CHAPTER 6

MEASURING INNOVATION

KEITH SMITH

6.1 Introduction

It is sometimes suggested that innovation is inherently impossible to quantify and to measure. This chapter argues that while this is true for some aspects of innovation, its overall characteristics do not preclude measurement of key dimensions of processes and outputs. An important development has been the emergence of new indicators of innovation inputs and outputs, including economy-wide measures that have some degree of international comparability. Following sections discuss first some broad issues in the construction and use of science, technology, and innovation (STI) indicators, then turn (briefly) to the strengths and weaknesses of current indicators, particularly R&D and patents. Final sections cover recent initiatives focusing on the conceptualization, collection, and analysis of direct measures of innovation.

New rather than “traditional” indicators are emphasized here because, as Kenneth Arrow remarked many years ago, “too much energy has gone into squeezing the last bit of juice out of old data collected for different purposes relative to the design of new types of data,” a point echoed by Zvi Griliches: “far too little fresh economics data is collected” (Arrow 1984: 51; Griliches 1987: 824). Innovation data producers have responded to this kind of challenge. The most important development has been new survey-based indicators, especially the Community Innovation Survey (CIS),
which has been carried out three times in all EU Member States. The basic format of
CIS has diffused to many other countries (including Canada, Australia, Hungary,
Brazil, Argentina, and China). Has this effort been justified? In answering this
much depends on the quality of analysis these surveys make possible, so the final
section discusses the rapidly growing research and publication efforts deriving
from CIS.

6.2 The Conceptual Background: Measurement Issues

Measurement implies commensurability: that there is at least some level on which
entities are qualitatively similar, so that comparisons can be made in quantitative
terms.

An immediate problem is that innovation is, by definition, novelty. It is the
creation of something qualitatively new, via processes of learning and knowledge
building. It involves changing competences and capabilities, and producing qualita-
tively new performance outcomes. This may lead to new product characteristics
that are intrinsically measurable in some way—new lift/drag aspects of an aircraft
wing, for example, or improved fuel efficiency of an engine. However, such technical
measurement comparisons are only rarely meaningful across products. More
generally, innovation involves multidimensional novelty in aspects of learning
or knowledge organization that are difficult to measure or intrinsically non-
measurable. Key problems in innovation indicators therefore concern the under-
lying conceptualization of the object being measured, the meaning of the
measurement concept, and the general feasibility of different types of measurement.
Problems of commensurability are not necessarily insoluble, but a main point
arising from recent work is the need for care in distinguishing between what can
and what cannot be measured in innovation.

Quite apart from the problem of whether novelty can be measured, a fundamental
definitional issue is what we actually mean by “new” (see Ch. 1 by Fagerberg in this
volume). Does an innovation have to contain a basic new principle that has never
been used in the world before, or does it only need to be new to a firm? Does an
innovation have to incorporate a radically novel idea, or only an incremental
change? In general, what kinds of novelty count as an innovation? These issues of
commensurability and novelty are basic problems for all S&T indicators—R&D in
particular—but have been most explicitly addressed in the development of direct
innovation indicators.
6.3 Theories of Innovation and their Use in Indicator Development

Although statistics are often treated as though their meanings are transparent, they always rest on some kind of (usually implicit) conceptual foundations. The system of national accounts, for example, derives from Keynesian macroeconomic concepts that seek to identify components of aggregate demand. R&D data has a complex background in the scientification of innovation—the notion that acts of research and discovery underpin innovation (Laestadius 2003). These conceptual foundations are rarely considered when indicators are used. Such issues are complicated by the fact that some key S&T indicators are by-products of other processes—legal procedures (as with patents), or academic institutions (as with bibliometrics, which rest on publishing conventions).

What kinds of ideas have formed the conceptual foundations of innovation indicators? An important figure here has been Nathan Rosenberg, whose work quite explicitly affected the OECD's Innovation Manual (OECD 1992, 1997). (This manual is usually called the Oslo Manual because much of the drafting and expert meetings on it occurred there.) First, Rosenberg challenged the notion of research-based discovery as a preliminary phase of innovation. Second, he challenged the idea of separability between innovation and diffusion processes, pointing out that most diffusion processes involve long and cumulative programs of post-commercialization improvements (see Rosenberg 1976 and 1982). Perhaps his best-known contribution, with Steven Kline, has been the so-called chain-link model of innovation, which stresses three basic aspects of innovation (Kline and Rosenberg 1986):

- innovation is not a sequential (linear) process but one involving many interactions and feedbacks in knowledge creation
- innovation is a learning process involving multiple inputs
- innovation does not depend on invention processes (in the sense of discovery of new principles), and such processes (involving formal R&D) tend to be undertaken as problem-solving within an ongoing innovation process rather than an initiating factor

The work of Rosenberg alone, and of Rosenberg and Kline, has at least two important implications for indicator development. The first is that novelty implies not just the creation of completely new products or processes, but relatively small-scale changes in product performance which may—over a long period—have major technological and economic implications. A meaningful innovation indicator should therefore be able to pick up such change. The second is the importance of non-R&D inputs to innovation—design activities, engineering developments and experimentation, training, exploration of markets for new products, etc. So there is a
need for input indicators that reflect this input variety and its diverse distributions across activities.

The CIS effort has in general been informed by ideas from recent innovation research. One in particular should be mentioned, especially because it has had a strong impact on research using the new data. This is the idea that innovation relies on collaboration and interactive learning, involving other enterprises, organizations, and the science and technology infrastructure. Data gatherers have been concerned to explore the networking dimension of innovation, and this has been an important conceptual issue in survey design (see Howells 2000, for an overview of research on this topic).

6.3.1 Existing and New Indicators: What Can Be Measured, and What are the Limitations?

What does it mean to measure qualitatively diverse phenomena? Clearly this is a serious problem for R&D data. Research is a knowledge-creating process for which both activities and outcomes are radically incommensurable—there is no meaningful way to assess the dissimilar actions and events that feed into research, let alone to compare the increments to knowledge that follow from research. This problem cannot be overcome—it can only be circumvented by carefully specifying aspects of the research process that are in some serious sense measurable. The solution adopted by the framers of the *Frascati Manual* (the OECD’s operating statistical manual for R&D data collection) has been to write definitions of research—comprising activities, and then seek data on either expenditure or personnel resources devoted to such activities. The measurement concept for R&D is therefore economic in character, and the datasets that result are collections of economic indicators compatible with industrial datasets, and indeed with the national accounts.²

This approach to measurement has also been taken with innovation surveys. The problem is that innovation is usually conceptualized in terms of ideas, learning, and the creation of knowledge (moreover knowledge creation of a far wider character than research), or in terms of competences and capabilities. As with “research,” innovation is a multidimensional process, with nothing clearly measurable about many aspects of the underlying process. Most modern innovation theory rests on some kind of “resource-based” theory of the firm, in which firms create physical and intangible assets that underpin capabilities (see Lazonick in this volume). Innovative learning can be seen as change in the knowledge bases on which capabilities rest. Neither learning, nor the capabilities which result, seem to be measurable in any direct way. However, just as “research” can be captured via expenditures on certain activities, or by the use of time by certain research personnel, so learning processes can to some extent be captured by activities such as design, training, market
research, tooling up, etc. Expenditure on such activities can in principle be measured (of course the practice may be difficult, since some of these innovation-related activities are not straightforwardly reflected in the accounting procedures of firms). On the output side, the question is whether capability outcomes can be measured by some tangible change in physical or economic magnitudes. Once again there are also potential measurement areas—experience (with pilot or experimental surveys in the 1980s) showed that firms can identify changes in their product mixes, and can estimate sales from new or changed products (Smith 1992). So it is possible to define product change, in terms of construction, use of materials, technical attributes, or performance characteristics, and then to look at the place of (differently) changed products in the sales of the firm. These considerations lead to expenditure measures of inputs to innovation, and sales measures of outputs of innovation. These economic measures of innovation are clearly analogous to the measurement of research. This similarity in approach incidentally suggests that it makes no sense to use R&D data while rejecting the use of more direct innovation data.

6.4 Current Major Indicators

This section outlines the major established indicators that have been used for innovation analysis, and provides a brief guide to further analysis of them. There are three broad areas of indicator use in STI analysis: first, R&D data; second, data on patent applications, grants and citations; and third, bibliometric data (that is data on scientific publication and citation).

In addition to this there are three other important classes of indicators:

- technometric indicators, which explore the technical performance characteristics of products (see e.g. Saviotti 1996 and 2001 for a theoretical view of this, and Grupp 1994 and 1998 for analysis and empirical specifications);
- synthetic indicators developed for scoreboard purposes mainly by consultants (see World Economic Forum 2003);
- databases on specific topics developed as research tools by individuals or groups (such as the large firm database used by Pavitt and Patel, or the MERIT-CATI database on technological collaboration developed by John Hagedoorn, or the DISKO surveys on technological collaboration emanating from the University of Ålborg (see Patel and Pavitt 1997 and 1999, Hagedoorn and Schakenraad 1990, and—for extensive reporting on the use of collaboration data—OECD 2001).
The following discussion concentrates on R&D and patents, since bibliometric analysis relates primarily to the dynamics of science rather than innovation (see Moed et al. 1995, and Kaloudis 1997 for reviews of the state of the art).

6.4.1 Research and Development (R&D) Statistics and Indicators

By far the longest-standing area of data collection is R&D. The key OECD document for the collection of R&D statistics is the Standard Practice for Surveys of Research and Experimental Development, better known as the Frascati Manual. The first edition was the result of an OECD meeting of national experts on R&D statistics in Frascati, Italy, in 1963. The manual has been continuously monitored and modified through the years: the current version of the manual, the Frascati Manual 2002, is the seventh edition (OECD 2002). The Manual defines R&D as comprising both the production of new knowledge and new practical applications of knowledge: R&D is conceived as covering three different kinds of activities: basic research, applied research, and experimental development—these categories are distinguished in terms of their distance from application.

It is often difficult to draw the dividing line between what should be counted as R&D and what should be excluded: “The basic criterion for distinguishing R&D from related activities is the presence in R&D of an appreciable element of novelty and the resolution of scientific and/or technological uncertainty, i.e. when the solution to a problem is not readily apparent to someone familiar with the basic
stock of commonly used knowledge and techniques in the area concerned” (OECD 2002: 33). Education and training in general is not counted as R&D. Market research is excluded. There are also many other activities with a scientific and technological base that are kept distinct from R&D. These include such industrial activities related to innovation as acquisition of products and licenses, product design, trial production, training and tooling up, unless they are a component of research, as well as the acquisition of equipment and machinery related to product or process innovations.

R&D is often classified according to multiple criteria, and data is collected in highly detailed forms. Beyond the distinction between basic research, applied research and development the data is classified into sector of performance: business enterprise, government, higher education, and private non-profit. It also distinguishes between sources of finance, both domestic and international. Then there is classification by socio-economic objectives, and a further classification by fields of research. These detailed classifications are usually ignored both by policy analysts and researchers, who tend to focus on gross expenditure only (at industry or country level), thereby missing most of the really interesting detail in the data. For example, a major issue is that, when looking at R&D by fields of research, ICT (information and communications technologies) turns out to be the largest single category in all countries that classify R&D data in this way. However most of the ICT research is actually performed outside the ICT sector, in the form of systems and software development by users. On the one hand, this raises interesting questions about the cross-industry significance of the ICT sector; but there are also questions about the extent to which such activity should be classified as R&D at all. Concerns have also been expressed about whether the R&D definitions are comprehensible to firms (especially SMEs), and whether or not there is systematic undercounting of small-firm R&D (Kleinknecht, Montfort, and Brouwer 2002).

R&D data is always constrained as an innovation indicator by the fact that it measures an input only (Kleinknecht et al. 2002). However, R&D also has fundamental advantages. These include the long period over which it has been collected, the detailed subclassifications that are available in many countries, and the relatively good harmonization across countries. Unfortunately a great deal of the literature consists essentially of an attempt to match aggregate R&D measures across time and across sectors or countries to some measure of productivity (see Griffith, Redding, and Van Reenen (2000) for a very thorough recent example; Dowrick (2003) is a recent survey of this very large literature). However this research effort is limited in two senses—on the one hand it tends to imply (along with the new growth theory, incidentally) that R&D is the primary source of productivity growth, and on the other it fails to exploit the basic complexity of the data that is actually available. The disaggregation processes that are possible with R&D data continue to offer rich and unexploited opportunities for researchers.