The use of literature-based innovation output indicators for research evaluation

By: Albert N. Link

Link, A.N. The use of literature-based innovation output indicators for research evaluation. *Small Business Economics* **7,** 451–455 (1995). https://doi.org/10.1007/BF01112465

Abstract:

This paper discusses the appropriateness of literature-based innovation output indicators in the evaluation arena, and it sets forth criteria against which to judge such appropriateness.

Keywords: arena | industrial organization | research evaluation | innovation output | output indicator

Article:[1]

I. Introduction

Fundamental to the evaluation of any research program is an ability to measure the output(s) from the research. Many of the traditional evaluation methods are based on analytical frameworks that require at least an informal quantification of such outputs.[2]

The evaluation of research is important for both public and private policy reasons. Accountability is critical in the public sector where research is often funded with public resources. Accountability is equally critical in the private sector where stakeholders' investments are being allocated in pursuit of a research strategy vis-a-vis alternative competitive strategies. In small-sized firms, in particular, resources are limited and strategic diversification is not always possible, thus making the evaluation process even more important. In addition, evaluation is a *per se* management tool. The evaluation process, and the results from that process, can provide insights and information that can improve both managerial efficiency and the scientific judgements of those involved in both funding and conducting the research that leads to innovation.

As Kleinknecht and Bain (1993, p. vii) have correctly noted:

---

[1] This paper was motivated by discussions with Zoltan Acs regarding new concepts in innovation output measurement and the implications of those concepts for small firms.
[2] See Link (1993).

> We know a great deal about inputs (or presumed inputs) such as expenditures on R&D but there have been, to date, no means of measuring how effectively such [research] spending is translated into actual innovation.

This observation by Kleinknecht and Bain is especially important in today's environment of collaborative research relationships. Evaluation is growing in importance as governmental agencies are becoming increasingly involved with the private sector in jointly-conducted research endeavors. Both parties in the collaborative relationship are finding it imperative to have at their discretion sound evaluation methods, and the soundness of these methods depends in large part on the accuracy with which their research output is measured and then evaluated.

The purpose of this paper is two-fold. First, it attempts to fill a conspicuous void in the research evaluation literature by offering a taxonomy of alternative evaluation methods; and second, it describes the potentially important role of literature-based innovation output indicators – as proposed by Kleinknecht and Bain (1993) and illustrated through the papers in their edited volume, *New Concepts in Innovation Output Measurement* – within the broader context of evaluation methodology.

II. How to evaluate research

Research evaluation is an art; not a science. A number of evaluation methods have been used by both public and private sector policy makers, as well as by academicians, often with an eye toward comparing the output of one research program against that of another. There are strengths and weaknesses associated with all evaluation methods. The evaluation process is inherently subjective, although some methods can be more technical than others. There is no single method that is applicable to all research situations.

Research evaluation methods can be classified into one of two broadly defined frameworks: a framework that utilizes a technical approach, or one that utilizes a normative approach.[3] The distinction between technical and normative evaluation approaches is based on how one views the underlying information used in the evaluation process. The underlying information can be technical in its nature, or normative judgements can be imposed on the interpretation of the underlying information.

1. Technical approaches to evaluation

Technical approaches to research evaluation are implemented on the assumption that there are technical criteria upon which to judge the outcomes of research. If there are such fixed objectives, then normative interpretations are removed when conducting the evaluation. Such evaluations are formulated on either a *technology-based evaluation standard* or a *cost-effective evaluation standard.*

Technology-based evaluation standards are generally applied in a comparative way, that is evaluating the output from one research project against the output from another. The question

---

[3] See, in particular, Lave (1981) and Smith (1986). See Smith for particular references of the application of these decision models.

asked by the evaluator is: Has the "best" technological result been obtained from a research project? Under this evaluation criterion, "best" is generally defined in the engineering sense of being state-of-the-art. The primary advantage of applying this standard is that neither quantifiable information on benefits nor quantifiable information on costs is needed to conduct the evaluation. All that is required is consistent engineering judgement. The primary disadvantage is that "best" is a subjective concept, even to engineers who are familiar with the technological dimensions of the innovative effort.

Cost-effective evaluation standards are also used for comparative purposes. The question asked is: Have the technical objectives of the research been met in the most cost-effective way? An advantage of this approach is that there is no need to measure benefits. Benefits are defined simply in terms of technical completion, that is "completed" or "not completed". Given technical completion, then the evaluation criterion is based on the single dimension of cost effectiveness. The primary disadvantage of this standard is that no consideration is given to the usefulness or applicability of the resulting innovations.

2. Normative approaches to evaluation

In contrast to technical approaches to research evaluation, normative evaluation approaches are based on the assumption that there are trade-offs among the evaluation-related dimensions of a research project. And, normative judgements are inherent when making such tradeoffs. Three common evaluation methods are *multiple programming analysis, economic impact analysis, and benefit-cost analysis.*

Multiple programming analysis requires the development and use of a mathematical model to rank projects based on a pre-established set of criteria, such as the objectives of the research, the cost constraints on the research, and/or the outcomes of the research. A strength of this method is its objectivity; however, a weakness is that this objectivity is only as good as the information programmed into the model. In other words, mathematical models are limited in their ability to "accept" normative considerations.

Economic impact analysis employs a variety of methods for associating research results with economic variables, such as productivity growth, employment, prices, economic development, competitiveness, and so on. These methods range from traditional operations research techniques, to production function applications, to input-output mappings. The primary advantage of this valuation technique is that it associates, within accepted statistical frameworks, the results from the research with objective economic variables. The primary disadvantage is that the cost of conducting the initial research is not considered in the analysis. Economic impact analysis can be used to evaluate a given research program, or to compare across programs.

Benefit-cost analysis relates the benefits, or outputs, from a research project to the cost of the research project. The benefits that result from a project fall into two broad categories, tangible benefits and intangible benefits. Tangible benefits are measurable, whereas intangible benefits are not. Within each category, benefits can be further disaggregated into the sub-categories of direct and indirect. Direct benefits flow directly to the stakeholders of the organization, while indirect benefits, like any externality, flow to third-parties. There are tangible and intangible costs as welt, and both have a direct and indirect component. In practice, benefit-cost analysis compares tangible direct and indirect benefits to tangible direct cost. The primary strength of this evaluation approach is that it is an easily understood tool, and one that can be utilized to compare across projects or to benchmark only one project. The primary weaknesses of benefit-cost analysis are that it requires a significant amount of information to implement, and that intangible benefits are often ignored because of measurement problems.

III. Selection of an evaluation method

The selection of an evaluation method should be based on objective criteria. While there is not a generally agreed upon set of criteria, several broad guidelines can be inferred from the literature and from practice.[4]

Guideline One: The evaluation method should be appropriate.[5] Because evaluation is an art rather than a science, there is a trade-off between the choice of a method that is sophisticated in concept and rigorous in design and a method that is appropriate and doable for the evaluation under consideration, The selection of an evaluation method should therefore be made only after the research being evaluated is fully understood. Associated with appropriateness is the consideration that the evaluation must be done within resource constraints, including time and financial constraints.

Guideline Two: The evaluation method should be complete. If there is more than one appropriate evaluation method, the one that has the capacity to quantify most completely the results from the research is the one preferred. Associated with completeness are metric-related components that are welt accepted in the evaluation community and easily understood both by those being evaluated and those who will use the evaluation results.

Guideline Three: The evaluation method should be replicable. Quite often research activities continue over time beyond the date of evaluation. As such, it may be desirable at a future point in time to conduct follow-on evaluations. When done, the evaluators should have the opportunity to replicate the methodology used in the earlier evaluation.

IV. The importance of literature-based innovation output indicators

As can be seen from the taxonomy of research evaluation methods set forth above, there is no uniformly-used, much less uniformly-accepted evaluation metric. The evaluation methods discussed differ primarily is the way that they measure research output. Some rely on technical judgements while others rely on more traditional economic-based data. That fact aside, the taxonomy clearly indicates that fundamental to the evaluation of research or research programs is the ability to measure the output of the research, and the art currently practiced for the measurement of research output has its limitations.

Traditional output indicators include:

---

[4] See Lave (1981), and Ormala (1989).
[5] According to Lave (1981, p. 27), "The felicitousness of the [evaluation] framework is more important that it comprehensiveness".

(i) patent counts,
(ii) productivity growth indices, and
(iii) major innovation counts.

The problems associated with patent data have been carefully reviewed by others.[6] Productivity-based studies have been shown to be useful when attempting to quantify an aggregate rate of return to R&D spending.[7] Kleinknecht and Bain (1993) contend that counting and then evaluating major innovations have a number of important measurement advantages. One advantage is that such measurements can facilitate international comparisons of the output from innovation programs, and another advantage is that the data are readily available and thus verifiable. One disadvantage, however, is that while such literature-based innovations can be objectively counted, they can only be subjectively valued in importance.

Nevertheless, were the criteria for the selection of a research evaluation method to be applied to the selection of an output indicator, literature-based measures would score welt. They are *appropriate* in the sense of being countable, and even valued given sufficient time and resources. They are *complete* in that they do act as a market test for success in the innovation process. And, they are *replicable* in that they come from verifiable sources.

In terms of the taxonomy of methods for evaluating research, literature-based innovation output indicators are applicable to both the technical and normative approaches reviewed above. Regarding the technical approaches, literature-based innovation output indicators can be used to validate, based on experts' opinions of "best", why the "best" research project is so judged from an engineering perspective. Regarding the normative approaches, literature-based innovation output indicators can be used as an independent economic impact measure as well as measurable component in a benefit-cost analysis.

More important than the conceptual fit of literature-based innovation output measures to either a taxonomy of evaluation methods or to criteria for selecting an evaluation method is the usefulness of such measures for analyses of the innovation process. The major contribution made by Kleinknecht and Bain (1993) in their edited volume, *New Concepts in Innovation Output Measurement,* is to provide specific examples of this usefulness of innovation output indicators.

The papers in their volume contribute to the innovation literature by offering insights into some important (and previously researched) issues. In particular, the papers in the volume illustrate:

- a general methodology for identifying and using appropriate literatures to assemble data on innovations;[8]
- the consistency with which R&D investments (inputs) can be related to counts of significant innovations (outputs), thus providing additional statistical insights into the stages of the overall innovation process;[9]

---

[6] See, in particular, Acs and Audretsch (1993).
[7] See, in particular, Link (1987).
[8] In particular, see Steward (1993).
[9] In particular, see Acs and Audretsch (1993).

- the various inter-industry and inter-sectoral patterns in innovative behavior as measured by the origin of published innovations;[10]
- the alternative metrics relevant for classifying the technical complexity of innovations, and thus for classifying the underlying R&D activity (given the Acs and Audretsch (1993) findings);[11]
- the pattern of inter-sectoral flows of innovation; and[12]
- the consistency with which innovation count data can be used for cross-country comparisons of technological advancements.[13]

With no intent to diminish the importance of these contributions to the innovation literature – and these contributions are significant especially given the multi-country focus of the assembled papers – it should be emphasized, however, that Kleinknecht and Bain (1993) fall short in their ability to demonstrate the relevance, much less usefulness, of literature-based innovation output indicators for the evaluation of private or public research activities.

Kleinknecht and Bain would be urged by this writer to respond to this observation of an area of omission by pointing out that the application of such indicators to evaluation methodology was beyond the scope of questions that they intended to address.[14] This writer would definitely agree, but would retort by expressing the opinion that perhaps the most important issue facing innovation process analysts (which includes researchers) relates to how one demonstrates the economic importance of research activity.

Not only is the evaluation issue of growing concern to private sector managers as they increasing are forced to meet technology-based competitive pressures, but also it is becoming a standard practice in many industrial nations for policy makers to document the economic impacts of public-sector research funding.[15] In fact, from this writer's perspective, economists have not begun to focus their creative talents toward this application area to the same degree as scholars in other disciplines. One might hope that evaluation methodology, and a critical part of it is how one quantifies research output, does capture the attention of economic researchers since they are generally so willing and able to offer proscriptions on competitive strategies and economic growth initiatives. When the issue of research evaluation does capture the profession's attention, one might hope that the research participants pay as close attention to the details of output measurement as have the contributors to the Kleinknecht and Bain (1993) volume.

Time will tell how well literature-based innovation output indicators perform in the evaluation arena. Although their use has been widespread for investigating other aspects related to the innovation process, as carefully noted in Kleinknecht and Bain (1993), their acceptance in public sector and private sector evaluation efforts remains to be seen.

---

[10] In particular, see Kleinknecht, Reijnen, and Smits (1993).
[11] In particular, see Fleissner, Hofkirchner, and Pohl (1993).
[12] In particular, see Cogan (1993).
[13] In particular, see Kleinknecht (1993).
[14] See Kleinknecht and Bain (1993, pp. 2-3).
[15] For example, the Government Performance and Results Act of 1993 (Public Law 103-62) was enacted in the United States.

References

Acs, Zoltan J. and David B. Audretsch, 1993, ‘Analyzing Innovation Output Indicators: The U.S. Experience’, in Kleinknecht A. and D. Bain (eds.),*New Concepts in Innovation Output Measurement*, New York: St. Martin's Press.

Cogan, D. J., 1993, ‘The Irish Experience with Literature-based Innovation Output Indicators’, in A. Kleinknecht and D. Bain (eds.),*New Concepts in Innovation Output Measurement*, New York: St. Martin's Press.

Fleissner, P., Wolfgang Hofkirchner and Margit Pohl, 1993, ‘The Austrian Experience with Literature-based Innovation Output Indicators’, in A. Kleinknecht and D. Bain (eds.),*New Concepts in Innovation Output Measurement*, New York: St. Martin's Press.

Kleinknecht, A., 1993, ‘Testing Innovation Indicators for Postal Surveys: Results from a Five-County Project’, in A. Kleinknecht and D. Bain (eds.),*New Concepts in Innovation Output Measurement*, New York: St. Martin's Press.

Kleinknecht, Alfred and Donald Bain (eds.), 1993,*New Concepts in Innovation Output Measurement*, New York: St. Martin's Press.

Kleinknecht, Alfred, Jeroen O. N. Reijnen and Wendy Smits, 1993, ‘Collecting Literature-based Innovation Output Indicators: The Experience in the Netherlands’, in A. Kleinknecht and D. Bain (eds.),*New Concepts in Innovation Output Measurement*, New York: St. Martin's Press.

Lave, L. B., 1981,*The Strategy of Social Regulations: Decision Frameworks for Policy*, Washington, D.C.: The Brookings Institution.

Link, Albert N., 1993, ‘Evaluating the Advanced Technology Program: A Preliminary Assessment of Economic Impacts’,*International Journal of Technology Management*, Winter, pp. 726–739.

Link, Albert N., 1987,*Technological Change and Productivity Growth*, London: Harwood Academic Publishers.

Ormala, Erkki, 1989, ‘Nordic Experiences of the Evaluation of Technical Research and Development’,*Research Policy*, December, pp. 333–342.

Smith, V. Kerry, 1986, ‘A Conceptual Overview of the Foundation of Benefit-Cost Analysis’, in J. Bentkover, V. Covello and J. Mumpower (eds.),*Benefit Assessment: The State of the Art*, Dordrecht, Holland: Kluwer Publishers.

Steward, Fred, 1993, ‘Extracting Significant Innovations from Published Sources in Great Britain’, in Kleinknecht A. and D. Bain (eds.),*New Concepts in Innovation Output Measurement*, New York: St. Martin's Press.