

The Impact of the New York Manufacturing Extension Program: A Quasi- Experiment

**Eric S. Oldsman and
Christopher R. Heye¹**

Eric Oldsman is President of Nexus Associates, Inc., an economic and management consulting firm based in Belmont Massachusetts. Christopher Heye is Research Director at Nexus Associates.

¹The authors wish to acknowledge the contributions of Douglas Welch of Nexus Associates to this project. He was responsible for designing and administering the survey, preparing case studies, and providing critical insights into the operations of the New York MEP and other manufacturing extension programs. The material presented in this article is based on work supported by the New York Science and Technology Foundation. Any opinions, findings, conclusions and recommendations are those of the authors and do not necessarily reflect the view of the New York Science and Technology Foundation.

Abstract

The New York Manufacturing Extension Partnership (New York MEP) was established to help manufacturers in New York compete more effectively and thereby generate economic benefits for the citizens of the state. This article presents an analysis of whether the program has accomplished this goal. Using a quasi-experimental design in which the performance of New York MEP clients is compared to that of similar, non-participating companies, the study finds that clients, *on average*, have changed critical business and manufacturing practices, improved their manufacturing performance, and increased their added value (profits plus labor compensation) as a direct result of the program. Enhanced client performance has led to substantial *public* benefits. The New York MEP program generated an additional \$30 million to \$110 million of value-added income in New York State in the two years between 1993 and 1994. This growth, in turn, led to the creation of approximately 510 to 1,920 jobs in New York during the same time period.

Introduction

Measuring the impacts of manufacturing extension programs is a difficult task. Many methodological, practical, and financial obstacles reduce the likelihood of producing meaningful results. The need for reliable program impact assessments, however, remains critical in light of current budget debates. At a time when virtually every state and federal government program is under pressure to provide justification for further funding, impact evaluations can supply supporting evidence and furnish tangible guidelines for policymaking. Rigorous evaluations are the

best and perhaps only way to deliver “bottom line” estimates of program costs and benefits.

In this article we present the results of the recently completed evaluation of the New York Manufacturing Extension Partnership (New York MEP) to illustrate one approach to measuring program impacts. The New York MEP was established to help manufacturers in New York compete more effectively and thereby generate economic benefits for the citizens of the state. The stated mission of the New York MEP is “*To provide a range of high-quality manufacturing modernization services within the reach of every manufacturer which will lead to improvements in their productivity and competitiveness.*”² The principal aim of the program is to help the 26,000 manufacturing companies in the state with fewer than 500 employees address problems ranging from narrow technical matters to broader concerns about plant management and operations. At the heart of the New York MEP is a regional network of field engineers who provide technical and managerial assistance to companies throughout the state. These individuals help companies identify areas requiring improvement and offer suggestions on steps they should take to enhance their capabilities. Field engineers offer direct advice to clients based on their own expertise and provide referrals to other public and private resources.

The aim of the evaluation is to assess whether and to what extent the New York MEP is having an impact on the performance of small and medium-sized manufacturers in the state.³ The approach employs a quasi-experimental design that

estimates program impacts by measuring New York MEP client performance against that of a group of similar firms that have not received manufacturing extension services.

Methodology And Model Specification

Quasi-experimental design

True experiments based on the random assignment of subjects are the gold standard of program evaluation, providing the strongest foundation for statements of causality. In this type of experimental design, companies are randomly assigned to “treatment” and “control” groups producing groups that are, on average, equivalent with respect to all characteristics. Because of this equivalence, random assignment washes out the influence of other extraneous factors that might contribute to observed outcomes. As such, differences in performance between the two groups can be safely attributed to the program.

However, the potential for utilizing an experimental design in evaluations of manufacturing extension programs is generally impractical for political and administrative reasons. Most difficulties stem from the inability of publicly supported program to deny services to eligible companies. As such, evaluators need to rely on an alternate approach -- so-called “quasi-experimental” designs. While this approach still involves comparing the performance of companies that received services to that of similar companies that have not received services, additional steps are needed to minimize the effects of potential influences unrelated to the treatment process, including selection bias. Selection bias arises when one group of companies differs from the other in terms of characteristics affecting program outcomes. To illustrate, consider the situation where managers of some companies are keen to undertake changes

² New York MEP Extension of Cooperative Agreement #70NANB4H1646, NYS Science and Technology Foundation, June 6, 1995, p.3.

³ The evaluation follows a previous study conducted by Oldsman (1996) of an earlier phase of the NYMEP program. The state of New York requires that its manufacturing assistance programs be evaluated every two years.

leading to improved performance. It is likely that these managers will seek out assistance such as that rendered by manufacturing extension programs more often than managers of companies who are content with the status quo. As a result, observed performance improvements among clients may be due more to pre-existing differences in management than to services provided by manufacturing extension centers.

To combat these threats to validity, care must be taken to construct a comparison group that is as similar as possible to the treatment group with respect to these non-observable characteristics as well as other variables. As discussed in more detail below, the analysis is based on a comparison group drawn from the Performance Benchmarking Service™ of the Industrial Technology Institute (ITI). The use of this dataset, we believe helps minimize selection bias. In addition, quasi-experimental designs require more sophisticated analytical techniques such as regression analysis to help rule out competing hypotheses for observed differences in performance among the treatment and comparison groups.⁴ The specific specification of the model used in the analysis is presented in the following section.

Model specification

In this case, the quasi-experiment was designed to estimate the impacts of the program on New York MEP clients by measuring the difference in changes in value added between a group of clients and a valid comparison group, controlling for other possible influences. Value added is the increase in the value of goods resulting from production. It is computed by deducting the costs of purchased inputs from the value of

the firm's total output. Value added is roughly equivalent to total income, i.e. employee compensation plus profits (before taxes and depreciation). The concept of value added underlies measures of regional economic activity such as gross state product.

The general approach was first to develop a model that could explain as much of the variation in performance as possible, focusing on variables other than New York MEP services. The goal in constructing this type of base model is to include strong control variables that are highly correlated with outcome measures but are uncorrelated with services. By helping to explain the underlying variation in performance, these control variables increase the ability to detect the "true" impact of New York MEP services on clients.

To this end, we began with a modified Cobb-Douglas production function. Our base model assumes that output (Q) is a function of labor (L), capital (K) and materials (M), as well as firm (F) and industry specific (I) effects:

$$1) \quad Q_i = f(L_i, K_i, M_i, F_i, I_i)$$

Subtracting materials (M) from each side, we get:

$$2) \quad Q_i - M_i = V_i = f(L_i, K_i, F_i, I_i)$$

where V_i is the value-added for firm i .

Because we are interested in estimating the program's contribution to the *change* in value-added over time, we employed a growth accounting framework. Taking the first difference of production functions in the two periods, we assume the following relationship:

$$3) \quad \Delta_t \log V_i = \Delta_t \log L_i + \Delta_t \log K_i + F_i + I_i$$

where Δ_t in this case is the change from year-end 1992 to year-end 1994. Firm and industry effects are assumed to be relatively

⁴ For a thorough explanation of quasi-experimental designs, see Cook and Campbell (1979) and Moffitt (1991). Heckman and Robb (1985) review and compare some of the more complex analytical techniques used in program impact evaluations.

fixed and therefore are not represented in terms of a change variable. To control for firm level effects, two additional variables were included: the log of the total number of employees in 1992 and the log of value-added per employee, also in 1992. The former variable controls for initial firm size effects, while the latter controls for initial levels of productivity.

Formally, the “base model” used was as follows:

$$4) \quad \Delta_t \log V_i = \Delta_t \log L_i + \Delta_t \log K_i + \log V_{emp} + \log E_i + I_i + e_i$$

where $\Delta_t \log V_i$ is the change in the log of value-added between 1992 and 1994, $\Delta_t \log L_i$ the change in the log of employees between 1992 and 1994, $\Delta_t \log K_i$ the change in the log of computer keyboards and keypads between 1992 and 1994, $\log V_{emp}$ the log of value-added per employee in 1992, $\log E_i$ the log of employees in 1992, I_i represents industry dummy variables, and e_i an error term.

Because the levels of some of our manufacturing performance measures vary somewhat by industry, the results will be influenced by variations in the industry mix of the two groups. To control for possible industry effects, we split the client and non-client groups into 11 industry categories, based on their SIC classifications as well as other business process measures. The SIC has the merit of being an accepted and available classification scheme but it may not control adequately for other important differences among firm. While the product produced is important, other characteristics such as the type of production process, production volumes and average piece prices are also important. ITI has done substantial work in defining comparison groups for its benchmarking program through the detailed and extensive benchmark survey. This

experience provided a foundation for defining industry categories used in the analysis.⁵

Finally we included program impact variables, $Client_t$, to estimate New York MEP program effects:

$$5) \quad \Delta_t \log V_i = \Delta_t \log L_i + \Delta_t \log K_i + \log V_{emp} + \log E_i + I_i + Client_t + e_i$$

In constructing our “Client” variable, three alternative specifications of program services were considered: total services (i.e., total number of services received by the client between April 1993 and December 1994),⁶ elapsed service time (i.e., the number of days between the first service date and the end of the period), and “service intensity” (an interactive variable capturing both service magnitude and elapsed time effects). Because they received no MEP services, zero values for all three service variables were assigned to all comparison group observations.

Of these three service measures, we hypothesized that service intensity would be the more significant. The total services measure, while important, fails to reflect how much time has passed to allow a given service to take effect. The elapsed service time variable is also somewhat deficient in that it fails to capture the number of services received during the study period. The service intensity variable, however, captures both, and therefore represents a more complete measure of program services.

⁵ This approach was also used in the construction of a valid comparison group as discussed below.

⁶ A more valid measure of the magnitude of services would be the total level of effort (person-hours) committed by service providers to each company during the study period. These data were, however, unavailable.

Data

New York MEP clients

The analysis draws on data gathered through a survey of New York MEP clients conducted at the end of 1995.⁷ The questionnaire asked respondents to report actual historical data on a variety of company performance measures for the years 1992 and 1994, including annual sales, annual expenditures for raw materials and other purchased inputs, annual payroll and total employment.⁸

We surveyed a stratified, random sample of 275 companies drawn from the total population of 1,117 surveyable New York MEP clients. To increase statistical power, five stratifiers were constructed using dichotomous variables expected to be the key predictors of program outcomes:

- **A priori prediction of “high impact.”** A critical characteristic of manufacturing extension services observed in previous evaluations is that impacts on firms are highly skewed. That is, a large percentage of clients appear to experience minimal to moderate impacts while a small percentage of firms experience very large impacts. It was feared that by taking a simple random sample of all clients that large impact firms might, by chance, not be sampled. If this chance occurrence took place,

⁷ The survey tool was developed through multiple iterations. It was reviewed by a number of extension service specialists, including staff at NIST/MEP, Dan Luria of the Industrial Technology Institute, and Philip Shapira of the Georgia Institute of Technology.

⁸ In addition, the survey asked clients to assess the benefits they received as a result of the receiving from the New York MEP program. Clients were asked to report actual business outcome measures for 1995, as well as estimates of what their performance would have been in that year in the absence of the program. These results were positive.

program impacts would be greatly underestimated. Therefore, the New York MEP regional affiliates were asked to review their full list of clients and designate the 10 percent they believed to have experienced the greatest impact in terms of increased value-added. All of these “High Impact” clients were included in the sample.

- **Participation in the Industrial Effectiveness Program (IEP).** Another factor believed to be an important determinant of program impact is whether the firm had received a grant under the IEP program to retain the services of a private consultant. All firms were therefore categorized as to whether they were on record as having received an IEP grant.
- **Firm size.** A third factor theorized to have some bearing on the impact of the program is the size of the client firm. Using information from New York MEP records, firms were identified as “large” if they employed more than 40 employees -- the median size in the sample population.
- **Region.** Another factor expected to influence impacts on firms was the size and age of the regional affiliate serving them. Four of the ten regional affiliates, known as “Manufacturing Outreach Centers,” are generally larger and older than the remaining six affiliates. Clients were, therefore, identified in terms of whether they received the services of an MOC or a non-MOC regional affiliate.
- **Date of initial participation.** Finally, it was suggested by New York MEP staff that the nature of services delivered to clients had changed substantially after August 1994. At that point, the program was expanded. Moreover, the duration of typical services was lengthened, focusing on more substantial projects. If in the survey period a client was first served after August 1, 1994, they were categorized as having received “Late” services.

Of the 275 surveys sent out, 152 were returned, representing an overall response rate of 55 percent. In all subsequent analyses of the survey data, each firm observation was given an analytical weight equaling the inverse of its probability of being sampled.⁹

Comparison group

For analytical purposes, we also needed to construct a comparison group dataset. Ideally, the comparison group should be as similar as possible to the client group with respect to variables affecting outcome measures. Similarity in this context refers to objective criteria such as industry classification, firm size, etc. However, it also refers to subjective criteria such as management's commitment to change.¹⁰

As noted above, the comparison group information is drawn from ITI's Performance Benchmarking Service™ dataset. As part of its benchmarking service, ITI collects historical information on thousands of manufacturing plants through the administration of detailed surveys on manufacturing operations and management. We constructed a subset of the ITI dataset to ensure comparability. Specific steps included:

- Eliminating all companies that had received services from the New York MEP or any other MEC during the study period from the comparison group; and
- Excluding companies that were in industries that differed from the New

York MEP client group from the comparison group.¹¹

The constructed comparison group has one additional feature that merits highlighting. As noted above, the best comparison group is one that mirrors the client group on important dimensions, including non-observable factors such as manager motivation and organizational behavior. It can be argued that the particular comparison group constructed for this project is likely to look like the client group with respect to these other variables. Similar to New York MEP clients, companies participating in the Performance Benchmarking Service select into the program because of their interest in performance improvements. In fact this type of benchmarking exercise is often a precursor to taking required actions among companies in both the treatment and comparison groups. Therefore, the use of the ITI dataset helps minimize the potential for selection bias.

The questions used in our survey of New York MEP companies for the purposes of conducting the quasi-experiment were identical to those in the ITI survey. In addition, the years 1992 and 1994 were chosen to match the availability of the Performance Benchmarking data.

Regression Results

Ordinary Least Squares (OLS)

Our base model was first estimated using standard ordinary least squares (OLS) regression techniques. The results are summarized in Table 1. In all cases, the sign on the service impact variable was positive. The adjusted r-squared was always over .60 -

⁹ Weights needed to be calculated because, as a result of over-sampling, not all clients had an equal chance of being sampled. For example, if clients in one cell had only a 15 percent chance of being randomly chosen, the weight assigned to their survey results equals 6.67 -- the inverse of its probability of being sampled (1/.15).

¹⁰ See Light (1990) for a discussion of comparison groups.

¹¹ As noted above, industry is defined in terms of the primary products manufactured at specific plants following the Standard Industrial Classification (SIC) scheme and other characteristics such as production volumes and typical piece prices.

high for cross-sectional data - suggesting that the independent variables are explaining a large portion of the variation in the dependent variable. As hypothesized, the service intensity coefficient estimate displayed the highest level of statistical significance.

To check for robustness, we estimated several additional equations using different firm and industry effect variables. In all cases, the service intensity variable, and often the other two program impact measures, appeared positive and significant. In many cases, the coefficient value for the service intensity variable was higher than in the estimation results reported in the table, indicating a larger program impact. Importantly, in all formulations, the estimated coefficient for the service intensity variable was contained within a confidence interval that did not include zero. These results strongly suggest that the New York MEP program is having a positive impact on the economic performance of its clients.¹²

Given the model's somewhat complex structure, the coefficient on the "service intensity" variable is difficult to interpret. Technically, the coefficient measures the increase in the log of the ratio of value-added in 1994 to value-added in 1992, per unit increase in service intensity. An easier way to interpret this coefficient is to calculate the change in value-added that

would result for a hypothetical client if its level of service intensity increased. Using the estimated coefficient, the analysis indicates that moving from the 25th percentile on the service intensity scale to the 75th percentile would result in an increase in value-added of \$162,091.

Median Regression Results

These results all point to the positive impact the program is having on small manufacturing firms in New York. However, closer inspection reveals that client value-added growth rates are not normally distributed (see Figure 1). A small group of clients have performed exceedingly well in recent years, growing much faster than other clients or non-clients. Nearly 20 percent of all clients generated value-added increases in excess of 100 percent between 1992 and 1994, a figure that is much higher than in the comparison group. This chart (and additional calculations) indicates that the difference between the mean and median values of the dependent variable is quite large.

Given this discrepancy, we estimated the base model using an alternative regression technique known as "quantile" or "median" regression. Median regression (MREG) techniques estimate a curve through the median, rather than the mean value of observations. Since we have evidence that our distribution of firm performance is skewed because of the presence of a few firms with large performance increases, the median regression approach represents a plausible alternative to OLS.

Median regression results are presented in Table 2. Once again, the service intensity variable is positive and statistically significant. In this case, however, while the total services variable appeared with a positive and significant sign, the elapsed time measure did not. Alternative model specifications yielded roughly similar results.

¹² The study also demonstrated that New York MEP clients, on average, increased the design content of products, broadened their use of computers, increased training expenditures, and adopted formal statistical quality assurance (SQA) techniques. Simple tests of the differences in means between New York MEP clients and the comparison groups revealed statistically significant differences with respect to changes various intermediate performance measures. For example, New York MEP clients, on average, reduced scrap rates by almost one percentage point more than other similar, non-participating companies, reduced manufacturing lead times by 5.5 days more, and increased productivity by 20 percentage points more than non-clients between 1993 and 1994.

Table 1. Dependent variable = Change in value-added 1992-94 (OLS results)

Independent variable	Base model	II	III	IV
Change in labor	0.663***	0.661	0.658***	0.653***
Change in computer capital	0.086***	0.081	0.078***	0.078***
Value-added per employee in 1992	-0.367***	-0.361	-0.361***	-0.345***
Employees in 1992	-0.037***	-0.031*	-0.031*	-0.025
Total services		0.008		
Elapsed service time			0.00012*	
Service intensity				0.000037***
Constant	4.15***	4.05***	4.05***	3.84***
Adjusted R-squared (n=380)	0.614	0.615	0.616	0.625

Note: Dependent and all independent variables, excluding program impact measures, represent logged values. Equation also includes variables (not shown) controlling for industry effects. A joint coefficient test of significance for all industry dummies indicated that these variables are significant beyond the 1% level.
 *** = significant at 1% level; ** = significant at 5% level ; * = significant at 10% level (based on 2-tailed, t-test)

Figure 1. Distribution of the change in value-added, 1992 to 1994

Source: Nexus Associates, Inc.

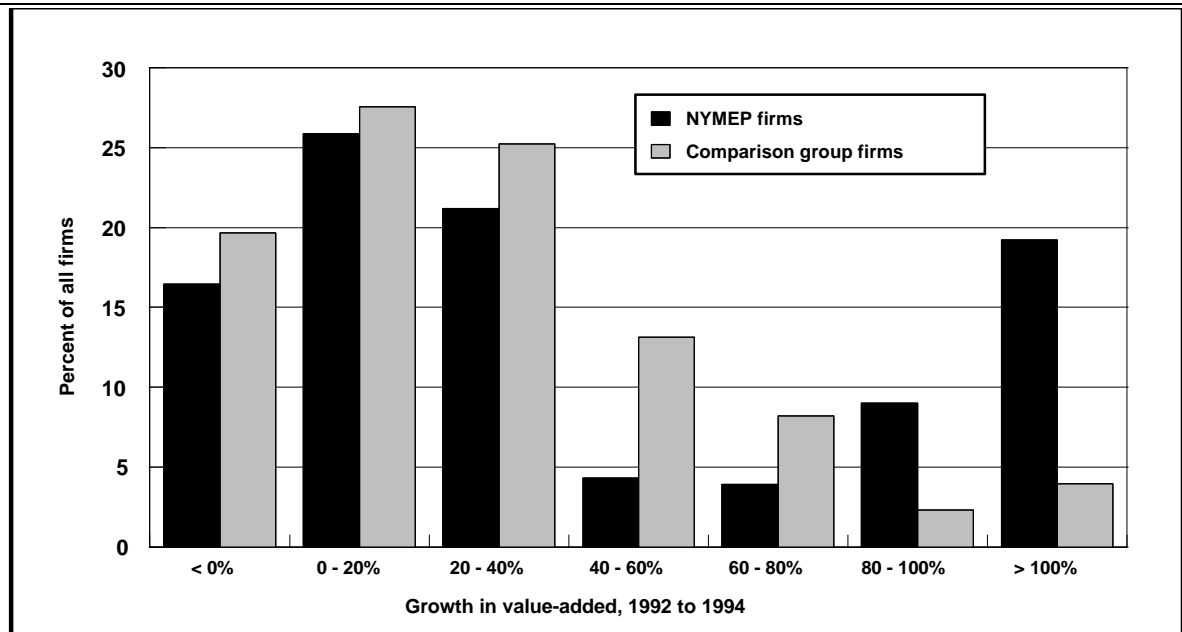


Table 2. Dependent variable = Change in value-added 1992-94 (MREG results)

Independent variable	Base model	II	III	IV
Change in labor	0.660***	0.661***	0.654***	0.667***
Change in computer capital	0.003	0.025***	0.005	0.019***
Value-added per employee in 1992	-0.247***	-0.245***	-0.244***	-0.239***
Employees in 1992	-0.008	-0.007***	-0.010	-0.008***
Total services		0.007***		
Elapsed service time			-0.00001	
Service intensity				0.00001***
Constant	2.79**	2.71***	2.78***	2.67***
Pseudo R-squared (n=380)	0.334	0.341	0.334	0.343

Note: Regression was calculated at the 50% percentile (median). Dependent and all independent variables, excluding program impact measures, represent logged values. Equation also includes variables (not shown) controlling for industry effects.
 *** = significant at 1% level; ** = significant at 5% level ; * = significant at 10% level (based on 2-tailed, t-test)

Economic And Fiscal Impacts

Using the results of these two sets of regression runs, we developed an estimate of the program’s impact on the New York state economy as a whole for the 1992-94 period. Specifically, we used the estimated coefficient values for the service intensity variable to construct an estimate of the increase in client value-added resulting from the program. Since the coefficient value for the service intensity variable, though always positive and significant, varied depending on the structural form of the model and the estimation technique, we constructed an impact range rather than a single point estimate. The range is based on the confidence intervals of the estimated coefficients generated from the two sets of regression results reported above.

Because we surveyed only a sample of New York MEP clients, our results needed to be extrapolated to the client population as a whole. Our extrapolation calculations were performed in the standard way, using weights equal to the inverse of the company’s probability of being sampled. We did not, however, assume that non-respondents had the same characteristics as respondents. In all cases, non-response to the survey were taken to indicate zero impact. That is, unlike many studies of this kind, response bias was assumed to exist such that non-respondents received no benefits from the program.¹³

¹³ This is a very conservative assumption that is known to be wrong in some cases. For example, two case studies we prepared as a separate part of our evaluation of the program revealed major impacts for two clients who never returned surveys. In one case, the firm’s management was overwhelmed with operational issues and did not have sufficient internal records to aid

Two additional calculations needed to be made in order to estimate the *net* impact of the New York MEP program on the New York economy as a whole.

- **Displacement.** First, we sought to control for economic displacement effects. The growth in client firms' value-added is likely to come, in part, at the expense of other companies that manufacture similar products in New York. To the extent that this occurs, the estimate of client impacts will overstate the net increase in economic activity in New York. Therefore, estimates were adjusted downward based on responses to the survey.¹⁴
- **Multiplier effects.** A second set of calculations was performed to estimate income multiplier effects. The multiplier effect refers to the process by which income generated by New York MEP clients in turn creates additional income in the state. To develop our projections, we used information supplied by the U.S. Bureau of Economic Analysis (BEA) regional input-output model (RIMS II). The BEA has calculated a series of multipliers for all states. Their estimate of the income multiplier for New York manufacturing industries is equal to roughly 2.0, meaning that for every \$1 earned by New York manufacturers, an additional \$1 in income is generated in the state. In contrast to the displacement effect, the income multiplier results in an

completion of the survey. In the second case, the firm was unwilling to share detailed financial data. Though they are known to exist, impacts on these companies were therefore not included in the analysis.

¹⁴ The survey asked clients to estimate the percentage of their sales that would have been fulfilled by other plants in New York if they were forced to shut down operations. Estimates ranged from 0 to 100 percent with a mean of approximately 14 percent. We assume the displacement in value-added is proportional to sales.

upward adjustment to total program impacts.

After taking into account displacement and multiplier effects, we estimate that the New York MEP generated between \$29.0 and \$108.7 million in additional value-added for the 1992-94 period. Since all costs incurred by clients in implementing changes recommended by field agents or other service providers are already captured in the value-added measure, these figures represent the total *net* economic benefit of the program with respect to the New York economy as a whole.¹⁵ Using average value-added per employee figures computed from the survey data, we also estimated a total jobs impact. The value-added totals imply that the program was responsible for creating between 513 and 1,923 jobs between the end of 1992 and the end of 1994.

A summary of the findings is shown in table 4. In this table we also present fiscal impacts, which represent the amount of estimated state tax revenues collected as a result of the increased economic activity generated by the New York MEP program. In all cases, the estimated impacts associated with the self-assessment technique are somewhat higher than those generated in the quasi-experiment.

Conclusions

The New York MEP has generated jobs, increased income, and yielded higher tax revenues in New York State. This conclusion is based on the results of a comprehensive evaluation of the program using a quasi-experimental design. This rigorous approach yields credible estimates of program impacts.

¹⁵ Estimated net economic benefits do not include any efficiency losses that might be associated with taxation or the impact of federal transfers.

However, in the absence of random experiments, causality can always be questioned, particularly because of potential selection bias. While the construction of the comparison group and specification of the model used in this evaluation helps minimize potential bias, future efforts need to address this problem in greater detail. Explicit modeling of the selection process using longitudinal data warrants further efforts by

researchers interested in measuring program impacts.

Finally, quantitative analysis is only one approach to estimating the impacts of programs. Statistical techniques should be combined with qualitative research to yield better insights into program dynamics as well as provide a more solid foundation for conclusions concerning program effectiveness.

Table 4. Total estimated economic and fiscal impacts for New York

Value-added impact	Employment impact	Fiscal impact
\$29.0 - \$108.7 mil	513 - 1,923 jobs	\$1.8 - \$6.7 mil

Source: Nexus Associates, Inc. estimates.

References

Cook, T. and Campbell, D., 1979, *Quasi-Experimentation: Design and Analysis Issues for Field Settings* (Rand McNally, Skokie)

Heckman, J. and Robb, R., 1985, "Alternative Methods for Evaluating the Impact of Interventions", *Journal of Econometrics*, 30, pp.239-267

Heckman, J. and Smith, J., 1995, "Assessing the Case for Social Experiments", *Journal of Economic Perspectives*, vol. 9, no. 2, pp. 85-110

Light, R, Singer, J and Willett, J., 1990, *By Design: Planning Research on Higher Education*, (Harvard University Press, Cambridge)

Moffitt, R., 1991, "Program Evaluation with Nonexperimental Data", *Evaluation Review*, vol. 15, no. 3, pp. 291-314

Oldsman, E., 1996, "Does Manufacturing Extension Matter? An Evaluation of the Industrial Technology Extension Service in New York", *Research Policy*, 25 (1996) pp. 215-232