



2005 Market Street, Suite 2800 P 215.575.9050
Philadelphia, PA 19103-7077 F 215.575.4939

901 E Street NW, 10th Floor P 202.552.2000
Washington, DC 20004 F 202.552.2299
pewtrusts.org

Memo

To: Chair Benavidez and members of the Tax Expenditure Evaluation Interim Study Committee

From: Alison Wakefield and Khara Boender, The Pew Charitable Trusts

Date: August 6, 2019

Subject: Follow-up to Tax Expenditure Evaluation Interim Study Committee Meeting

In response to discussions during the July 23, 2019 Tax Expenditure Evaluation Interim Study Committee meeting, we provide examples of state approaches to the following topics:

- Approaches to answering the “but-for” question;
- Establishing goals for tax incentives;
- Evaluating incentive-related job and economic impacts between different regions within states.

Approaches to answering the “but-for” question

Considering the extent to which tax incentives changed behavior is crucial to measuring their economic impact. Tax incentives provide economic benefits to states to the extent they encourage businesses to create new jobs and make new investments. If a company would have acted the same way even without the incentives, the state is no better off.

While it is difficult to know exactly what happened because of incentives, numerous states have successfully come to reasonable estimates of the extent to which the programs are changing business behavior. These evaluations have used a variety of different techniques, which are described below.

Sensitivity analysis:

Some evaluations have calculated results using multiple scenarios of cause-and-effect (e.g. 30% of business activity occurred because of the incentive or 50% of business activity occurred because of the incentive). Evaluators can then analyze which scenario is most likely.

A 2014 Connecticut [evaluation](#) used this approach. For several of the state’s incentives, the analysts modeled four scenarios attributing varying degrees of business activity to the incentive: 100%, 50%, 20%, and 0% (page 4).

Utilizing academic literature on the responsiveness of businesses to cost reductions:

Academics and other economic researchers have generated an extensive literature around the extent to which state and local taxes affect business activity. As a result, if an analyst knows the extent to which incentives are reducing business taxes (or business costs more generally), then he or she can use the economic literature to estimate the extent to which the incentives change business behavior.

The Minnesota Legislative Auditor's [evaluation](#) of the state's Job Opportunity Building Zone program is one example of this approach. The evaluation estimated that 21% of the new jobs at participating businesses were a result of the program (pages 94-95). Their basis for this estimate was "reasonable assumptions from the literature on tax incentives on the responsiveness of employment to tax incentives." An [evaluation](#) of Tennessee's Job Tax Credit used a similar approach (pages 36-38).

An [evaluation](#) in Kentucky used an interesting twist of this strategy (pages 97-108). Rather than starting by estimating the extent to which incentives changed business behavior, the analysts started by measuring the extent to which the incentive would *need* to change business behavior to be superior to the estimated outcomes of an alternative policy. For this purpose, analysts compared the tax incentive to reductions in the corporate income tax rate (though other tax reductions or spending increases could be used as the alternative scenario as well). This comparison included a calculation of what they called the "threshold effectiveness," defined as the minimum level of job creation needed for the tax incentive to be more effective than the alternative policy. For example, if the threshold effectiveness was 40%, this means that the tax incentive had to be responsible for at least 40% of new job creation at incentivized firms for the policy to be more effective than the alternative. The Kentucky evaluators then consulted academic literature on the responsiveness of businesses to cost reductions to help assess whether it was plausible to think the incentives were better than the alternative.

Project-specific financial analysis:

The question of cause-and-effect often comes down to whether incentives are large enough to matter. The evaluations from Minnesota, Tennessee, and Kentucky mentioned above study this question on a programmatic level: Overall, how large is the cost reduction for businesses participating in the program? However, another option is to analyze the cost reduction for individual companies or projects and then use the findings to draw conclusions about the effectiveness of the program as a whole.

In Oregon, consultants [studied](#) tax credits for renewable energy projects such as wind and solar farms. To do so, the consultants examined what return on investment would make the various types of projects financially viable. These feasibility thresholds recognized that whether an energy project makes sense for a company depends on the amount of energy a project can produce, the cost of producing it, and the price that buyers will pay. Then, using financial models for representative companies (e.g. a utility-scale wind farm), they studied whether the incentives were large enough to make otherwise unfeasible projects financially viable. This approach allowed the evaluation to conclude that smaller-scale projects needed the incentives in order to go forward, but larger ones likely did not. Later, lawmakers amended the program to focus the program more on smaller-scale projects, an outcome that was consistent with the findings of the evaluation.

Comparing incentivized companies to control groups:

In theory, the ideal way to determine whether incentives are successfully influencing business decisions would be to approach the question like a science experiment: by comparing the companies that received incentives to a control group that did not, but that was otherwise identical. In the real world of economic development, it's rarely possible to conduct that type of experiment. Occasionally, though, the design of an incentive is conducive to a similar analytic approach.

Under one Michigan program, companies apply, the state scores each application, and then companies above a certain score receive incentives. An [evaluation](#) (page 9) compared the performance of companies that scored just barely above the threshold and got the incentives to those that scored just below it and didn't—the idea being that these two groups of companies were basically the same, except that some received incentives and others didn't. The evaluation found that the incentives did make a difference in improving certain measures of company performance.

In other cases, analysts have used statistical methods, such as regression analysis, to make similar comparisons even if beneficiaries have not been randomly assigned. For example, a 2015 Iowa Department of Revenue [study](#) of the Beginning Farmer Tax Credit compared beneficiaries of the program to a control group of farmers who did not participate to help estimate to what extent the incentive was helping farmers succeed (pages 35-40). The evaluation used a variety of statistical methods (t-tests, regression analysis, and chi-squared tests) to assess whether the different outcomes between program participants and the control group were statistically significant and a result of the incentive. Based on this analysis, they conclude that “in some respects, program participants became more established in farming between 2008 and 2013 than did members of the comparison group.”

Business surveys:

Surveys of program participants are sometimes parts of effective evaluations, but they need to be carefully designed. It's often not especially useful to ask businesses directly whether they value the incentives they receive—most will almost certainly say yes—but more thoughtful questions can reveal important information about how the programs are influencing decisions.

An [evaluation](#) of Minnesota's angel investor tax credit included a survey of participating investors and businesses. One thing the survey showed is that many program participants were “inside investors” who had a personal stake in the companies in which they were investing. The evaluation concluded that these inside investors—often executives or board members—were less likely to need the incentives to encourage them to invest than were venture capitalists not affiliated with the companies. In response, the Legislature limited the ability of inside investors to participate in the program.

Common sense assumptions:

Massachusetts' [evaluation](#) of its film tax credit is a good example of a state considering what was taking place before the incentive was offered to help determine how much economic activity can be attributed to the incentive. On page 6, the study notes:

Because this report is attempting to measure only new Massachusetts economic activity that results from the film production tax incentives, we do not include economic activity that, while eligible for the film tax incentives, was already taking place before the tax incentives were implemented and presumably would have continued to take place had the incentives not been enacted.

...

We identified long-running shows and specifically local programming that claimed the tax incentives, and assumed that these would have continued to be produced even without the incentives. For the most part, these consisted of educational, public affairs, and sports-themed productions connected to long-established local institutions.

The evaluation also goes a step further by considering the possibility that some of the new activity would still have taken place regardless of the tax credit. For example, they assume that new locally-themed documentaries would have been filmed in Massachusetts regardless of the incentive (page 7) and therefore do not consider economic benefits from these productions as part of their analysis. This seems like a reasonable common-sense assumption—it's not likely that a documentary about Boston would be filmed anywhere other than Massachusetts regardless of the availability of incentives.

Establishing goals for tax expenditures

A clear, measurable goal helps policymakers understand how the proposed incentive is intended to benefit the state economy. A very general goal such as “encouraging economic development” won't facilitate a conversation about the incentive nor does it provide guidance on how the results of the incentive could be evaluated. In contrast, including a specific, measurable goal carries multiple benefits. First, lawmakers can determine if the new program properly aligns with a state's current economic development objectives. Second, a clear goal helps lawmakers track progress and is a critical factor in evaluating the program's effectiveness.

A 2013 Vermont law set up a process to draft proposed goals for each of the state's existing tax credits, exemptions, and deductions. To do so, legislative staff studied the statutory description of the incentives and available legislative records. The following year, lawmakers adopted [goals](#) for each program based on this research. In this way, Vermont developed purpose statements for all tax incentives, rather than take a piecemeal approach.

Between 2012 and 2013, Nebraska's Legislative Audit Office (LAO) released a series of reports examining the state's tax incentives. Because the programs lacked measurable goals, LAO's ability to assess the success of the state's incentives was impaired. In a later report, LAO suggested possible goals for some of the state's incentive programs based on a review of goals as stated in statutes, statements made during legislative debates when the programs were created, and other related legislative actions. LAO also surveyed non-governmental groups to hear their perspective on appropriate goals. That report helped encourage lawmakers to clarify the objectives of Nebraska's major incentives.

Definitions of key terms is another valuable component. LAO's evaluation of Nebraska's Advantage Act was sometimes hampered by a lack of clarity on the meaning of key terms. For example, the LAO was

charged with analyzing whether the Advantage Act was helping to revitalize distressed areas. As the report explained, however, Nebraska law includes two separate definitions of distressed areas and the LAO lacked guidance on which definition to use. Under one definitions, more than 99% of Nebraskans live in distressed areas, while the other definition is far less inclusive.

In 2013, Washington enacted a [law](#) requiring that proposed tax incentives include a “performance statement” designed to help the state’s Joint Legislative Audit and Review Committee (JLARC) —which evaluates incentives on a 10-year schedule—determine whether the programs in question succeeded. Each statement documents the purpose of the tax incentive, how the state will know whether the program accomplished its goal, and what data evaluators will need to conduct the review.

From Washington’s [law](#):

(1) As provided in this section, every bill enacting a new tax preference must include a tax preference performance statement.

(2) A tax preference performance statement must state the legislative purpose for the new tax preference. The tax preference performance statement must indicate one or more of the following general categories, by reference to the applicable category specified in this subsection, as the legislative purpose of the new tax preference:

(a) Tax preferences intended to induce certain designated behavior by taxpayers;

(b) Tax preferences intended to improve industry competitiveness;

(c) Tax preferences intended to create or retain jobs;

(d) Tax preferences intended to reduce structural inefficiencies in the tax structure;

(e) Tax preferences intended to provide tax relief for certain businesses or individuals; or

(f) A general purpose not identified in (a) through (e) of this subsection.

(3) In addition to identifying the general legislative purpose of the tax preference under subsection (2) of this section, the tax preference performance statement must provide additional detailed information regarding the legislative purpose of the new tax preference.

(4) A new tax preference performance statement must specify clear, relevant, and ascertainable metrics and data requirements that allow the joint legislative audit and review committee and the legislature to measure the effectiveness of the new tax preference in achieving the purpose designated under subsection (2) of this section.

(5) If the tax preference performance statement for a new tax preference indicates a legislative purpose described in subsection (2)(b) or (c) of this section, any taxpayer claiming the new tax preference must file an annual survey in accordance with RCW 82.32.585.

Instead of relying on intuition or legislative history to determine lawmakers' intentions, JLARC now uses the specific criteria a tax preference proposal sets out in its performance statement to determine whether or not an incentive is successful. In addition to helping JLARC produce high-quality analyses, the performance statements have also encouraged lawmakers and advocates for incentives to think carefully about the design of proposed preferences and what they are trying to achieve.

In 2010, the Minnesota Legislature passed [legislation](#) requiring a "statement of intent" in any legislation to create, renew, or continue a tax expenditure. As Pew has found in the course of its research, "to determine the effectiveness of tax incentives, national studies recommend setting goals and objectives for the incentives." Without them there is no standard against which evaluation results may be compared to determine if they are meeting, surpassing, or falling short of expectations.

[Evaluating incentive-related job and economic impacts between regions within states](#)

When states provide incentives to some businesses, other businesses, such as their competitors, may be harmed. This is often an important consideration for evaluations, especially when states provide incentives to businesses that have in-state competitors. Displacement is a measure that does not necessarily require economic modeling to estimate. A Louisiana [evaluation](#) of the state's Enterprise Zone program measured "net" new jobs to take into account the effects of the incentive on the state's economy as a whole. The evaluation found that in certain economic sectors—retail, restaurants, hotels, and health care—many of the jobs for which companies received incentives were likely to have come at the expense of existing Louisiana jobs in those same sectors. As a result, while companies participating in the program reported creating more than 9,000 jobs, the evaluation found that the program had only created around 3,000 *net* jobs.

In a [2009 study of TIFs in the St. Louis region](#), researchers from the University of Missouri - St. Louis found that TIFs lead economic activity to shift from areas not using TIFs to TIF-using areas. This finding was echoed in a [2011 report on St. Louis TIFs](#), which found that while positive for the incentive-using municipality, TIFs had negative impacts on sales tax revenue in neighboring municipalities. Perhaps because neighboring localities feel the need to remain competitive, [a separate study on Missouri TIFs](#) found that TIF adoption in one Missouri city makes adjacent cities 2.5 times as likely to respond with TIFs of their own.

Minnesota's Job Opportunity Building Zones (JOBZ) program was designed to help economically distressed communities, so the evaluators needed to determine whether the areas being impacted were truly disadvantaged. Their [evaluation](#) of the program assessed whether the zone designations had been effectively targeted to areas in need (pages 39-46) by comparing measures of distress inside and outside of the zones. These measures of distress included: 1) the percentage of the population living under 200% of the federal poverty level; 2) the average unemployment rate from 2000 through 2003; 3) the percentage change in population between 1982 and 2002; 4) median household income for 1999; and 5) median sales prices for existing homes in 2001 and 2002 (page 42). The evaluators concluded that more prosperous areas were just as likely to have received zone designations as less prosperous ones. They

explained that part of the reason for this was that the initial zone selection process was flawed (page 31).