias within the general regression framework outlined here. If selection bias is found to be a problem, the analyst must learn about the process through which plants become clients. Once this is done, the regression model can be respecified in order to obtain unbiased estimates of the impact of MEP services.

Third, it is easy to incorporate information on the type and intensity of the MEP services into the regression model (provided that this information is available and reliable). Thus, the evaluation analyst can test whether clients which receive a lot of assistance from the MEP perform better than those receiving less assistance.

Finally, the regression model can control the influences of non-MEP programs and secondary information flows, or spillovers. However, the matched MEP-Census dataset described above is probably not detailed enough to address these issues. Although the data can be used to estimate spillovers from client to non-

client plants/firms, we could probably not confidently say whether the information being "spilled over" originated from a MEP engagement or not. Also, the evaluation dataset contains no information on non-MEP services. In the short term, data form other sources, such as the Benchmarking Panel Dataset, could be used to examine this issue. However, this dataset has very limited coverage and may not be representative of client and non-client populations. Longer term solutions include expanding the coverage of such outside data sources and/or adding questions that deal with these issues to the ASM and CM.

V. SUMMARY

In this paper, we have outlined a framework for evaluating the MEP. This framework calls for combining MEP client records with plant level Census data. This work will be done at the Center for Economic Studies of the U.S. Census Bureau.

Once the evaluation dataset is constructed, CES staff will use it to test a number of evaluation hypotheses. The goal is to see, while controlling for other influences, if MEP engagements are related in a systematic way to the Process, Intermediate, Business and Policy Outcomes discussed in McGuckin and Redman (1995).

The evaluation framework has two components. The first is the evaluation dataset created by matching client records from

the MEP Centers with the Census data. The advantages of using the Census datasets (the SSEL and LRD) for MEP evaluation include: 1) the datasets track a number of outcome and control measures at the plant level and over time, 2) the datasets provide consistent measurement across establishments and time, 3) the datasets cover all manufacturing establishments in the U.S., 4) the data already exist, and 5) the datasets can also be used develop competitive profiles of establishments in MEP service areas. Together with the client data, these data provide a rich and comprehensive source of information for MEP evaluation.

The second component of the evaluation framework is the regression based methodology. CES analysts will analyze the evaluation dataset using regression techniques. They will test a number of hypotheses of interest to MEP stakeholders. The regression methodology combined with the evaluation dataset permits us to avoid many of the shortcomings associated with previous attempts to evaluate extension programs (e.g,, lack of a proper control group, selection bias, no measures of the type and intensity of extension services, and the failure to control for the influence of non-MEP service and secondary information flows).

VI. RECOMMENDATIONS

If the MEP chooses to move forward with this proposal, the following short-term actions are recommended.

First, MEP/NIST circulate a description of the proposed activity to MEP stakeholders including MEP Center representatives, NIST/MEP staff, and representatives of state and federal governments funding MEP activity. This review would solicit input from these individuals regarding the desirability of this evaluation framework, and its technical merit.

A first issue is, of course, whether this exercise is one which would provide useful information to MEP stakeholders. If not, the reasons for its lack of utility should be clarified. If these reservations relate to the proposed methodology, the review should consider ways to modify this approach to better satisfy technical requirements.

Among other technical issues this review should address are:

* To what extent would obtaining client EIN's be impractical, due to cost or other prospective difficulty in obtaining the identifier?
* Are there issues of client confidentiality which might preclude centers from providing the Census Bureau with the required information?
* Which centers would most fully fit the ideal profile for a participating center?

* Are there other performance measures or control variables available which should be added to those delineated above?
* What would be the best types of data among those available with which to profile clients and non-clients?
* Do reviewers have a sense of whether MEP client firms are systematically different from non-clients? What evidence (e.g., the criteria used by centers in soliciting or selecting clients, intuition that certain types of firms are more likely to seek services) of these differences might exist?
* What is the typical period of time (e.g., immediately, one-year lag, two-year lag) that might elapse before MEP engagement might manifest itself in the performance measures? How might this differ by the type/intensity of the project?
* What types of client project data would be the best to use in this analysis?
* What are the best ways to assess whether the pilot project has been sufficiently successful to warrant the wider use of this technique within the MEP system? What types of limitations can the analysis tolerate, and still provide a useful evaluation tool?

If the review indicates that this pilot activity should proceed, the next step would be to identify centers which wish to participate. Discussion would commence with those centers regarding the type of data to be provided and the method of transfer.

The analysis would best begin when the 1992 CM, the 1993 SSEL and the 1993 ASM data become available in 1995. With NIST funding support, Census analysts would link the client information from participating centers with the LRD and SSEL, and conduct the proposed regression analysis. If records were made available the first part of 1995, the analysis could be undertaken during late Winter/early Spring, 1995.

Census staff would report on the results of this analysis during late Spring/early Summer, including a discussion of the strengths and weaknesses of the approach, and the cost associated with this activity. The review team could then reconvene to discuss the report and recommend whatever further action is desirable.

If the regression activity itself proves less useful than anticipated, the review team still could recommend that Census continue providing selected services to MEP upon request, such as a continuing profiling of clients and non-clients or one-time analyses on specific topics of interest to MEP.

If the project proves successful, and it appears desirable to expand matching activity to other Centers, NIST/MEP could

explore the utility of providing resources to the Census Bureau to help accelerate the processing of newly collected Census and SSEL data, and their linkage to the LRD. The object of this inquiry would be to determine whether additional support could reduce the lag time between the collection of data and when the data are available for analysis.

Table 1
List of Variables to be Provided by the MEP Centers

Type of Measure	Variable (Measured at the Firm or Plant Level)
Measures of the Quantity and Type of Engagements	# of engagements performed during the year
	# of engagements by type (e.g., CAD/CAM, EDI/Communications/ LAN, Business Systems/Business Management, Environmental, Quality/ Inspection, Plant Layout/Manufacturing Cells, Automation/Robotics, Control Systems/Integration, Market Development, Material Engineering, Process Improvement, Product Development and Design, Human Resources, Other)
	# of engagements started during the year
	# of engagements started by type
	# of engagement carried over from previous year
	# of engagements carried over by type
	# of engagements completed during the year
	# of engagements completed by type
Measure of Engagement Intensity	MEP Center hours spent on all engagements
	MEP Center hours spent by type of engagement
	MEP expenses on all engagements
	MEP expenses by type of engagement
	Service fees received from all engagements
	Service fees received by type of engagement
Identifiers	Company name
	Plant or facility name (for plants of multi-plant firms)
	Complete address
	Zipcode
	Phone, Fax
	EIN

Table 2
Partial List of MEP Evaluation Variables

	Variable	Dataset
Process Outcomes	Technology Adoption	Client Records, SMT
	Investment[1,2]	LRD
Intermediate Outcomes	Labor Productivity[1]	LRD
	Total Factor Productivity[1]	LRD
Business Outcomes	Sales[1,2]	LRD
	Market Share[1]	LRD
	Operating Profits[1]	LRD
	Survival (0,1)	SSEL, LRD
Policy Outcomes	Exports[1,2]	LRD
	Employment[1,2]	SSEL, LRD
	Wages[1,2]	LRD
Treatment Variables	MEP Participation (0,1)	Client Records
	Service Type (dummies)	Client Records
	Service Intensity	Client Records
Control Variables	Size	SSEL, LRD
	Capital Intensity	LRD
	Geographic Information	SSEL, LRD
	Industry	SSEL, LRD
	Growth Rates Prior to Program	SSEL, LRD
	General Economic Conditions	Other Sources

1) These could also be measured as growth rates.
2) These could also be used as control variables.

Appendix A. Data Availability

The SSEL is available for 1977 and 1982 through 1991. The 1992 SSEL is currently available, but not in its final form. Some editing remains to be done and the final version will be out some time next year (much more processing is done for Census years). The 1993 SSEL should be available for use within the next couple of months. During non-Census years, the SSEL files generally become available slightly more than a year after the end of year for which the data were collected (e.g., the 1993 file will come out near the start of 1995).

The LRD contains data for 1963, 1967 and 1972 through 1991. The 1992 CM is available but not in its final form. Soon after the final version of the 1992 CM is released (next year) the 1993 ASM will also become available.

The variables included in the LRD varies somewhat over time and across establishments. The main reason for this is that ASM plants are asked a larger set of questions. Also, some questions are asked in census years only. Table A.2 lists the variables in the LRD and their availability. A "C" denotes variables that are available in census years only, whereas an "A" denotes those that are available for ASM cases only (even during census years). If nothing is entered in the availability column, the variable exists for all plants sent a form (i.e., for all non-administrative cases in census years and all ASM plants in non-census years).

Table A.1
SSEL Variables of Interest for MEP Evaluation

Variable	Description	Use
Permanent Plant Number (PPN)	Unique and permanent identifier for every plant.	Permits linking plants over time and owners.
Census File Number (CFN)	Unique 10-digit number consisting of a 0 followed by the plants EIN in the single unit case and of a 6-digit Alpha number and 4-digit plant number for multiunits.	This variable is useful as a firm identifier.
Employer Identification Number (EIN)	9-digit taxpayer ID assigned by the IRS.	This variable can be used to match the SSEL and/or LRD to client records which contain an EIN.
Activity Code	Shows whether an establishment was added, deleted or ghosted.	This variable is useful for survival analysis.
Coverage Control Code	Describes the operational status of the plant.	This variable is useful for survival analysis.
Type of Operation	Shows the type of activity the establishment is engaged in.	
SIC	4-digit industry	
State Code	State where the plant is located.	
County Code	County where the plant is located.	
Place Code		
Payroll	Sum of quarterly payroll from SSEL.	
Employment	Total number of employees reported to the IRS.	

Source: Doms and Peck, (1994).

Table A.2
LRD Variables

Description of Variable	Availability
Identification Items	
Permanent Plant Number	
Establishment Identification Number	
Tab Number	
Tabulated Industry Code (SIC Based)	
Derived Industry Code	
Primary Product Class Code	
Primary Industry Specialization Ratio	
Primary Product Specialization Ratio	
Status of Establishment	
Employer Identification Number	
Establishment Type (ASM or NonASM)	C
Administrative Record Case	C
Coverage Code	
Source Code	
Legal Form of Organization Code	C
Location of Establishment	
Census State Code	
FIPS State Code	
Census Region Code	
SMSA Code	
County Code	
Place Code	
Employment	
Total Employment	
Production Workers: March	
Production Workers: May	
Production Workers: August	
Production Workers: September	
Production Workers: Average	
Workers-hours for Production Workers	
Manhours: January-March imputed 81-	
Manhours: April-June imputed 81-	
Manhours: July-September imputed 81-	
Manhours: October-December imputed 81-	
Total Manhours: January-December	
Labor Costs	
Total Salaries and Wages	
Production Worker Wages	
Other Worker Wages	
Total Supplemental Labor Costs	A
Legally Required Supplemental Labor Costs	A
Voluntary Supplemental Labor Costs	A
Costs of Materials, Services and Energy	

Cost of Materials, parts, etc.				
Cost of Resales				
Cost of Fuels				
Cost of Purchased Electricity				
Quality of Purchased Electricity				A
Cost of Contract Work				
Cost of Purchased Communication		C	A	77-87
Inventories-Beginning of Year				
Total				
Finished Products				
Work-in-Process				
Materials and Supplies				
Inventories-End of Year				
Total				
Finished Products				
Work-in-Process				
Materials and Supplies				
Depreciable Assets				
Gross book value (beginning of year)				
Structures			A	74-87
Machinery			A	74-87
New Capital expenditures				
Structures				
Machinery				
Used capital expenditures				
Total	72-76		A	77-
Retirements				
Structures			A	77-85
Machinery			A	77-85
Gross book value (at end of year)				
Structures				
Machinery				
Depreciation Charges				
Structures			A	77-85
Machinery			A	77-85
Rental Payments				
Structures				A
Machinery				A
Repair Charges				

Structures	C	A	77-87
Machinery	C	A	77-87

Receipts

Total Value of Shipments	
Value Added	
Value of Resales	
Receipts of Contract Work	
Miscellaneous Receipts	

Detailed Materials Consumed (Inputs)

Material Code	C
Quantity Produced and Consumed	C
Quantity Received and Consumed	C
Material Delivered Cost	C

Detailed Products Shipped (Outputs)

Product Code	C
Product Quantity Produced	C
Product Quantity Shipped	C
Product Value Shipped	C
Quantity of Interplant Transfers	C
Value of Interplant Transfers	C
Quantity Produced and Consumed	C
New Product Code (current SIC)	C
Product Class Code	A
Product Class Value of Shipments	A
New Product Class Code (current SIC)	A

Purchased Fuels(Quantity, Cost and Stock

Fuel Code	A 74-81
Fuel Quantity Consumed	A 74-81
Fuel Cost	A 74-81
Fuel Stock End-of-Year	A 74-81

Special Inquires

Inventory Valuation Method

First-In-First-Out	A 75-83
Last-In-First-Out	A 75-83
Average Cost	A 75-83
Specific or Actual Cost	A 75-83
Standard Cost	A 75-83
Other Valuation Method	A 75-83
Lower of Cost or Market	A 75-83
Market Cost	A 75-83

Inventory Adjustment

End-of-Year

Amount not subject to LIFO	A 82
Amount Subject of LIFO	A 82

LIFO Research	A 82
LIFO Value	A 82
Valuation Method not Reported	A 82
Amount Subject to LIFO without Reserve	A 82

Beginning of Year

Amount not subject to LJFO	A 82
Amount subject to LIFO	A 82
LIFO Reserve	A 82
LIFO Value	A 82
Valuation Method not Reported	A 82
Amount subject to LIFO without Reserve	A 82

Miscellaneous Variables

ASM Weight	A
Establishment Impute Flag	

Plant History

Year Company Began to Operate	75 & 81
Status in Initial Year	75 & 81
Year of Purchase	75
Year Plant Began Operation at Present	75 & 81

Other Variables

Exports	A Some Years
Unfilled Orders	A 76-83
Interplant Transfers	A 76

Source: LRD Technical Documentation Manual
Notes: C denotes available in Census years and A denotes available for ASM cases.

Appendix B. MEP Evaluation Regression Model

To evaluate the effectiveness of the MEP, one would want to set up a controlled experiment with randomly chosen treatment and control groups. Unfortunately, we do not have the option of performing such an experiment. However, there are several options available for quantitatively evaluating the MEP with the data described above. All the options discussed in this appendix are special cases of a general regression framework. A regression framework such as this has been employed to evaluate agricultural extension and it has been used in countless studies of plant and firm performance in the general economics literature. The advantage of a regression framework for MEP evaluation is that it allows the analyst to measure the impact of extension services while controlling for factors such as size, industry, capital intensity and so on. This general regression framework is described first. Then it is shown how a specific evaluation model can be obtained by placing structure on the general framework.

A general regression model for evaluating the performance of MEP clients is given by

$$y_{it} = f\ (\mathbf{Z}_{it},\ \mathbf{X}_{it},\ (_{it},\ \$_{it},\ g_{it}) \quad \text{for i 0 } N \text{ and t 0 } T \quad (1)$$

where i indexes firms and/or plants and t indexes time periods. The dependent variable, y_{it}, is some measure of performance. The type and intensity of MEP service provided is measured by the vector $\mathbf{Z}_{it}$. Other variables that affect performance are included

in $\mathbf{X}_{it}$. Empirical research at CES and elsewhere suggests that $\mathbf{X}_{it}$ should include plant/firm characteristics such as size, capital intensity, wages, industry and location to name a few. This research is summarized in Jensen and McGuckin (1994). One can view $\mathbf{Z}_{it}$ as the vector of treatments and $\mathbf{X}_{it}$ as the vector of controls. Table 1 lists some of the more important variables available for estimating (1). The vectors $\$_{it}$ and $(_{it}$ contain parameters to be estimated and g_{it} is a random error term. The set *N* contains the firms/plants and *T* contains the time periods to be included in the estimation of the model.

Depending on the evaluation hypothesis to be tested and on the available data, the analyst must choose a performance metric, a set of treatment variables, a set of control variables, and functional form for (1) and determine how the error term enters the model. These choices dictate which estimation procedure is most appropriate (e.g., OLS, logit, probit, NLLS etc).

Say, for example, the analyst is interested in testing whether participation in the program is associated with increases in productivity. She might specify the following linear model

$$\ln y_{it} = \alpha + (Z_{it} + \$\ln(X)_{it} + g_{it} \quad \text{for } i \in N \text{ and } t \in T, \tag{2}$$

where y_{it} measures productivity (or any of the other treatment variables listed in table 1) and Z_{it} is a dummy where a value of one indicates participation in the MEP. The vector of controls might include size, capital intensity, industry and geographical

dummies (or any of the other control variables listed in table 1).

The coefficient (is intended to measure the impact of MEP participation on productivity. The interpretation of this coefficient depends on the nature of the dataset. First, if the set *N* contains only MEP clients and *T* includes years both before and after they receive MEP services, then (measures the percentage change in productivity clients experience after enrolling in the program. The control group, in this case, consists of the clients before they received MEP services. If, on the other hand, *N* contains both clients and non-clients and *T* included only years after clients began receiving service, then (simply measures the percentage difference in productivity between clients and non-clients. In this case, the non-client firms/plants make up the control group. The non-client controls should be as similar to the clients in industry, location, size and other characteristics as possible. If the panel is sufficiently long and contains both client and non-client firms, the regression equation (2) most closely mimics a controlled experiment and the coefficient (comes closer to measuring the true impact of program participation.[19]

[19] Note that this requires that we be able to identify all MEP clients in the data since we do not want any clients placed mistakenly in the control group.

An important estimation issue for the evaluation project is whether clients are selected randomly or not. If they are, regression models such as (2) will yield unbiased estimates of the parameters measuring the impact of MEP programs. If not, the model is misspecified and the parameter estimates are likely to be subject to selection bias. This can happen if good (or bad) firms/plants select themselves to become MEP clients. It may also occur if the MEP centers choose to interact more intensively with better performing firms/plants.

This issue has been discussed in the literature on evaluating job training programs (see LaLonde (1986) and Anderson et. al. (1993)). The direction of the selection bias, in these studies, tends to indicate that the programs are more successful than they actually are. Therefore, more needs to learned about how the MEP obtains clients and how the amount of service provided them is determined. Given this information, a two step estimation procedure can be constructed to correct for the selection bias.[20]

The selection issue brings up another point worth considering as the MEP is about to expand. This concerns how the centers themselves are evaluated. Assume, for example, they are

[20] LaLonde shows that the use if longitudinal data and/or a two step estimation procedure can reduce the potential for misspecification. These do not, however, alleviate the potential for misspecification. He also shows that econometric models which pass standard specification test failed to replicate the experimental results.

evaluated on the job growth of their clients. Knowing this one would expect them to accept as clients only those firms which have good prospects for growth. Clearly this would lead to the appearance that the program was more successful than if clients were chosen randomly. Thus, NIST should be aware that the criteria by which the MEP centers are evaluated may affect our ability to reliably evaluate the program as a whole.

REFERENCES

Anderson, K.H, R.V. Burkhaueser, and J.E. Raymond, "The Effect of Creaming on Placement Rates Under the Job Training Partnership Act, Industrial & Labor Relations Review, vol 46, no.4, pp. 613-624, (1993).

Birkhaeuser, D., R.E. Evenson and G. Feder, " The Economic Impact of Agricultural Extension: A Review," Economic Development and Cultural Change, vol. 39, no. 3, pp. 607-650, (1991).

Doms, M. and S. Peck, "Examining the Employment Structure of Firms in Manufacturing," Mimeo, Center for Economics, U.S. Bureau of the Census, 1994.

Feller, I., "What Agricultural Extension Has to Offer as a Model for Manufacturing Modernization," Journal of Policy Analysis and Management, Vol 12, no. 3, pp 574-581, (1993).

Jensen, J.B. and R.H. McGuckin, "Firm Performance and Competitiveness: Evidence from Microdata Research," mimeo, Center for Economic Studies, (1994).

LaLonde, R.J., "Evaluating the Econometric Evaluations of Training Programs with Experimental Data, American Economic Review, vol 76, no. 4, pp. 604-620, (1986).

McGuckin, R.H., and S.V. Nguyen, "On Productivity and Plant Ownership Change: New Evidence from the LRD," RAND Journal of Economics, (forthcoming Summer, 1995).

McGuckin, R.H., and J. Redman, "MEP Program Evaluation," mimeo, (1995).

Shapira, P., Modernizing Manufacturing, New Policies to Build Industrial Extension Services, Washington: Economic Policy Institute, 1990.

True, A.C., A History of Agricultural Extension Work in the United States, 1785 - 1923, New York: Arno Press and The New York Times, 1969.

U.S. Department of Commerce, "Longitudinal Research Database Technical Documentation Manual " mimeo, Center for Economic Studies, (1992).

CENTER FOR ECONOMIC STUDIES DISCUSSION PAPERS

88-1 "Long-Run Expectations and Capacity," by Robert H. McGuckin & Peter A. Zadrozny, 4/88. Forthcoming in Proceedings, Conference on Capacity Utilization, (1988). (29 pages, $7.25)

88-2 "The Longitudinal Research Database (LRD): Status and Research Possibilities," by Robert H. McGuckin & George A. Pascoe, Jr., 7/88. Survey of Current Business, Volume 68, Number 11, pp. 30-37 (1988). (29 pages, $7.25)

88-3 "Public Use Microdata: Disclosure and Usefulness," by Robert H. McGuckin & Sang V. Nguyen, 9/88. Journal of Economic and Social Measurement, Volume 16, Number 1, pp. 19-40, (1990). (29 pages, $7.25)

88-4 "Modelling Technical Progress and Total Factor Productivity: A Plant-Level Example," by Edward C. Kokkelenberg & Sang V. Nguyen, 10/88. Journal of Productivity Analysis, Volume 1, Number 1, pp. 69-90, (1989). (37 pages, $9.25)

88-5 "Analytic Derivatives For Estimation of Linear Dynamic Models," by Peter A. Zadrozny, 11/88. Journal of Computers and Mathematics with Applications, Volume 18, pp. 539-553, (1989). (41 pages, $10.25)

89-1 "Measuring the Trade Balance in Advanced Technology Products," by Thomas A. Abbott, III, Robert H. McGuckin, Paul Herrick, and Leroy Norfolk, 1/89. Proceedings, 1989 American Statistical Association. (33 pages, $8.25)

89-2 "Using Linked Census R&D-LRD Data to Analyze the Effect of R&D Investment on Total Factor Productivity Growth," by Frank R. Lichtenberg and Donald Siegel, 1/89. Economic Inquiry under title "The Impact of R&D Investment on Productivity - New Evidence Using Linked R&D-LRD Data," Volume 29, pp. 2-13, (1990). (46 pages, $11.50)

89-3 "The Effect of Takeovers on The Employment and Wages of Central Office and Other Personnel," by Frank R. Lichtenberg and Donald Siegel, 6/89. Journal of Law and Economics, Volume 33, pp. 383-408, (1990). (36 pages, $9.00)

89-4 "Entrepreneur Factor Inputs and Small Business Longevity," by Timothy Bates, 6/89. Review of Economics and Statistics under the title "Entrepreneur Human Capital Inputs and Small Business Longevity," Volume LXII, Number IV, (1990). (40 pages, $10.00)

89-5 "The Effects of Leveraged Buyouts on Productivity and Related Aspects of Firm Behavior," by Frank R. Lichtenberg and Donald Siegel, 7/89. Presented at the Conference on Innovation in Corporate Organization, Control, and Incentives, 1990. #3022, NBER Working Paper Series. Journal of Finance and Economics, Volume 27, pp. 165-194, (1990). (53 pages, $13.25)

89-6 "Firm Entry and Post-Entry Performance in the U.S. Chemical Industries," by Timothy Dunne, Mark J. Roberts, and Larry Samuelson, 7/89. Journal of Law and Economics, (1989). (49 pages, $12.25)

89-7 "Price Dispersion in U.S. Manufacturing," by Thomas A. Abbott, III, 10/89. (16 pages, $4.00)