2002

Evaluating the Impacts of Local Economic Development Policies On Local Economic Outcomes: What Has Been Done and What is Doable?

Timothy J. Bartik
W.E. Upjohn Institute, bartik@upjohn.org

Upjohn Institute Working Paper No. 03-89

Published Version
In Evaluating Local Economic and Employment Development: How to Access What Works Among Programmes and Policies, Paris: Organisation for Economic Co-Operation and Development, 2004, pp. 113-142

Evaluating the Impacts of Local Economic Development Policies On Local Economic Outcomes: What Has Been Done and What is Doable?

Upjohn Institute Staff Working Paper No. 03-89

by

Timothy J. Bartik
Senior Economist
The W.E. Upjohn Institute for Employment Research
300 South Westnedge Avenue
Kalamazoo, MI 49007 USA
(269) 343-5541
bartik@upjohn.org

November 2002

JEL Classification Codes: R110, R500, H430, H700

Paper prepared for presentation on November 20, 2002, in Vienna, Austria at a Conference on "Evaluating Local Economic and Employment Development", organized by OECD's Local Economic and Employment Development (LEED) Programme, in collaboration with the European Commission and the Austrian Ministry for Economic Affairs and Labour. I thank Claire Black for her assistance with this project. I appreciate the financial support provided by OECD and the W.E. Upjohn Institute in preparing this paper. I alone am responsible for the findings and conclusions of this paper.

Evaluating the Impacts of Local Economic Development Policies
On Local Economic Outcomes:
What Has Been Done and What is Doable?

Abstract

This paper argues that more rigorous evaluations of local economic development policies are feasible. Programs that aid selected small firms can be rigorously evaluated using an experimental approach, without excluding firms from assistance, by randomly assigning some firms to receive more intense marketing efforts by the program. Programs that aid distressed local areas can be rigorously evaluated by random assignment of the program among eligible distressed areas. If an experiment cannot be done, a variety of statistical approaches can be used to compare firms or areas that use the program with comparison groups of firms or areas that do not use the program. These statistical analyses should be supplemented with surveys and focus groups with businesses that use the program, which give some insight into why the program works or doesn't work. Evaluations should go beyond the effects of programs on business growth to effects on local fiscal health and the earnings of the unemployed. Evaluations using rigorous approaches suggest that programs providing information services to small manufacturers are frequently effective. Programs targeting distressed areas are ineffective unless great resources are used over a lengthy period.

FOREWORD

This paper argues that local economic development policies can and should be more rigorously evaluated. The evaluation should attempt to determine the impact of the policy on local economic outcomes—that is, how local economic outcomes differ compared to what would have happened "but for" the policies.

Programs that provide services or financial assistance to small and medium-sized enterprises (SMEs) can be rigorously evaluated using experimental methods. In such a random experiment, the program would be selectively marketed to randomly chosen SMEs, the "treatment" group, while a control group of SMEs would still be eligible for services but would not receive special marketing efforts . The policy's impact on SMEs can be evaluated by comparing economic outcomes and program usage in the treatment and control groups.

Programs that target distressed local areas for assistance, such as enterprise zones, can also be rigorously evaluated using experimental methods. Areas designated for assistance can be randomly chosen among eligible distressed areas, and the scarce available resources can be more concentrated on these randomly chosen "treatment" areas. The policy's impact on designated areas can be evaluated by comparing economic outcomes in these treatment areas to the eligible distressed areas that were not randomly chosen for assistance.

If experimental data are unavailable or an experiment is infeasible, local economic development programs can and should be evaluated by statistical analyses of economic outcomes in firms or areas using the programs, or more intensively using the programs (the "treatment" group) and economic outcomes in comparison firms or areas (the "comparison" group). There are a variety of well-developed statistical techniques that attempt to determine how much of the

differences in economic outcomes between treatment and control groups is attributable to the program.

These statistical analyses should be supplemented with surveys and focus groups targeting the business clients that use economic development programs. Surveys that are independently administered, ensure anonymity, and ask specific questions can provide additional evidence on the effectiveness of the program in affecting business actions. Surveys and focus groups can also give some insight into how and why a program is effective, and suggest how the program can be improved.

Evaluations should seek to go beyond the impact of policies on increasing local business growth to the benefits of the policy for the public. These benefits include the fiscal benefits for government, and increased earnings for the unemployed or underemployed. Fiscal and employment benefits can be estimated using regional econometric models which are combined with special modules that consider the structure of local taxes and government budgets, and the local labor market.

In the United States, these more rigorous evaluation approaches have been extensively—but by no means universally—used by federal, state, and local organizations concerned with economic development. The results of these evaluations suggest that economic development programs that provide information, training, and consulting services to small and medium-sized manufacturers are frequently effective in improving local business performance. However, programs that target distressed areas, such as enterprise zones, tend to be ineffective if the services and financial assistance offered are too modest to offset the economic disadvantages of the distressed area; more effective economic development programs for distressed areas, such as

the Appalachian Regional Commission, mobilize greater resources over a longer time period. Economic development programs frequently have significant fiscal and employment benefits, however the extent of these benefits varies widely, depending on local conditions. Models can estimate these fiscal and employment benefits if the models incorporate the effects of special local conditions.

Encouraging more rigorous evaluation of local economic development policies probably requires the intervention of higher units of government. These higher units of government should provide funding for evaluations and require evaluations when funding local programs. Such intervention by higher units of government is necessary and appropriate because the benefits of evaluating a particular program go well beyond the organization running the program, and accrue to all organizations that either run or would consider running similar programs, and to the public.

INTRODUCTION

This paper considers the best approaches to evaluating the impacts that local economic development policies have on desirable local economic outcomes.[1] The paper is largely based on my knowledge of state and local economic development policies in the United States, but presumably, similar issues arise in evaluating local economic development policies in other OECD countries.

The paper tries to answer nine questions:

- What are the economic development programs that we are trying to evaluate, and why are they important?
- What type of evaluation of these programs is most needed?
- What biases arise in evaluating these programs?
- Can we effectively use experiments with randomization to evaluate economic development programs?
- Can we use statistical methods to make nonrandom comparison groups truly comparable?
- If a local area has an economic development approach that is truly "unique," can it be evaluated?
- Is there other evidence than statistical comparisons with control or comparison groups that might indicate program impact?
- Can we determine why and how a program has impacts or fails to have impacts?
- Can we determine a program's impacts on ultimate rather than proximate economic objectives?

What are the Economic Development Programs that We are Trying to Evaluate, and Why are They Important?

By "local economic development programs," I mean programs that provide assistance to businesses that is more or less customized or targeted to the needs of that type of business, with

[1] I have previously considered these issues in Bartik and Bingham (1997). The present paper updates my thinking on this topic and considers more recent research findings.

the immediate goal of increasing business activity in the local economy. (There are, of course, ultimate economic objectives to be achieved by increasing local business activity, which I will address later.)

There are many ways of classifying such local economic development programs. Table 1 provides one classification scheme that classifies programs in a way that will later be shown to be relevant for appropriate evaluation techniques. The first type of local economic development programs are those that provide services or financial assistance to only some eligible firms, usually small and medium-sized enterprises (SMEs), with firms either self-selected for assistance or selected by the programs. Such services or financial assistance may include information or training for the enterprise's managers or workers, or public financial support for the enterprise's startup or expansion.

A second type of program provides financial assistance or services to all firms located in a specified area that has been designated as distressed by some higher level of government that helps finance the program. Examples in the United States include the enterprise zone programs sponsored by many state governments, the "Empowerment Zone" program enacted by the federal government under President Clinton's administration, and the Appalachian Regional Commission started in the 1960s.

A third type of program provides assistance throughout the area sponsoring the program, and to all or almost all firms eligible for assistance, although often firm eligibility guidelines target assistance towards the types of firms that are thought to provide the greatest economic benefits. A common target for such programs are manufacturers or other "export-based" firms that export their product outside the area sponsoring the program, although sometimes programs

are more narrowly targeted towards a particular industry, such as some high tech industry. These programs include: marketing an area as a location for new corporate facilities; helping resolve government regulatory problems with new facilities or facility expansions; providing tax breaks, site-specific infrastructure, or customized worker training for new or expanded facilities; and working with networks or clusters of firms in an area to enhance local services or infrastructure.

In the United States, it is estimated that roughly $20–30 billion in state and local government spending or tax expenditures is devoted to such "customized" economic development programs annually, with perhaps another $6 billion annually in support from the federal government.[2] The overwhelming bulk of such resources go to whole area programs, mostly in the form of tax incentives. For example, a recent study of the state of Michigan suggests that, of the $700 million in resources (about $70 per capita) devoted to economic development programs annually, over $600 million is devoted to programs that operate throughout the state for almost all eligible firms, and over three-quarters of this $600 million is in the form of tax breaks, most notably reduced property taxes on new or expanded manufacturing facilities (Bartik, Eisinger, and Erickcek 2003).

[2]The sources for these estimates are discussed in Bartik (2001, p. 251), and are consistent with more recent estimates in Bartik, Eisinger, and Erickcek (2003). In the United States, unlike in Europe, there is no systematic collection of data on state and local economic development program budgets and tax expenditures (Thomas 2000). These figures must be extrapolated from individual state studies.

However, though $20 or $40 billion in resources is significant enough, the importance of local economic development in the United States goes well beyond this relatively narrow definition of local economic development policy. Such a narrow definition focuses on policies that are clearly customized to individual firms or targeted on particular groups of firms, and excludes many more general state and local policies.[3] In state and local debates over taxes, spending, or regulatory policy, the effects of the policy on the state or local area's economic development is always an important consideration (Peterson 1995, 1981). For example, in recent years, almost three-fourths of all states in the U.S. have shifted their approach of apportioning a multi-state corporation's income among the states to an approach that bases half or more of the formula on the state's share of the corporation's "sales," which often enormously reduces corporate income tax collections for firms that export a sizable share of their product outside the state's boundaries (Mazerov 2001; McLure and Herllerstein 2002) . This dramatic change in state business tax policy is usually rationalized as a way to promote the state's economic development. Promoting economic development is also used to rationalize many other changes in state and local policies: other methods of lowering state or local business taxes; lower personal taxes, particularly those paid by high income individuals; reduced welfare benefit levels; changes in workers' compensation laws or unemployment compensation laws; and changes in environmental or health or safety regulations.

Therefore, evaluating local economic development policies is important, not only because of the billions of dollars of resources involved, but also because economic development activity is clearly one of the most important functions of state and local governments in a federal

[3]For example, Thomas (2000) gets state/local business subsidies in the United States of close to $50 billion

system. Distinguishing between strong and weak claims for the effects of some proposed policy in providing economic development benefits is clearly crucial in having well-functioning state and local governments.

What Type of Evaluation of these Programs is Most Needed?

The type of evaluation of local economic development policies that is most needed are estimates of the impact of the policies on desirable local economic outcomes. I will call this "outcome impact" evaluation. Ideally, such an evaluation should include estimates of how outcome impacts will vary with any possible change in the scope, scale, design, or management of these policies, or in other words, that from the evaluation we understand fully how and why the policy has its estimated impacts. In addition, an ideal evaluation would not only tell us the policies' impact on local business activity, which is the proximate goal of local economic development policies, but also the policies' impact on the economic well-being of local residents, the ultimate goal of local economic development policies.

Why is "outcome impact" evaluation needed? Only outcome impact evaluation gives us the information needed if policymakers are to make an informed choice regarding the policy option that will maximize social benefits.

In the United States, a great many reports or studies purport to provide "evaluations" or "performance assessments" of economic development policies, but do nothing of the sort. It has become increasingly common for state and local economic development agencies to produce considerable data on program activities, such as numbers of jobs created by assisted firms.

---

annually by including some of the more general state and local tax expenditures for business.

Agency reports sometimes claim that this job creation is a "program impact," which erroneously assumes that none of the economic activity would have occurred "but for" the program assistance. Also, state and local economic development agencies also often report data on local economic conditions, such as jobs created during a particular time period or reductions in the unemployment rate. Sometimes these reports claim such improvements in local economic conditions as "program impact," which erroneously assumes that any improvements in the local economy are due to local economic development policies.

For example, a study of "business incubators" in the United States, which provide low-cost space, shared support services, and some consulting help to start-up businesses, claimed that "the business incubation programs studied in this project have stimulated the creation of thousands of new jobs throughout the country." (Molnar et al. 1997, p. 12). The study goes on to admit that "some jobs credited to the incubator would have been created even if the incubator did not exist, because a certain number of entrepreneurs will always go into business " (ibid., 13). However, the study claims that "it is impossible to know after the fact what a firm would have done without the assistance of its business incubator program. Consequently here, as in most research on the impact of business assistance programs, analysis focuses upon gross, as opposed to net, impact" (ibid., 13). In contrast to the claims of this business incubation paper, I argue that we can estimate the net impact of the program by estimating what would have happened, on average, if the program did not exist. Furthermore, I believe that the terminology "gross impacts" is misleading, because such numbers are not necessarily impacts of the program.

To avoid confusion, I should emphasize that data on program activities and local economic conditions is often useful. Program activity data helps in managing programs, and

local economic condition data helps in understanding the local economy. These data may even be part of the information that is needed to do a true "outcome evaluation" of local economic development policies, which seeks to identify a cause and effect link between program activities and local economic conditions, and quantitatively estimate its magnitude. By itself, however, data on program activities or local economic conditions do not tell us the impacts of policies on outcomes.

Outcome impact evaluation is often expensive in its demands for more data and expertise in statistics and economic modeling. Because such outcome impact evaluation is expensive, it is not clear that such evaluations need to be performed on each and every program run by each local economic development agency. Individual local economic development agencies are probably best advised to reserve outcome impact evaluation for their most expensive programs, for which the possible gains from better policy choices are the greatest. Higher levels of government may provide a useful service by paying for the evaluation of smaller programs, and ensuring that the results are widely disseminated to the local economic development agencies that use, or might use, similar programs.

What Biases Arise in Evaluating these Programs?

The ideal—but impossible—study of a government program would borrow a time machine from H.G. Wells or some other science fiction writer, go back in time and eliminate the program but make no other direct intervention, and then compare the outcomes in this induced alternative world *without* the program to the outcomes in the original world *with* the program. Absent a time machine, the next best alternative is to find some group of entities that are

comparable to the group of entities receiving the effects of the program, but this comparison group has no involvement with the program. For local economic development policies of type 1 (see Table 1), in which only a subset of eligible firms receive assistance, the comparison group would consist of firms that do not receive assistance. For local economic development policies of type 2, which target distressed areas, the comparison group would consist of areas that are not officially designated as distressed. For local economic development policies of type 3, which serve all eligible firms in the area sponsoring the program, the comparison group would consist of other areas.

For such comparisons to immediately and easily reveal, without statistical torture, the causal effects of the local economic development policies on local economic outcomes, the comparison group will have to be the same, on average, in observed and unobserved characteristics that affect local economic outcomes. Absent experimental data, which will be discussed later in the paper, the group receiving program assistance will generally differ from the comparison group in ways that affect local economic outcomes. Therefore, the assisted group and the comparison group would be expected to experience different changes in economic outcomes, even if neither group received program assistance. As a result, a simple comparison of the two groups will provide a biased measure of program effects.

What are the likely direction of these biases? For local economic development policies that selectively aid firms (policies of type 1), our intuition is that rapidly growing firms are more apt to self-select into participation in the program, precisely because their growth leads them to be more in need of financial assistance and services. There is some evidence that rapidly growing firms are more likely to use selective firm services provided by local economic

development agencies (Jarmin 1999). Furthermore, there is some evidence that firm growth is positively correlated over time (Nexus 1999). Under these conditions, firms that participate in the program would have been likely to grow more rapidly in the future even if they had never participated in the program, which will bias evaluations towards overestimating the positive effects of the program. (Of course, particular local economic development programs may have different biases in their evaluations if the programs select firms for assistance in a different way, or if the change in economic outcomes variable that is examined is not positively correlated over time.)

For local economic development programs that target distressed areas (policies of type 2), these distressed areas—by definition—are likely to have higher levels of economic distress than non-designated areas. (For evidence, see Bondonio and Engberg 2000; Greenbaum 1998; Greenbaum and Engberg 1998.) Therefore, any study that compares the levels of economic outcomes for targeted areas versus some comparison group of areas is likely to be biased towards finding negative effects of the program, as levels of economic outcomes are obviously positively correlated over time, and therefore the targeted areas would have higher levels of distress than their comparison group in the future without the program's intervention. It is not as obvious that changes in economic outcomes will differ between targeted areas and comparison non-targeted areas. In fact, some evidence suggests that, in the United States, the correlation between area designation as an enterprise zone and prior area growth is slight (Bondonio and Engberg 2000). (Again, the bias tendencies in evaluations of a particular program will depend on the targeting rules of the program.)

For local economic development programs that serve all eligible firms throughout the area sponsoring the program (policies of type 3), the bias tendencies in evaluations will depend upon what types of areas are more likely to aggressively pursue economic development. The available evidence suggests that, in the United States, incentives do tend to be somewhat higher in states or cities with higher unemployment and previous slow growth (Fisher and Peters 1998). However, these incentives do no more than offset the generally higher effective basic state and local business taxes that prevail in these high unemployment and slow growth areas, so the effective state and local business tax rate after incentives is not strongly correlated with state and local unemployment rates or employment growth. Therefore, studies of the effects of incentives may be biased towards finding less positive effects of incentives on local economic growth, effects, as state and local areas that heavily use incentives would be more likely to grow slowly even without incentives. On the other hand, studies that look at the effects of basic state and local business tax rates on growth may be biased towards finding more positive effects of lower business taxes, as slow growth states tend to have higher state/local business tax rates (for confirming evidence for the same state over the business cycle, see Reed and Rogers 2000).

Can We Effectively Use Experiments with Randomization to Evaluate Economic Development Programs?

The best feasible way to avoid bias in estimating the outcome impacts of economic development programs is to experiment with the programs by creating some random process which will help determine which entities (firms or areas) will use the program and which will not. Because the process determining the use of the program is random, we know that the program and treatment groups must be the same, on average, in observed and unobserved

variables affecting economic outcomes. Any remaining differences in economic outcomes between the program and treatment groups are either due to the program, or to random factors affecting economic outcomes for a particular firm or area. With a sufficient sample size, these random factors will average out to zero, and we will be able to precisely estimate the true impact of the program on economic outcomes.

To my knowledge, the only economic development evaluation that has relied on data generated from an experiment using random assignment is a study, sponsored by the U.S. Department of Labor, of the effects of entrepreneurship training for UI recipients (Benus, Wood, and Grover 1994). In this experiment, UI recipients in the states of Massachusetts and Washington were first invited to orientation sessions explaining the entrepreneurship training program. The three percent of UI recipients who expressed interest in such training after the orientation were then randomly assigned to a treatment group that received such training, and a control group that did not. Forty-nine percent of the treatment group ended up with some self-employment experience, compared to 28 percent of the control group, with no sign of a different business failure rate in the two groups. Because the treatment and control group, on average, should be the same in observed and unobserved characteristics, we can be confident that, except for random noise, the 21 percent differential in self-employment experience is due to the entrepreneurial training program. Note that the usual program practice of claiming credit for all business activity associated with the program would exaggerate the effects of the program more than twofold, claiming credit for all 49 percent of the treatment group that had self-employment experience. Economic development programs cannot legitimately claim credit for all jobs and

other business activity that are assisted by the program, because at least some—perhaps all—of this business activity would likely have occurred even without the program.

Random experimentation methods could readily be used with other local economic development policies that only assist a select group of eligible firms. One concern about such experimentation is a reluctance to exclude some firms from services, which is what is done in classical experiments with the control group. Such exclusion can be avoided if the experimentation takes the form of random selection of firms for targeted marketing of the program. Randomization methods would be used to choose which firms would receive an intensive marketing effort, such as letters, phone calls, and personal visits, informing the firm of the services or financial assistance provided by the economic development program. If this marketing is intensive enough, the result should be some significant difference in usage of the program between firms in the treatment group (the group receiving targeted marketing efforts) and the firms in the control group (the group not receiving targeting marketing efforts). However, no firm in the control group that requested services would be arbitrarily denied services. The difference in economic outcomes (job growth, productivity growth, etc.) between the treatment and control groups of firms, divided by the difference in program usage between the two groups, provides an estimate of the effects of the program. For example, consider a manufacturing extension program designed to improve firms' productivity growth, and a random experiment that intensively marketed the program to a randomly chosen treatment group of firms. If productivity in the treatment group increased 10 percent, productivity in the control group increased 5 percent, and program usage in the treatment group was 35 percent, versus 10 percent in the control group, then the estimated productivity effect of the program is a 20 percent

improvement in productivity (20 = (10 - 5)/(0.35 - 0.10)).[4] Because the treatment and control groups on average only differ in what random number they were assigned, and thereby whether the program was marketed to them, we can be confident that with sufficient sample size this calculation will reveal the impacts on economic outcomes of the program.

Random experimentation could also be done with economic development programs that target distressed areas. In general, there are more economically distressed local economies than a higher unit of government can afford to target with sufficient resources to realistically help turnaround a distressed area's economic fortunes. Furthermore, it is unclear whether, among distressed areas, one should target the most or least distressed: the most distressed areas may need help more, but the least distressed may be easer to affect with the right program. Therefore, any effort by program managers to select target areas among all distressed areas are likely to reflect fairly arbitrary judgements. Finally, in practice it is often the case that higher levels of government use political criteria to select which distressed areas will be designated for assistance. For example, during the Clinton administration, in selecting which areas of the United States would be targeted for an "Empowerment Zone" or "Enterprise Community," the final targeted zones were chosen by political appointees, and did not rigidly follow the ranking developed by a selection panel. Given the inherent arbitrariness and political nature of current

[4]This discussion glosses over the issue that the program effect may vary across firms. Technically, all that this experiment can estimate is the effect of the program in the extra 25 percent of all firms that are induced by marketing to use the program, which may differ from the productivity effect among the 10 percent of firms which use the program without marketing, or the remaining 65 percent of firms that are unaffected by marketing. This is sometimes called the "local average treatment effect" (LATE), or the "marginal treatment effect" (Heckman and Vytlacil 2001; Imbrens and Angrist 1994).

procedures for designating distressed areas for assistance, there should be no serious ethical issues for such designation to be done using random assignment. If this were done, the designated areas and the undesignated areas would, on average, be the same in observed and unobserved characteristics and growth prospects, and the difference in economic performance of the two groups would be an unbiased estimate of the effects of the program. For such estimates to be precise enough to be useful, there would have to be a sufficiently large number of randomly chosen designated and undesignated areas so that random factors average out. How large the sample size would have to be depends upon how large a program effect one is trying to detect, and on how much natural variation there is in the economic performance of distressed areas; standard statistical methodologies allow such issues of sample size to be systematically answered. As a rule of thumb, it seems unlikely that in most cases much could be learned without a sample size of at least 20 in each of the groups, the designated and undesignated distressed areas.

For local economic development policies that serve all eligible firms in the entire area, random experimentation is not possible by definition, as these programs are sponsored by the area, and the area government will not control what programs are adopted by the governments of other areas. As mentioned before, in the United States, whole area programs receive the majority of resources devoted to local economic development policies. Thus, for many economic development policies, random experimentation is infeasible.

Can We Use Statistical Methods to Make Non-Random Comparison Groups Truly Comparable?

Because experimentation isn't often done with local economic development policies, and is infeasible with some policies, it is important to explore alternatives. We will often have some data on economic outcomes for firms or areas that use a local economic development program or use a program more intensively (the "treatment" group) and those that do not (the "comparison" group). Because the treatment and comparison groups differ in observed and unobserved variables that affect economic outcomes, a simple comparison of outcomes for the two groups may not reveal true program effects. However, there are a number of statistical techniques that can be used to limit or even eliminate the biases resulting from these differences between the treatment and comparison groups.[5]

This is not the appropriate place to go into all the technical details of the appropriate statistical techniques, but briefly, there are at least five statistical techniques, not necessarily mutually exclusive, that can be used to detect the true effect of a program on some outcome variable when the program users differ in other ways than program use from the nonusers. First, we can simply statistically control for observed variables that affect the economic outcome and might be correlated with program use by including these observed variables in the estimation equation that is used to predict the outcome variable. This approach is most effective in reducing the bias in estimation of program impacts when we have data on as many variables as possible that affect the economic outcome of interest and are correlated with program use. This approach

[5]Much of the following discussion in this paper is phrased as if the program is measured by a zero-one dummy, in which the treatment group uses the program and the comparison group does not. However, in general the discussion is generalizable to a situation in which there are different levels of use of a program, and the program usage variable is a continuous variable.

cannot correct for biases that might be caused by unobserved variables that are correlated with both economic outcomes and program use. This approach also assumes that we know the functional form by which the observed variables affect economic outcomes.

A second approach that is a variant of the first goes under the label of "difference-in-differences" estimation, or "difference-in-differences-in-differences" estimation (DD and DDD for short) (Meyer 1995). Under a DD approach, we compare the difference before and after the program of the differences between users and non-users of the program or policy. Under a DDD approach, if we have reason to think that some types of users are likely to be more affected by the policy than another, we can compare the difference between the likely high impact and low impact groups in the user and non-user group before and after the policy. A DD approach is equivalent to assuming that one can do a good job for controlling for other factors affecting economic outcomes by allowing for effects of the time period, and for whether the entity is in a user or non-user group or a high-impact or low-impact group. The limitation of this approach is that there may be many other variables, both observed and unobserved, that also affect economic outcomes and are correlated with program use. The second approach can be combined with the first approach by adding some of these other observed variables to the estimation equation.

A third approach is matching program users with non-users who are similar in observed characteristics. Recent research has revealed that this matching should focus on finding users and non-users who are as similar as possible in their estimated "propensity score," which is an estimated probability given observed variables that a given entity will use the program (Smith and Todd 2001; Heckman, Ichimura, and Todd 1999).

This propensity score should be estimated using variables that predict program use and have a correlation, independent of program use, with economic outcomes. Variables that predict program use but do not independently predict economic outcomes should not be included in the prediction of program use. Such variables, both observed and unobserved, provide the variation in program use that is independent of non-program factors affecting economic outcomes.

The propensity score approach works well if we have data on all the variables that do a good job of predicting program use and are also correlated with non-program factors affecting economic outcomes. This matching approach will not work well if there are many unobserved variables that predict program use and are correlated with economic outcomes. In addition, in many cases there may be no reasonably close matches for some users with non-users, and the estimates from a matching approach hence are only valid as average program effects for the types of program users for which we can find good matches among non-users.

A fourth approach is explicitly modeling selection into the program and how it is correlated with unobserved variables affecting economic outcomes (Murnane, Newstead, and Olsen 1985). This requires the estimation of three equations: one equation explaining economic outcomes for program users, a second equation explaining economic outcomes for non-users, and a third equation explaining whether a given entity is a program user. The estimation of the third equation allows a "selection bias correction" term to be added to each of the first two equations, which—in theory—corrects for the bias caused by unobserved variables that affect economic outcomes and are correlated with program use. This approach assumes that we have accurately specified the variables and functional form that should enter all three equations. In

addition, this approach assumes a particular statistical distribution for the unobservable factors (the "error terms") that enter all three equations.

A fifth approach requires finding some "instrumental variable" that predicts program use and is uncorrelated with unobservable variables that affect economic outcomes (Angrist and Krueger 2001). Under this instrumental variable approach, we only examine the change in economic outcomes that can be attributed to shifts in program use that are statistically associated with shifts in the instrumental variable. The intuition is that the effects on economic outcomes of these instrument-induced shifts in program use show the true effects of the program because these shifts in program use will be uncorrelated with unobservable variables predicting economic outcomes, as these shifts are generated by an instrumental variable that is uncorrelated with unobservable variables predicting economic outcomes. The problem is finding such instruments. The instrumental variable must do a good job of explaining program use. Otherwise, the estimation approach throws away too much information. But the variable must have little (ideally, zero) correlation with unobservable variables affecting economic outcomes, and it is difficult to test assumptions about the correlation of a proposed instrument with unobservable variables. Good instruments are hard to find and may not be convincing to all readers.

These five approaches can be combined in different ways. For example, researchers can create matched data sets, include controls for various observed variables in the estimation, and use instrumental variables for the program variable.

There are many good examples of impact outcome evaluations for local economic development policies that use non-experimental data. For programs providing assistance to selected firms, Holzer et al. (1993) implicitly used an instrumental variable approach to study a

program providing customized training to a firm's workers. This program, run by the state of Michigan, provided grants to manufacturing firms for worker training. The comparison group was firms that applied too late in the fiscal year to receive a grant. The implicit assumption is that the time a firm applied for a grant is an "instrument" that explains participation in the program, but is uncorrelated with unobservable variables affecting a firm's performance. The study found that firms that received grants had significantly lower scrappage rates after that training was completed than firms that applied for grants but did not receive them. The study would yield biased results if firms that applied late in the fiscal year differed in ways we cannot control (e.g., if such firms were more poorly managed).

Another good study of a firm selective program is Jarmin's evaluation of the federally-sponsored Manufacturing Extension Partnership, which provides consulting advice to small and medium-sized manufacturers in improving their productivity (Jarmin 1999). Jarmin's paper first does a rough match by only including non-clients in the data if they were located in the two states where all his clients were located. The paper then controls for selection bias by estimating an equation predicting whether a given manufacturing firm becomes a client of the MEP (for example, this is affected by whether the firm happens to be in a metropolitan area that has a MEP center), and including a selection bias correction in two equations predicting a firm's productivity growth, one equation for firms that are clients of MEP, and another equation for firms that are not clients of MEP. Jarmin's study finds that MEP increases productivity by 3 to 16 percent.

For targeted areas, a number of studies in the United States have attempted to evaluate enterprise zones by comparing the performance of enterprise zones to matched non-zone areas.

Several studies by researchers at Carnegie-Mellon University (Bondonio and Engberg 2000, Greenbuam 1998, Greenbaum and Engberg 1998) have explicitly made such matches using estimates of the "propensity score," that is estimates of the probability of a given area (in this case, a postal "zipcode" or routing code) being designated as an enterprise zone. These studies find little or no effect of enterprise zone designation on the growth of local business activity. In addition, as mentioned before, the propensity score estimation suggests that enterprise zone designation is not strongly correlated with previous area growth, which increases the odds that the estimates reveal the true effect of enterprise zone designation. Other studies have also examined the performance of enterprise zones with non-zones (e.g., Papke 1993, 1994; Hebert et al. 2001), but published versions of the research do not contain sufficient information to judge the validity of the matching. Finally, one forthcoming study (Peters and Fisher 2002) evaluates state government-designated enterprise zones by comparing the performance of enterprise zones with high versus low levels of incentives. The assumption is that unobservable factors affecting an area's performance might be correlated with the area's designation as a zone, but will not necessarily be correlated with the magnitude of the zone incentives, which depend on the various political compromises in whatever enterprise zone bill was enacted by that particular state. This study finds little effect of the magnitude of a zone's incentives on firm start-ups or expansions.

For local economic development programs that assist all eligible firms in an area, there is a huge literature on the effects of state and local taxes on business location and growth, which has been summarized by Bartik (1991, 1992) and Wasylenko (1997). These studies typically deal with possible correlations of taxes with other variables affecting business location and growth by including as many relevant location and growth factors as possible as explanatory variables in

the estimating equations. As summarized by Bartik and Wasylenko, these studies generally come up with an elasticity of state and local business activity with respect to state and local business taxes in the range from -0.1 to -0.6. That is, a 10 percent reduction in overall state and local business taxes will eventually increase a state's business activity by 1 to 6 percent.

Over the last 10 years, Andrew Isserman and his colleagues have done a number of papers that evaluate various economic development interventions in the U.S. by matching counties with the interventions with comparison counties without the interventions, using pre-intervention data. One such study indicated that the Appalachian Regional Commission increased growth in Appalachian counties compared to matched counties outside Appalachia, with this growth effect strongest in counties in which the ARC built highways (Isserman and Rephann 1995). Another study evaluated a large tax cut in the state of Illinois by matching each county in Illinois with similar counties outside Illinois, and found that the tax cut had some short-run economic growth effects but no significant long-run effects (Rogers and Reed, forthcoming).

Finally, one of the best studies of the effects of state taxes compares the business location decisions among U.S. states of foreign firms from two groups of countries: countries in which U.S. state taxes can be credited against the firm's tax liability in its home country; and countries in which U.S. state taxes can only be deducted against taxable income subject to home country taxation (Hines 1996). For firms from the former countries, U.S. state taxes should be irrelevant to business location decisions for any firm with positive tax liabilities in its home country. Hines (1996) found that firms from the first group of countries located in higher tax U.S. states than firms from the second group of countries. This can be seen as a form of DD estimation or

instrumental variable estimation . The implicit assumption is that the only difference between the two groups of firms that is relevant to their business location choices is how their home country treats U.S. state taxes. If this assumption holds, then the resulting estimates are convincing evidence that state and local taxes do affect the business location decisions of large corporations in the United States.

If a Local Area Has an Economic Development Approach that is Truly "Unique," Can It be Evaluated?

Experimentation or statistical analysis using comparison groups assumes that one has data on a significant number of firms or areas using the program. These statistical methods need a sufficient sample size of program users so that one can assume that unique factors affecting economic outcomes of program users average out over the sample. But what if an area has a unique economic development program? For example, what if the area has some unique package of economic development programs that are believed to have a synergistic effect, so that the entire effect of the package cannot be accurately predicted even if one knows the effects on economic outcomes of each individual program?[6]

If this package of economic development programs is offered to a select group of eligible firms, then firms that do not use the program can be used as a control or comparison program, using either experimental methods or non-experimental statistical analysis, depending upon what evaluation resources and will power are available. The procedures would be identical to what has previously been described. Similarly, if the package is offered to a targeted group of

[6]If there is no special synergistic effect of multiple programs, then the effects of a "unique" combination of programs can be extrapolated from studies of the effects of each individual program.

distressed areas, then distressed areas that aren't targeted can be used as a control or comparison group, as described previously.

On the other hand, what if the package is offered to all eligible firms in the entire area? In that case, then the best that any statistical analysis can say is whether the area's economic outcomes differ significantly from the average performance that one would predict, based on the performance of matched comparison areas or based on a prediction equation using characteristics of comparison areas. With only one area offering this program, its economic outcomes will have to differ quite a bit more from the average predicted for other areas for its outcomes to be statistically significantly different, compared to a situation where a sizable number of areas offer the same program. In addition, even if the area's performance is statistically significantly different from what would be predicted, all one can conclude is that the net effects of the area's unique economic development programs, and any other special characteristics of the area during this time period, result in a net effect on economic outcomes that is significantly different. Separating out what is clearly due to the unique program is impossible.

Is There Other Evidence than Statistical Comparisons with Control or Comparison Groups that Might Indicate Program Impact?

Given the difficulties and uncertainties associated with experiments or statistical comparisons of programs, it is important to consider whether there are alternative methods that can substitute or supplement for experiments or statistical comparisons, and allow us to make some inference of a link between economic development program activities and economic outcomes. I would argue that there are at least three alternative methods to link program activities and economic outcomes. First, in some cases, if one believes one can model how

different programs affect business decisions, then it may be possible to extrapolate from results obtained for other programs to new programs. For example, at current U.S. state and local business tax rates, an elasticity of state business activity with respect to state/local business taxes of -0.25, which is close to the median result in the research literature, implies that it costs roughly $9,000 annually in foregone state and local business tax revenue to create one job, or discounting at a real discount rate of 10 percent, a present value of foregone state and local business tax revenue of $90,000 (Bartik, Eisinger, and Erickcek 2003; Bartik 1992). If one is willing to assume that all that matters in an economic development subsidy is its cost, and that all cost reductions have roughly similar effects on business location probabilities, then one can infer a likely effect of different business subsidies on economic development. For example, suppose that a new branch plant with 1,000 employees is given economic development subsidies whose present value is $30 million. Then, to be fully consistent with the business tax and location literature, these subsidies would be expected to increase the odds of the branch plant choosing the state by one-third, because this effect on the location probability would yield a present value of cost per job created of $90,000 (= $30 million / [(1/3)(1000)]). Of course, in any particular case, the subsidy was either decisive in tipping the location decision or it wasn't. If one has information that makes it more or less likely that the subsidy was decisive, it should be used. But in the absence of other information, it is unclear why the effects of economic development subsidies on business location probabilities should differ from those of general state and local business taxes.[7]

[7]This assumption that all that matters are overall business costs is based on the assumption that output effects dominate in determining business location and expansion behavior, and that factor substitution effects are of secondary importance.

Second, surveys of firms receiving assistance of economic development programs can, if properly run, be used to get a rough idea of the effects of some economic development programs on business decisions. It has become very common for state and local economic development agencies to use some sort of "customer satisfaction" survey of clients. The more useful surveys, however, ask specific questions about how the assistance provided to the firm has affected its behavior. Responses to such "outcome impact" survey questions are more credible when asked for economic development assistance that is provided in the form of in-kind services rather than cash, because firms have an incentive to claim that cash assistance had an impact to keep the cash coming, whereas it is unclear why a firm would claim an in-kind service was useful if it was actually useless. A good example of economic development surveys to determine the impact of economic development services are the regular surveys of program clients of the Manufacturing Extension Partnership, whose local centers provide assistance to small and medium-sized manufacturers in improving their productivity. In the most recent MEP surveys, conducted in 2001 for clients whose projects closed a year earlier, about 64 percent of surveyed clients reported that their involvement with MEP had led to productivity improvements (NIST 2002). The average MEP client reported that the MEP services led to sales increases of $143,000 and cost savings of $50,000. It is unclear why MEP clients would seek to fabricate such responses, particularly since they were provided anonymously to a third-party survey organization.

The credibility of firms' responses to surveys about the impact of tax breaks and other financial assistance is more questionable. Firms clearly have some incentive to claim the financial assistance affected their location or expansion behavior, in the hopes of keeping the assistance going. In some cases, the program may even have required that the firm sign a

stipulation that the subsidy was essential to the location or expansion decision in order to receive a subsidy. However, it is certainly not the case that firms will always claim that financial assistance was crucial. For example, the Colorado state legislative audit agency, in an audit of the state's enterprise zone program, surveyed 18 businesses that had located or expanded in Colorado enterprise zones, and found that 10 of these businesses reported that the enterprise zone's incentives had no effect on their location or expansion decision (Hinckley and Hsu 2000). Surveys about financial assistance are more credible when administered anonymously by an independent agency, especially when the surveys ask specific and definitive questions (for example, "did you consider other locations?" or "would the location you chose have been clearly inferior in profitability to these other locations without the subsidy?")

A third method of inferring a link between program activities and local economic outcomes is determining whether administrative data on the program and its clients are consistent with the program's stated purpose. For example, the Capital Access Program in Michigan was designed to encourage banks to provide higher-risk loans to small business borrowers. For each small business loan program, the bank and borrower would each put 1.5 to 3.5 percent of the loan's value into a loan loss reserve fund, and CAP would provide a 150 percent match of the bank and borrowers' contribution to the fund. Administrative data suggested that the resulting program had a loss rate of about seven times that of a normal bank loss rate on small business loans, which means that the program—at the very least—is probably encouraging loans that otherwise would not have been made (Rohde, Cash, and Ammerman 1990). This finding is consistent with the hypothesis that the program is expanding the supply of credit to small business, although it doesn't definitely prove that the program would pass a

benefit-cost test. As another example, the MEGA program in Michigan provides very large refundable tax credits to a select group of new firms or firm expansions in Michigan, but only if the firm can present financial information showing that without the subsidy, the firm would locate outside the state. Although it is obviously possible for businesses to make up such financial data, the state is free to ask probing questions about the firms' data analysis, and refuse to provide the MEGA credits if the firm's responses are insufficiently convincing. The requirement that the firm financially demonstrate that the subsidy will be decisive at least increases the difficulties and costs for firms with a relatively weak case of applying to the MEGA program. Officials in Michigan's economic development agency claim that they screen out over 90 percent of firms expressing an interest in the MEGA program (Bartik, Eisinger, and Erickcek 2003).

Can We Determine Why and How a Program Has Impacts or Fails to Have Impacts?

One concern about outcome impact evaluations is that they are often perceived, even if done well, as only telling us whether a program works, and leaving the workings of the program a "black box": we don't know why or how the program works, so we don't know how to improve the program. In principle, statistical analysis using control or comparison groups can give insights into why and how a program works if a sufficient variation in program designs is observed and accurately measured. With data on many evaluations, statistical comparison with control or comparison groups can suggest which program designs are most effective, which, in a practical sense, is as good as knowing how or why a program works.

In the real world, however, one rarely observes a sufficient variation in program designs to adequately answer all the important questions about how or why the program works. In particular, it is impossible in principle to have data on program variations that have not yet been tried.

This suggests that surveys of clients and client focus groups may often be valuable in opening the "black box," and getting more insight into the strengths and weaknesses of the program. Statistical analysis using control or comparison groups, and surveys and focus groups, should be seen as complementary approaches to evaluating a local economic development program. The statistical comparisons are more likely to give objective quantitative evidence on the bottom line for the program—its impact on local economic conditions—whereas surveys and focus groups are more apt to give information on how that bottom line can be improved.

Can We Determine a Program's Impacts on Ultimate Rather Than Proximate Economic Objectives?

Most of the discussion so far has not considered what economic outcomes should be evaluated. In practice, the economic impacts that are easiest to evaluate are the proximate impacts on various dimensions of business activity, such as the number of business start-ups or expansions, job growth, productivity growth, etc. But public subsidies for local economic development programs cannot be justified by these programs' effects on local business activity alone. Public subsidies for local economic development require that the changes in local business activity lead to broader public benefits. The most plausible of such public benefits are fiscal benefits to state and local governments, and employment benefits to local residents.

There are possible efficiency rationales for some programs that promote increased business activity, even if there are no broader public benefits. For example, some local economic development policies can be justified by various "market failures" in information markets or financial markets (Bartik 1990). Private information markets may sometimes fail to provide businesses with information that is more valuable than the cost of providing the information, justifying public provision or public subsidy of the information. Financial markets may sometimes fail to make business loans or investments whose private return exceeds the costs, potentially justifying some public investment or subsidies for private loans or investments. But if all that these public interventions do is promote greater business activity, with no broader public benefits, it is unclear why the public at large should pay for these interventions. The business community at large would be a more justifiable source of funds.

The public receives benefits from local economic development if the increased local business activity leads to fiscal benefits or employment benefits (Bartik 1991). Fiscal benefits occur when the increased business activity, and the spinoff effects of this increased business activity on the local economy, result in tax revenue that exceeds required public expenditure increases. Employment benefits occur when the wages of the newly created local jobs exceed the "opportunity costs" of the non-working time foregone by local residents who obtain jobs because of the newly created jobs. The new jobs must either be filled by employed local residents, non-employed local residents, and in-migrants. Jobs filled by employed local residents lead to vacancies that are filled in this same way, so ultimately the newly created jobs are either filled by nonemployed local residents or in-migrants. If these jobs hadn't been created, in-migrants could have moved to another similar local area and obtained a job, so the opportunity

cost of their time is close to their wage rate. For local residents who are non-employed, the opportunity cost of their time—their "reservation wage"—may be considerably less than the wage rate of the new jobs.

To calculate fiscal and employment benefits of local economic development policies requires an economic model that takes the initial effects of the policies on local business activity, and calculates the impacts on the overall local economy, including multiplier effects on suppliers and retailers, and effects on local population growth. A variety of such regional models are commonly used by economic development agencies in the United States, most prominently the REMI model and the IMPLAN model. Once the overall impacts on all local business activity and population growth are determined, these impacts need to be translated into impacts on state and local government budgets, and local employment benefits.[8]

Fiscal impact models that translate economic impacts into budget impacts need to be specially constructed for each state or local area, given variations in the local economy and local tax and budget structure. Fiscal impacts on state and local governments depend on several factors. First, fiscal impacts depend, in part, on how different tax bases are taxed by the state and local governments. Second, fiscal impacts depend on the amount of population in-migration compared to increased business activity, as it is generally the case in the U.S. that the average business pays more in normal taxes than it directly requires in public services, whereas the average household uses more public services than it pays in taxes (Oakland and Testa 2000).

[8]For more extensive discussion of state econometric models and associated fiscal impact and employment impact models, and their use in evaluating economic development incentives, see the report by Poole, Erickcek, Iannone, McCrea, and Salem 1999.

Third, fiscal impacts depend, in part, on the share of state and local spending that goes towards purposes related to income redistribution, such as welfare or Medicaid (the United States' medical assistance program for the poor), as such spending will not respond proportionately to local economic growth. Fourth, fiscal impacts depend, in part, on whether the existing state and local infrastructure has some unused capacity that will allow economic expansion without requiring expensive construction of new infrastructure.

In the United States, positive fiscal impacts of local economic development may often be significant. For example, one recent study calculates that positive fiscal impacts of state economic development in Michigan may offset as much as half of the gross costs of the state's economic development subsidies (Bartik, Eisinger, and Erickcek 2003).

Studies of local labor markets in the United States suggests that for every one percent in extra employment growth in a metropolitan area, local employment rates increase by about 0.2 percent, and average earnings per job increase due to occupational upgrading by about 0.2, so average earnings per local resident increase by about 0.4 percent. (Bartik 1991). These earnings effects may be long-lasting if the extra employment experience for previously employed local residents increases their job skills, self-confidence, and reputation with employers (Bartik 2001). Other estimates suggests that these positive effects on local earnings are greater if the extra employment growth is concentrated in jobs that pay well relative to the skills required, such as manufacturing jobs (Bartik 1993). Theoretical models of local labor markets suggests that the reservation wages of newly employed local residents are likely to be greater in metropolitan areas with higher unemployment rates, in which the unemployed are more likely to be desperate to get a job (Bartik 1991). Some attempts to apply these local labor market models in the

evaluation of local economic development policies have been made by some state economic development agencies, most notably in New York state (Poole et al. 1999). However, we know less than we should about how the employment benefits of local economic development policies are affected by other factors. For example, we might expect that local employment benefits will be greater when there is a better match between the newly created jobs and the job skills of the local unemployed, but there is no direct empirical evidence to support this expectation. We might also expect that local employment benefits will be greater when local labor market institutions are more efficient in job training and job matching, but again there is no direct empirical evidence. More studies of the factors affecting the link between local economic development and local employment benefits would help improve the quality of evaluations of local economic development policies.

CONCLUSION

This paper has tried to show that more rigorous evaluation of economic development programs is feasible. Such rigorous evaluations can be done through random experimentation, statistical analysis of program users and comparison groups, surveys and focus groups, and linking regional econometric models with fiscal impact and local labor market models.

Such rigorous evaluations have been done extensively in the United States. These studies often find that services to small and medium-sized manufacturers can be effective in improving the performance of these firms. Programs that target distressed areas tend to be ineffective if, like enterprise zones, they provide modest resources, but are more successful if, like the Appalachian Regional Commission, they provide extensive resources over a lengthy time period.

Programs that provide financial incentives or services to encourage the economic development of a whole area, such as state tax incentives, have modest effects in increasing employment growth, which is reasonable given the modest share of costs that can be influenced by state or local governments. However, if a state or city has extensive underused infrastructure or labor, then even modest increases in job growth may offer considerable fiscal and employment benefits.

However, rigorous evaluation is still the exception rather than the rule. There are far too many cases where state and local economic development organizations claim credit for any state or local job growth, or at least for any of the growth that the organizations happen to have subsidized.

How can more rigorous evaluation be encouraged? I see little prospects of significant increases in rigorous evaluations without some outside pressure and funding. Economic development organizations, at least in the United States, face some significant disincentives in rigorously evaluating themselves. In the U.S. context, with suspicion of government activism, negative evaluations are a common excuse for terminating a program. Hence, for U.S. policymakers, the advice of economists Gary Burtless and Robert Haveman often seems reasonable "If you advocate a particular policy reform or innovation, do not press to have it tested." (Burtless and Haveman 1984, p. 128).

Therefore, I was not surprised a few years ago when one state economic development official told me that the trouble with universities evaluating his programs is that the universities seemed to think that negative evaluations should be made public. One can deplore this attitude, but should understand the real fear of budget cuts and program extinction that motivates it.

Rigorous evaluation of economic development policies is only likely to occur if funded or required by outside groups. These outside groups could include legislatures, governmental audit bureaus, and higher levels of government. It makes sense for these groups to require and pay for evaluations, because the benefits of evaluations largely accrue outside of the agency managing the program. Evaluations of a state or local agency's economic development programs provide benefits to the general public in the local area, who benefit from any improvements that result in government effectiveness, and to local areas elsewhere that have similar programs or might consider similar programs.

Rigorous evaluation is also more likely to occur if the results are more frequently used to improve programs rather than kill the programs. If the basic rationale of the program makes sense, in that the program is addressing some problem that may benefit from government intervention, then negative evaluation results should be used to motivate the creation of a new approach to addressing the problem. This is more likely if the rigorous evaluation is accompanied by data from surveys and focus groups that give some insights into how the program can be improved. A balanced mix of rigorous and "softer" evaluation techniques, and a judicious use of evaluation results, will encourage economic development agencies to be more open to rigorous evaluation approaches.

REFERENCES

Angrist, Joshua D., and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15(4): 69–85.

Bartik, Timothy J. 1990. "The Market Failure Approach to Regional Economic Development Policy." *Economic Development Quarterly* 4(4): 361–370.

———. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

———. 1992. "The Effects of State and Local Taxes on Economic Development: A Review of Recent Research." *Economic Development Quarterly* 26: 102–110.

———. 1993. "Economic Development and Black Economic Success." Upjohn Institute Technical Report no. 93-001. Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research.

———. 2001. *Jobs for the Poor: Can Labor Demand Policies Help?* New York and Kalamazoo, Michigan: Russell Sage Foundation and W.E. Upjohn Institute for Employment Research.

Bartik, Timothy J. (with Peter Eisinger and George Erickcek). 2003. "Economic Development Policy in Michigan." In *Michigan at the Millennium*, Charles Ballard, Paul Courant, Doug Drake, Ron Fisher, and Elizabeth Gerber, eds. East Lansing, Michigan: Michigan State University Press.

Bartik, Timothy J., and Richard D. Bingham. 1997. "Can Economic Development Programs be Evaluated?" In *Dilemmas of Urban Economic Development*, Richard D. Bingham and Robert Mier, eds. (pp. 246–277). Thousand Oaks, California: Sage Publications, Inc.

Benus, J. M., M. Wood, and N. Grover. 1994. *A Comparative Analysis of the Washington and Massachusetts UI Self-Employment Demonstrations*. Unpublished report prepared for the U.S. Department of Labor, Employment and Training Administration, Unemployment Insurance Service under Contract No. 99-8-0803-98-047-01.

Bondonio, Daniele, and John Engberg. 2000. "Enterprise Zones and Local Employment: Evidence from the States' Programs." *Regional Science and Urban Economics* 30: 519–549.

Burtless, G., and R. H. Haveman. 1984. "Policy Lessons from Three Labor Market Experiments." In *Employment and Training R&D*, R. Thane Robson, ed. Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research, p. 128.

Fisher, Peter S., and Alan H. Peters. 1998. *Industrial Incentives: Competition Among American States and Cities*. Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

Greenbaum, Robert T. 1998. "An Evaluation of State Enterprise Zone Policies: Measuring the Impact on Business Decisions and Housing Market Outcomes." Doctoral dissertation, Carnegie Mellon University, December.

Greenbaum, Robert T., and John B. Engberg. 1998. "The Impact of State Urban Enterprise Zones on Business Outcomes." Discussion paper no. 98-20. Washington, DC: Center for Economic Studies, Bureau of the Census.

Hebert, Scott, Avis Vidal, Greg Mills, Franklin James, and Debbie Gruenstein. 2001. *Interim Assessment of the Empowerment Zones and Enterprise Communities (EZ/EC) Program: A Progress Report*. Washington, DC: Office of Policy Development and Research, U.S. Department of Housing and Urban Development.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64(4): 605–654.

Heckman, James J., and Edward Vytlacil. 2001. "Policy-Relevant Treatment Effects." *The American Economic Review* 91(2): 105–111.

Hines, James R., Jr. 1996. "Altered States: Taxes and the Location of Foreign Direct Investment in America." *American Economic Review* 86(5): 1076–1094.

Hinkley, Sara, and Fiona Hsu. 2000. "Minding the Candy Store: State Audits of Economic Development." Washington, DC: Good Jobs First, September.

Holzer, H. J., R. N. Block, M. Cheatham, and J. H. Knott. 1993. "Art Training Subsidies For Firms Effective? The Michigan Experience." *Industrial and Labor Relations Review* 46 (July): 625–636.

Imbens, Guido, and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467–475.

Isserman, Andrew, and Terance Rephann. 1995. "The Economic Effects of the Appalachian Regional Commission: An Empirical Assessment of 26 Years of Regional Development Planning." *Journal of the American Planning Association* 61(3): 345–364.

Jarmin, Ronald S. 1999. "Evaluating the Impact of Manufacturing Extension on Productivity Growth." *Journal of Policy Analysis and Management* 18(1): 99–119.

Mazerov, Michael. 2001. *The "Single Sales Factor" Formula for State Corporate Taxes: A Boon to Economic Development or a Costly Giveaway?* Washington, DC: Center on Budget and Policy Priorities.

McLure, Charles E., Jr., and Walter Hellerstein. 2002. "Does Sales-Only Apportionment of Corporate Income Violate the Gatt?" NBER working paper no. 9060. Cambridge, Massachusetts: National Bureau of Economic Research.

Meyer, Bruce D. 1995. "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistics* 13(2): 151–161.

Molnar, Lawrence A., Donald R. Grives, Jack Edelstein, Rocco DePietro, Hugh Sherman, Dinah Adkins, and Lou Tornatzky (with Yolanda Batts, George Fulton, and Sally Hayhow). 1997. *Impact of Incubator Investments*. Athens, Ohio and Ann Arbor, Michigan: NBIA Publications and The University of Michigan.

Murnane, R. J., S. Newstead, and R. J. Olsen. 1985. "Comparing Public and Private Schools: The Puzzling Role of Selectivity Bias." *Journal of Business and Economic Statistics* 3:23–35.

National Institute of Standards and Technology. 2002. "The Manufacturing Extension Partnership: Delivering Measurable Returns to Its Clients." Washington, DC: U.S. Department of Commerce, Technology Administration.

Nexus Associates. 1999. *A Record of Achievement: The Economic Impact of the Ben Franklin Partnership*. Belmont, Massachusetts: Nexus Associates, Inc.

Oakland, William H., and William A. Testa. 2000. "The Benefit Principle as a Preferred Approach to Taxing Business in the Midwest." *Economic Development Quarterly* 14(2): 154–164.

Papke, L. E. 1993. "What Do We Know About Enterprise Zones." In *Tax Policy and Economy, Vol. 7*, J. M. Poterba, ed. Cambridge, Massachusetts: MIT Press, pp. 37–72.

———. 1994. "Tax Policy and Urban Development: Evidence from the Indiana Enterprise Program." *Journal of Public Economics* 54: 37–49.

Peters, Alan H., and Peter S. Fisher. 2002. *State Enterprise Zone Programs: Have They Worked?* Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

Peterson, Paul E. 1981. *City Limits*. Chicago: University of Chicago Press.

———. 1995. *The Price of Federalism*. Washington, DC: The Brookings Institution.

Poole, Kenneth E., George A. Erickcek, Donald T. Iannone, Nancy McCrea, and Pofen Salem. 1999. *Evaluating Business Development Incentives*. Washington, DC: National Association of State Development Agencies.

Reed, W. Robert, and Cynthia L. Rogers. 2000. "Measurement Error and Endogeneity in Studies of State Tax Policy and Economic Growth." Unpublished paper, Department of Economics, University of Oklahoma, Norman, Oklahoma.

Rogers, Cynthia, and W. Robert Reed. Forthcoming. "Quasi-Experimental Control Group Methods for Estimating Policy Impacts." *Regional Science and Urban Economics*.

Rohde, S., J. Cash, and K. Ammarman. 1990. *Study of the Capital Access Program*. Unpublished working paper, Michigan Strategic Fund, Lansing, Michigan.

Smith, Jeffrey A., and Petra E. Todd. 2001. "Reconciling Conflicting Evidence on the Performance of Propensity-Score Matching Methods." *The American Economic Review* 91(2): 112–118.

Thomas, Kenneth P. 2000. *Competing for Capital: Europe and North America in a Global Era*. Washington, DC: Georgetown University Press.

Wasylenko, Michael. 1997. "Taxation and Economic Development: The State of the Economic Literature." *New England Economic Review* (March/April): 37–52.

Table 1 A Classification of Local Economic Development Policies

1. **Selective Firm-Assistance** (predominantly to small and medium-sized enterprises)
- Training in how to start-up or manage a business
- Public loans/investments or public support for private loans/investments for business start-ups or expansions
- Information/training on implementing new technology or new management techniques
- Firm- or industry-customized training for new workers
- Information/training on exporting

2. **Distressed Area Assistance** (enterprise zones and other programs that are typically designed and designated by higher levels of government)
- Tax breaks in local and higher-level government taxes for firms locating or expanding in the designated area
- Enhanced services or infrastructure in the designated area, whether firm-specific or general

3. **Whole Area Programs** (typically targeted at manufacturers or other "export-based" firms; sometimes targeted to particular industries)
- Marketing an area and providing site information to new branch plant prospects
- Providing existing businesses and new businesses with help in resolving government regulatory problems
- Expedited provision of site-specific roads and utilities for new plants or expansions, or previous development of industrial parks
- Tax incentives for new or expanded branch plants or corporate headquarters
- Firm-customized training for new workers as incentive for new corporate facilities or expansions
- Support for networks or clusters of firms in an industry to develop better support services such as training
- Technology or industry twist to any of above programs, for example technology-oriented industrial parks, or tax incentives, or training