

Policy Evaluation: Assessing Innovation Systems and Programs

As is the case with all public policies, science, technology, and innovation programs and policies need proper evaluation. Evidence-based policy making has become increasingly important in recent years, with new techniques being developed to assess individual programs and policies and new methods being adopted to undertake cross-country comparisons of innovation systems. Evaluation has to take place throughout the policy process. Rigorous evaluation is not always easy to achieve, especially for developing countries, because it is resource and data intensive. The benefits, however, can be considerable and can enable public funds to be put to their best use.

It is very difficult to measure the impacts and benefits of innovation policies. Many policies are related, and innovation systems are complex and evolve continuously. As a result, it is often hard to measure outcomes, but such evidence is crucial to establishing whether government policy successfully tackles market failures and provides a positive net stimulus to innovation. Effective evaluation programs should therefore seek to combine a range of evaluation methods, including national and international, quantitative and qualitative, and micro- and macrolevel.

This chapter reviews a range of techniques used to gain information on innovation by firms and economies as a whole and to evaluate programs and policies. It first reviews several widely used international benchmarking indexes that attempt to quantify innovation or innovation capabilities of economies as a whole. The chapter then turns to look at innovation surveys as a means of

This chapter was prepared by Désirée Van Welsum, with the contribution of Derek Chen for the section on macro-benchmarking methods.

obtaining information on what level and types of innovation firms actually engage in. It then proceeds with program evaluation, with particular emphasis on the evaluation of public research and development (R&D), and examines the use of field experiments. Last, the chapter reviews both national and regional policy evaluation. The importance of such evaluations is twofold: not only do they help maximize the economic and social efficiency of public spending, but they also provide information on which programs and policies work and under what circumstances that information can be used to learn from experience and inform future policy decisions. In particular, the section on field experiments discusses quantitative methods that have been developed for policy evaluations in emerging and less-developed economies. The particular focus of these techniques is to help understand which policy initiatives are especially effective and might be introduced on a larger scale.

Benchmarking Innovation at the Country Level

In this section, we review some of the more commonly used measures of country-level innovation. By determining the type of data or indicators that are being used by the innovation indexes, we can place the indexes into two broad groups: the first group of indexes is based exclusively on hard or objective data, and the second group of indexes uses a combination of both objective and subjective or opinion-based data. The scarcity of objective data tends to lead to different indexes' gravitating toward the same few available "hard data" indicators. It is unfortunate that all of these hard data indicators are relatively more R&D oriented.

The benefit of using data from opinion surveys is that it provides some indication of performance in an area for which hard data do not exist. One very valid concern with using such data, however, is that the ratings across individuals can never be truly consistent. For example, one person's "very good" may be another's "average," and this inconsistency may not be statistically rectified by increasing the survey sample size. Nevertheless, simple comparisons of country rankings in selected objective- and subjective-based data innovation indexes reveal that the rankings are, to a large extent, correlated. This correlation indicates that both types of indexes provide more or less similar pictures of the level of innovation in countries.

The World Bank Knowledge Economy Index

The World Bank produces one of the longest-running country-level innovation indexes available for a large number of countries. Its innovation index is one of several indexes generated by the Knowledge Assessment Methodology (KAM) (see <http://www.worldbank.org/kam>). The other KAM indexes include the Education Index, the Information and Communication Technologies (ICT) Index, the Economic Incentive and Institutional Regime Index, and the

Knowledge Economy Index (KEI), which is a simple average of the four preceding indexes. The KAM was launched in 2000 and has been updated at least once every year since. The most recent version of the KAM, KAM 2008, provides data and indexes for 140 countries. The KAM's innovation index is based on countries' performance over three indicators: (a) the number of patents granted by the U.S. Patent and Trademark Office (USPTO); (b) the number of scientific and technical journal articles; and (c) the amount of royalty payments and receipts. These indicators are relatively more R&D-oriented, but this orientation is due to the difficulty of quantifying innovation, as well as to problems in data availability in this domain. Being based on these three indicators allows the innovation index to be more consistently measured across a larger number of countries. In addition, the inclusion of royalty payments widens the definition of innovation to include technological adoption, thereby creating a broader index.

The UNCTAD Innovation Capability Index

In its World Investment Report 2005, the United Nations Conference on Trade and Development (UNCTAD) introduced the UNCTAD Innovation Capability Index (UNICI), which provides a measure of national innovation capabilities. The UNICI itself, which is based completely on hard data, comprises two equally weighted subindexes: the Technological Activity Index, which measures innovative activity, and the Human Capital Index, which measures skills availability for that innovative activity. The UNICI is available for 117 countries for the years 1995 and 2001. The UNICI Technological Activity Index was constructed using three innovation indicators: researchers in R&D, the number of USPTO patents granted, and the number of scientific and technical journal articles; the UNICI Human Capital Index employs the adult literacy rate and the gross secondary and tertiary enrollment rates. Note that the World Bank's KAM education index is also based on the same three education variables.

The UNDP Technology Achievement Index

The Technology Achievement Index (TAI) was developed by the United Nations Development Programme (UNDP) in 2001 to measure a country's ability to create and diffuse technology and build a human skills base, reflecting national capacity to participate in the technological innovations of the network age. The index focused on outcomes and achievements rather than on effort or inputs and is based on four dimensions of technological capacity: creation of technology, diffusion of recent innovations, diffusion of old innovations, and human skills. Each dimension was measured by two objective data indicators. The TAI, published in the *Human Development Report 2001* (UNDP 2001), covered 72 countries but has since been discontinued.

The Arco Technology Index

The Arco Technology Index (ATI) was developed by Daniele Archibugi and Alberto Coco (see Archibugi and Coco 2004) to measure the technology capacities of developed and developing countries. The ATI was built on the basis of the UNDP TAI, but the advantage of ATI is its larger country coverage and the capacity for comparisons across time. ATI 2004 covers 162 countries for two time periods: 1987–90 and 1997–2000. The ATI is composed of eight hard data indicators organized into three domains: technology creation, technological infrastructures, and level of human skills.

RAND Science and Technology Capacity Index

The Science and Technology Capacity Index (STCI) was produced by RAND in 2001 to measure a country's capacity to absorb and use science and technology knowledge. The STCI was updated by Wagner, Horlings, and Dutta (2004) and covers 76 countries. The STCI is constructed from eight hard data indicators that are divided into three domains:

- *Enabling factors*, measuring the environment conducive to the absorption, production and diffusion of knowledge
- *Resources*, contributing directly to science and technology (S&T) activities
- *Embedded knowledge*, measuring the output of S&T knowledge.

The European Innovation Scoreboard Summary Innovation Index

The Summary Innovation Index (SII) has been published annually with the European Innovation Scoreboard (EIS) by the European Commission since 2000 as a comparative assessment of the innovation performance of the member states of the European Union (EU) and other selected countries. The EIS 2007 includes innovation indicators and trend analyses for 37 countries, covering the EU-27 member states as well as Australia, Canada, Croatia, Iceland, Israel, Japan, Norway, Switzerland, Turkey, and the United States.

The SII is constituted of 25 innovation indicators, which are classified into five dimensions:

- *Innovation drivers*, measuring the structural conditions critical to innovation potential
- *Knowledge creation*, measuring the investments in R&D activities
- *Innovation and entrepreneurship*, measuring the efforts toward innovation at the firm level
- *Applications*, measuring the performance of labor and business activities and their value added in innovative sectors
- *Intellectual property*, measuring the achieved results from innovation activities.

Similar to the Global Innovation Index (see below), the EIS arranges the five dimensions into two groups: inputs and outputs. Innovation inputs cover

the first three dimensions, and innovation outputs cover the last two dimensions. Similar to the KAM, the SII uses only objective data.

The WEF Global Competitive Index

As 1 of the 12 components of its Global Competitiveness Index (GCI), the World Economic Forum (WEF) produces an innovation index for 134 economies. The GCI, which is published annually with the Global Competitiveness Report, is an index that measures national competitiveness.^{1,2} In contrast to the innovation indexes presented above, which are based completely on hard data, the GCI and the associated innovation index are based on both hard data and opinion survey–based data. The WEF innovation pillar is composed of one hard data variable (utility patents) and six opinion survey–based variables (capacity for innovation, quality of scientific research institutions, company spending on R&D, university–industry research collaboration, government procurement of advanced technology products, and the availability of scientists and engineers).

The World Business and INSEAD Global Innovation Index

The Global Innovation Index (GII), currently available for 130 countries, was developed by INSEAD and World Business in 2007 to show the degree to which nations and regions are responding to the challenge of innovation. The GII is made up of 84 variables divided into eight pillars, which are grouped as five input pillars and three output pillars. The five input pillars include institutions and policies, human capacity, infrastructure, technological sophistication, business markets, and capital. These pillars represent factors that enhance innovative capacity. The three output pillars include knowledge, competitiveness, and wealth. They measure results from successful innovation. The GII uses both objective data drawn from various public and private sources such as the World Bank and the International Telecommunication Union and subjective data drawn from the World Economic Forum’s annual Executive Opinion Survey.

Country Rankings across Selected Innovation Indexes

While it is important to note differences in the various indexes in terms of their construction and underlying data, it is more important to get a sense of the similarities and differences in terms of their outcomes. To this end, the country rankings of four innovation indexes that are relatively similar in the time period and number of countries covered are compared.³ Table 7.1 presents the top 10th percentile, the bottom 10th percentile, and the 50th and 60th percentile country rankings for the KAM Knowledge Economy Index, the KAM Innovation Index, WEF’s GCI Innovation Index, and the Global Innovation Index.

Table 7.1 Rankings of Economies for the Knowledge Assessment Methodology, Global Competitiveness, and Global Innovation Indexes, 2008–09

Rank	KAM knowledge economy index 2009	KAM innovation index 2009	GCI innovation index 2008–09	Global innovation index 2008–09
<i>Top 10th percentile</i>				
1	Denmark	Switzerland	United States	United States
2	Sweden	Sweden	Finland	Germany
3	Finland	Finland	Switzerland	Sweden
4	Netherlands	Singapore	Japan	United Kingdom
5	Norway	Denmark	Sweden	Singapore
6	Canada	United States	Israel	Korea, Rep.
7	United Kingdom	Netherlands	Taiwan, China	Switzerland
8	Ireland	Canada	Germany	Denmark
9	United States	Israel	Korea, Rep.	Japan
10	Switzerland	Taiwan, China	Denmark	Netherlands
11	Australia	United Kingdom	Singapore	Canada
12	Germany	Japan	Netherlands	Hong Kong, China
<i>50th and 60th percentile</i>				
50	Ukraine	Armenia	Jordan	Brazil
51	Kuwait	Brazil	Ukraine	Turkey
52	Serbia	Serbia	Italy	Oman
53	Brazil	Trinidad and Tobago	Thailand	Barbados
54	Armenia	Turkey	Lithuania	Greece
55	Trinidad and Tobago	Ukraine	Chile	Jordan
56	Macedonia FYR	Mexico	Vietnam	Poland
57	Argentina	Thailand	Slovak Republic	Azerbaijan
58	Russian Federation	Romania	Senegal	Sri Lanka
59	Turkey	Jordan	Malta	Latvia
60	Jordan	Venezuela, R.B.	Colombia	Mexico
61	Thailand	China	Kazakhstan	Croatia
62	Mauritius	Uruguay	Greece	Philippines
63	South Africa	Panama	Poland	Vietnam
64	Oman	Georgia	Nigeria	Trinidad and Tobago
65	Mexico	Jamaica	Turkey	Mauritius
66	Saudi Arabia	Kuwait	Egypt, Arab Rep.	Panama
67	Georgia	Oman	Jamaica	Russian Federation
68	Panama	Moldova	Romania	Romania
69	Moldova	Guyana	Serbia	Nigeria
70	Kazakhstan	Macedonia, FYR	Kuwait	Kazakhstan
71	Jamaica	Tunisia	Uganda	Jamaica
72	Colombia	Colombia	Panama	Bulgaria
73	Peru	Egypt, Arab Rep.	Guatemala	Colombia
<i>Bottom 10th percentile</i>				
111	Zambia	Madagascar	Mozambique	Burkina Faso
112	Mali	Tanzania	Kyrgyz Republic	Moldova
113	Lesotho	Nicaragua	Bangladesh	Cambodia
114	Benin	Cambodia	Guyana	Paraguay
115	Nigeria	Zambia	Mauritania	Ethiopia

Table 7.1 continued

Rank	KAM knowledge economy index 2009	KAM innovation index 2009	GCI innovation index 2008–09	Global innovation index 2008–09
116	Nepal	Tajikistan	Nepal	Albania
117	Burkina Faso	Guatemala	Nicaragua	Kyrgyz Republic
118	Cameroon	Mali	Bosnia and Herzegovina	Bolivia
119	Mozambique	Burkina Faso	Ecuador	Nepal
120	Cambodia	Mozambique	Albania	Mozambique
121	Bangladesh	Bangladesh	Bolivia	Zimbabwe
122	Ethiopia	Ethiopia	Paraguay	Lesotho

Source: Author compilation.

Notes: KAM = Knowledge Assessment Methodology; GCI = Global Competitiveness Index.

Three observations regarding the rankings are noteworthy. The first striking fact is that many economies have broadly similar rankings across the different indexes. These similarities in rankings imply that the indexes are largely correlated and provide similar rankings, even though they are based on different indicators and even on different types of indicators (objective and subjective). That fact consequently suggests that even though innovation is inherently difficult to measure, relative achievements in innovation can be broadly measured so long as reasonable metrics are used.

The second fact is that the positive correlation across the indexes appears to be weaker the further one moves down the rankings. More specifically, nine economies make it to the top 10th percentile (the top innovators) of three or more indexes. In rankings around the median in the 50th and 60th percentile, 15 economies appear in three or more indexes. In the bottom 10th percentile, six economies appear in three or more indexes. One possible reason for the apparent weakening in correlation may be that the collection and availability of data for less advanced countries are more challenging in terms of accuracy and consistency. Poor-quality data could result in a lack of consistency among the different metrics and could lead to significant differences in country rankings, depending on the metrics used.

Third, there does not appear to be a distinct difference in the rankings between the narrower innovation indexes (KAM Innovation Index and GCI Innovation Index) and the broader ones (KAM KEI and the GII). Not surprisingly, this absence of differences suggests that innovation and a strong broader economy go hand in hand or, more specifically, that the latter is a requirement for the former to take place, a point that is heavily emphasized in other chapters of this volume.

It is also important to keep in mind the appropriate use of these innovation indicators and indexes. Various institutions have proposed these indexes and indicators as an easy way to measure the amount of innovation taking

place in an economy. However, it should be clear that these indicators are far from ideal, in particular because they may be less relevant for economies that have not attained certain levels of economic development. For example, using fiscal incentives to encourage higher patent counts may be an unsound economic development strategy if the new knowledge associated with the patent is not usable by domestic industries because of their lack of technical know-how. In general, generating new knowledge may not be the best use of scarce resources in economies that are just embarking on the industrialization process. Instead, a focus on assimilating and adapting existing technical knowledge from abroad to enhance domestic industries would generally be a more appropriate path. In light of the above, policy makers should exercise caution when using these metrics as policy or target variables.

Microlevel Innovation Surveys

Innovation surveys collect information about innovation inputs and outputs in firms. They go beyond S&T statistics such as R&D surveys or patent data by also collecting information on nontechnological innovation and on factors that support or impede innovation efforts. Innovation surveys thus collect data on various types of innovations, the reasons for innovating (or not), the impacts of innovation, collaboration and linkages among firms or public research organizations, and flows of knowledge.

The benchmark definitions of innovation in most surveys reflect those of the *Oslo Manual*, published by the Organisation for Economic Co-operation and Development (OECD), which provides a harmonized framework for innovation surveys to ensure their comparability and quality. While the first edition covered mainly technological product and process innovations, the manual has evolved over time to take new forms of innovation into account.

The current version of the OECD's *Oslo Manual* (2005a) defines *innovation* as the implementation of a new or significantly improved product (good or service) or process; a new marketing method; or a new organizational method in business practices, workplace organization, or external relations. It distinguishes four types of innovation: product, process, marketing, and organizational. Furthermore, the concept of *new* can mean new to the firm, new to the market, or new to the world. Finally, the manual considers the role of linkages and collaboration in innovation: that is, whether innovations are developed mainly by the firm itself, together with others, or mainly by others.

In Europe, the first Community Innovation Survey (CIS) was carried out in 1993, and the sixth in 2008. The survey is now conducted in all EU countries. Similar innovation surveys based on the *Oslo Manual* are conducted in many other countries, including Australia, Canada, the Republic of Korea, New Zealand, Norway, the Russian Federation, South Africa, Switzerland, and Turkey, as well as most Latin American countries. However, the United States

does not so far have an official innovation survey based on the *Oslo Manual* framework.⁴

An annex was added to the current edition of the *Oslo Manual* to cover innovation surveys in developing countries. It takes into account the fact that most innovation in developing countries involves dissemination mechanisms and incremental change. Aspects discussed include differences in the structure and functioning of firms and markets, market failures and barriers to innovation, weaknesses in macroeconomic and institutional framework conditions, and the relative weakness of statistical systems.

Innovation Surveys in Developed Countries

Many of the most widely used international statistics on innovation activities do not provide direct evidence on the extent to which innovations are actually introduced. R&D-related indicators are an input measure into the overall innovation process, and patents and scientific publications are intermediate outputs. The formal, internationally agreed definition of innovation, that of the *Oslo Manual*, is output based and reflects the successful commercial introduction of new products and processes.

Direct output indicators of firm-level innovation performance are typically obtained through innovation surveys. The CIS, for example, provides indicators of direct relevance to European policy makers (see box 7.1 for examples) and increases the evidence available for evaluating the need for, and the performance of, innovation policies. The latest available survey, CIS4, covers EU member states and candidate countries as well as Norway and Iceland.

Box 7.1 Examples of Indicators from Innovation Surveys

The Community Innovation Survey provides a very rich data set. Examples of indicators that can be constructed include the share of firms that are

- involved in innovation;
- involved in process, product, marketing, or organizational innovation;
- introducing a good or service new to the firm, new to the market, new to the world;
- performing R&D;
- applying for a patent;
- receiving public funding;
- engaging in innovation cooperation (distinguishing among cooperation with universities, higher education or government research institutes, or foreign partners);
- reporting on important impacts of innovation (improved products, increased range of products, entering new markets);
- reporting impacts of organizational innovation (improved products, reduced response time, reduced costs, improved employee satisfaction).

Additional data cover expenditure indicators such as innovation expenditure and the share of turnover from different types of innovation.

Source: Author compilation.

The survey uses a harmonized questionnaire and survey method based on the *Oslo Manual* and agreed with Eurostat, the European Statistical Office. To maintain the secrecy of enterprise-level information, the microlevel database is confidential and can be accessed only by Eurostat staff. The most recent survey, launched in 2008, is gathering data for the reference years 2006–08 and will collect additional information about “nontechnical” aspects of innovation, such as management techniques, organizational change, design, and marketing issues.

The CIS4 found that around 40 percent of all EU firms undertook some kind of process or product innovation (European Commission 2007). Innovative companies were larger on average than noninnovative companies, accounting for around two-thirds of total employment. Three-quarters of the innovative companies invested in advanced machinery and equipment, and around half in R&D and training. R&D expenditure accounted for only a little over half their total expenditure on innovation. CIS4 data also indicate that around two-fifths of European firms introduced organizational or marketing innovations, a proportion similar to the introduction of other forms of innovation. However, many firms that introduced nontechnological innovations did not introduce product or process innovations.

Business services companies, particularly in computing and related activities and in financial intermediation, comprised a large share of innovative companies. The share of innovative firms in manufacturing differed little from that of the economy as a whole. When companies were questioned about the factors hampering innovative activities, they indicated that the four most important were the costs of innovation, uncertainty about demand, a lack of qualified personnel, and a lack of potential partners, in that order.

In addition to the microlevel indicators from the CIS surveys, related indicators on framework conditions and factors that enable innovation and its diffusion are also very important. Box 7.2 provides some examples.

Innovation Surveys for Developing Countries

The annex of the latest *Oslo Manual* (OECD 2005a) offers guidelines for the implementation of innovation surveys in developing countries. Many developing countries conducted innovation surveys on the basis of the preceding edition. These tended to require adaptations of the proposed methodologies to capture the particular characteristics of innovation processes in countries with economic and social structures different from those of the more developed OECD countries. The first effort to compile these particulars and guide the design of cross-country comparable innovation surveys was made by RICYT (Ibero-American Network on Science and Technology Indicators, or Red Iberoamericana de Indicadores de Ciencia y Tecnología) and resulted in the publication of the *Bogotá Manual*, which was later used in most innovation surveys conducted in Latin America and other regions. The importance

Box 7.2 Additional Sources of Innovation-Related Indicators

Innovation-related indicators include human resources, skills, and knowledge, as these are crucial enablers of innovation, absorptive capacity, and spillovers. They also include “knowledge infrastructure” indicators (such as public sector funding of research, universities, science and technology personnel, number of researchers, and the like). Others include business enterprise expenditure on research and development (R&D) and gross domestic expenditure on R&D as a percentage of gross domestic product (GDP), gross domestic expenditure on R&D financed from abroad, broadband penetration, education and training, participation in lifelong learning, employment in high-technology industries, employment in occupations requiring human resources in science and technology, scientific articles, science and engineering degrees, patents, patents with foreign co-inventors, trademarks, and venture capital, among other indicators. International organizations provide rich data sets with this type of information, in particular the OECD’s *Science, Technology and Industry Scoreboard* (OECD 2007d), the OECD’s *Science, Technology and Industry Outlook* (OECD 2008d) and the underlying databases, and the European Commission’s *European Innovation Scoreboard* (various editions available at <http://www.proinno-europe.eu/>).

The Information Technology and Innovation Foundation (ITIF 2009) provides another source of cross-country innovation-related indicators. It uses 16 indicators to benchmark countries’ competitiveness, including their innovative capacity, based on various innovation-related indicators (1–4) as well as on more general indicators of the macroeconomic framework conditions for competitiveness (5–6):

1. *Human capital*. Higher educational attainment in the population aged 25–34 and the number of science and technology researchers per 1,000 employed
2. *Innovation capacity*. Corporate investment in R&D, government investment in R&D, and share of the world’s scientific and technical publications
3. *Entrepreneurship*. Venture capital investment and new firms
4. *Information technology infrastructure*. E-government, broadband telecommunications, and corporate investment in IT
5. *Economic policy*. Effective marginal corporate tax rates and ease of doing business
6. *Economic performance*. Trade balance, foreign direct investment inflows, real GDP per working-age adult, and productivity.

Intangible assets surveys (such as the Intangible Asset Monitor, the Skandia navigator, and the International Accounting Standards Board) also provide useful complementary data, including indicators of human capital, intellectual capital, organizational capital, and relational capital.

Additional firm-level indicators can also be found in private sources, such as the reports and data sets from business consultants such as the Boston Consulting Group, McKinsey & Company, and Booz Allen Hamilton, which also report indicators companies use to measure innovation (outputs). These include revenue growth from new products, sales from new products, profit increase from new products, customer satisfaction with new products, and the return on investment for new products.

Source: Author compilation.

and impact of this standard-setting work were the inspiration for the *Oslo Manual* annex, the preparation of which was coordinated by the UNESCO Institute for Statistics (UIS). Many of the annex's recommendations are based on the experience of countries that have already conducted innovation surveys, most of which are among the medium- and higher-income countries of the developing world where innovation has already become a policy issue. The knowledge gained by these countries should help other developing countries acquire their own experience without having to build exclusively on innovation measurement exercises carried out in developed countries.

The Measurement of Innovation in Developing Countries

The *Oslo Manual* stresses the need for the measurement of innovation in developing countries to be comparable to results obtained in developed countries to enable benchmarking and construction of a coherent international system of innovation indicators. At the same time, the innovation surveys need to take account of the characteristics of innovation in developing countries. In addition to the difficulties of applying existing definitions, challenges can arise in measuring incremental changes that may not result in "new or significantly improved" products or processes or in defining the scope of innovations, since concepts such as "new to the market" may be interpreted differently in environments with less developed infrastructures.

A main reason for conducting innovation surveys in developing countries is to inform public policy making and the design of business strategies. The main focus is on the generation, diffusion, appropriation, and use of new knowledge in businesses. Measurement exercises should focus on the innovation process, rather than on its outputs, and emphasize how capabilities, efforts, and results are dealt with. The efforts of firms and organizations (innovation activities) and their capabilities are equally or even more important than the results (innovations). Factors hindering or facilitating innovation are considered key indicators. "Potentially innovative firms" may be of particular interest in developing countries. They are a subset of "innovation-active firms," that is, those "that have had innovation activities during the period under review, including those with ongoing and abandoned activities." A key element of innovation policies in developing countries is to help potentially innovative firms overcome obstacles that prevent them from being innovative and to convert their efforts into innovations.

Different measurement priorities in developing countries (why, what, and how to measure) guide the design of innovation surveys. Developing countries use surveys to obtain information on the innovation strategies present in the innovation system and to understand how they contribute to strengthening the competitiveness of particular enterprises and to enhancing economic and social development more generally. This effort requires linking the analysis of micro-, meso-, and macroeconomic innovation data to issues such as the

technological content of exports, strengths and weaknesses of particular industries or innovation systems, the absorptive capacity of innovation systems, networks, links between education and employment, and indicators of the effectiveness of different public instruments for supporting and promoting innovation.

The *Oslo Manual* finds the concept of innovation capabilities useful for classifying firms and industrial sectors in developing countries. It argues that the most significant innovation capability is the knowledge accumulated by the firm, which is embedded mainly in human resources but also in procedures, routines, and other characteristics of the firm. Because such intangible assets are notoriously difficult to measure, particular attention should be given to the parts of the surveys that directly connect knowledge development and diffusion with innovation capabilities, such as human resources, linkages, and the diffusion and use of ICT. In addition, more complex issues, such as the types of decision-making support systems put in place by the firm's direction and management and the firm's actual potential for knowledge absorption, should also be examined. For an accurate measurement of firms' innovation efforts, it is also essential to collect information about the intensity of innovation activities. Therefore, details should be obtained about the innovation activities undertaken by the firm and, where possible, data on expenditure by innovation activity.

Organizational change is extremely important in developing countries, where the absorption of new technologies, mostly incorporated in machinery and other equipment, can require significant organizational change. Questions on the implementation of organizational innovations should therefore be supplemented with questions on human resources and training and on the incorporation of ICT. This process can help provide an indication of an enterprise's innovative capabilities.

Adapting Existing Innovation Surveys

Innovation surveys should be adapted for developing countries in three main areas: ICT, linkages, and innovation activities. Surveys should specifically address ICT. If specific surveys on ICT in businesses are not available, however, innovation surveys should inquire about available infrastructure, the purpose and use of ICT (separating front- and back-office activities), the existence of internal ICT management and development capabilities, and ICT expenditure and its relation to organizational innovation.

For a firm's external linkages, a proxy measure of complexity can be developed by crossing the "type" and "objective" of the linkages. This process establishes a matrix of *linkage agents* (universities, technical and vocational training institutions, technological centers, test labs, suppliers, clients, head office, enterprises belonging to the same group, other firms, consultants, R&D firms, public S&T agencies) and *types of linkages* (open information sources, acquisition

of knowledge and technology, and innovation cooperation, supplemented by complementary activities, particularly access to new sources of financing and to commercial information).

For measurement of innovation activities, “hardware purchase” and “software purchase” should be covered separately as should “industrial design” and “engineering activities,” “lease or rental of machinery, equipment and other capital goods,” “in-house software system development,” and “reverse engineering.” Data should also be collected on the composition (by qualification, type of occupation, and gender) and management of human resources. For the latter, information on actions taken by firms with regard to training, including the resources involved, is important. Data can be collected to obtain information on the innovative capabilities of enterprises, not only on training activities linked to innovation but also on general training in areas such as management and administrative training, ICT, industrial security, and quality control.

Methodological Issues

The design and planning of innovation surveys in developing countries need to take account of the relative weakness of statistical systems. Because linkages between surveys and data sets tend to be weak or nonexistent, it is difficult to use information from other surveys in the design of the exercise and in the analysis of its results. The weakness, or sometimes lack, of official business registers, which are normally used as sample sets, is another example of this type of problem. It is important to involve national statistics offices in innovation surveys. If the statistical system lacks appropriate data about firm performance, some basic variables (for example, questions on sales or on turnover) can be included in the innovation survey to enable analysis of the relation between actions taken by firms for innovation and market performance (competitiveness).

Personal interviews (instead of mail or phone surveys) by adequately trained staff (for instance, undergraduate or graduate students) are recommended, since they have a positive impact on the response rate and on the quality of the results obtained. Interviews conducted by qualified staff can also provide the respondent with help in completing the questionnaire and increase response rates, particularly in countries where postal services may not be reliable.

The questionnaire can be designed with separate sections so that different persons in the firm can reply to different sections. Guidance for the respondent should be included with the questionnaire. It may be necessary to clarify certain concepts and provide a definition of terms used, and the wording needs to be adapted to the knowledge and experience of an “average” respondent. In certain cases, questionnaires may need to be formulated in more than one language.

It is generally recommended that innovation surveys be conducted every two years, but in developing countries every three or four years may be more appropriate. If possible, they should be timed to coincide with the major international innovation surveys, such as Europe's CIS, to obtain comparable data for similar time periods. It is also good to update a minimum set of variables every year (such as the main quantitative ones) if resources permit. A less costly strategy is to attach a significantly reduced questionnaire to an existing business survey. In some cases, simplified questionnaires can be designed to cover small firms to encourage their participation in innovation surveys.

The results of the innovation surveys should be published and distributed widely to encourage businesses to participate in future rounds and to increase awareness and use by researchers and policy makers. Diffusion mechanisms need to be included in the budget early in the exercise. Finally, an adequate legislative basis for collecting innovation statistics can help ensure the success of such exercises.

Program Evaluation

Evaluation of government action for innovation can be focused primarily on specific programs. After discussing methodological trends and issues, we will detail countries' experience in evaluating tax incentives for business R&D and public R&D support in various forms, including the support of R&D institutes.

Trends and Issues

Evidence-based policy making and the effective evaluation of public policy have become increasingly important in recent years, especially for science, technology, and innovation. These areas are becoming widely recognized as key drivers of economic growth and competitiveness. They also help reach socioeconomic objectives. It is important to evaluate not only whether the policy was implemented as planned but also whether it had the expected impact.

Program and policy evaluation is important, because it is necessary to make optimal use of public funds by maximizing the desired outcomes and ensuring that scarce resources are efficiently allocated. It is also important to gather information on which programs and policies work, and under what circumstances, to learn not only from success but also from failure. Insights into the determinants and magnitude of successful program and policy outcomes can be used to inform future policy decisions.

However, it is difficult to measure the impacts and benefits of government policy. Because many programs, policies, and policy areas are related, and because innovation systems are complex and evolve continuously, outcomes are not easy to measure because they may differ in the short, medium, and long term and because it is a challenge to establish causality. Effective evaluation programs therefore seek to combine a range of evaluations: national and

international, quantitative and qualitative, micro- and macrolevel (box 7.3). Moreover, as innovation programs and policies vary greatly across countries and agencies in their content and objectives, evaluation methods and approaches often need to be adapted to individual circumstances.

In a program evaluation, it is important to bear in mind the immediate problem or market failure it aims to address, as well as the wider context, such as the program's contribution to overall innovation goals and the

Box 7.3 Effective Policy and Program Evaluation Challenges for Designing Evaluation Schemes

The difficulty of designing an effective evaluation can be illustrated by looking at the evaluation of public programs that support research, as these are an important aspect of innovation policies. The difficulties include selection bias in the attribution of funding, the attribution of research results that may draw on previous work to varying degrees, potential knowledge spillovers and absorptive capacity, the "additionality" of public funding, and potential "crowding out" effects.

A further challenge is to identify and measure not only the economic but also the wider socioeconomic benefits and to determine the time scale and time lag of the evaluation as the outcomes identified may depend on whether short-, medium- or long-term effects are examined. Isolating the effects of particular policies can also be difficult, given the increasing interaction between different policies and programs across a wide range of areas not limited to innovation policy.

Deciding on the evaluation criteria and benchmarks and who should carry out the evaluation is also not straightforward. For example, for a peer group evaluation, innovation users (business, government) and international representation are ideal, but those who also competed for the funding should not be included.

In evaluations of larger programs or projects, it becomes difficult to identify impacts, especially when the beneficiaries of the research are not those that perform it. In addition, one or a small number of successful projects can skew the distribution of the impact results of a portfolio of projects. In evaluating programs, the so-called project fallacy is another common problem, as the sponsored organization, which carries out the publicly funded research as part of a larger program of work, may be inclined to overattribute effects or deliverables to the funded part of the research to please the sponsor (OECD 2006).

Quantitative versus Qualitative Assessment

Quantitative assessment does not tell the whole story, and it is difficult to disentangle relations and correlations. Furthermore, no fully specified dynamic general equilibrium models of innovation exist. Econometric models tend to assume a linear process that does not take into account the complexity of innovation. Qualitative assessment is important but is difficult, as innovation performance depends on the characteristics of the country's economy, innovation systems, and institutions and because cultural factors also play an important role (corporate culture, entrepreneurship culture, "general" culture). The two approaches are complementary, and an effective policy evaluation should combine elements of both.

Box 7.3 continued**National versus International Assessment**

National assessments can combine a detailed analysis of microdata with survey results. Overall assessment of national innovation policies and systems (such as the evaluations carried out in the OECD's national innovation policy reviews) can complement more detailed studies.

International assessment may include international benchmarking; for example, the European Commission's Innovation Scoreboard and the World Bank's Knowledge Economy Index, Knowledge Index, and scorecards as well as the ITIF innovation and competitiveness benchmarking study (ITIF 2009). Cross-country empirical studies, such as the OECD's Going for Growth studies and Innovation Microdata Project, can complement the international benchmarking, as can qualitative surveys, such as that of the World Economic Forum (WEF 2008).

The efforts of international organizations greatly facilitate and enhance the scope for international policy learning. The OECD and UNESCO have a long-standing commitment, dating from the 1960s—alongside the European Union, UNCTAD, and the World Bank—to provide a platform for international learning. The recent success of emerging economies is likely to have been aided by international policy learning.

Source: Author compilation.

possibility of unintended effects, interactions, or trade-offs with other programs. Furthermore, in impact assessment, it is important to understand why and how these occurred in order to gauge whether the program would lead to similar outcomes elsewhere. However, in the absence of strong institutional support and encouragement, underinvestment in program evaluation is likely, especially when the desired outcomes of policy interventions are not narrowly or precisely defined and when the impacts concern broader sectoral or economy-wide outcomes or take relatively longer to materialize (Ravallion 2009). Nonetheless, support for rigorous evaluation is increasing, as illustrated by the evaluation of government support mechanisms, publicly funded research, and use of field experiments.

Fiscal Incentives for R&D

Many governments continue to offer fiscal support for private sector R&D through grants or R&D tax incentives. These are longstanding and widely used policy instruments for stimulating innovation. Evaluating the outcomes of government-supported projects is difficult because it may be necessary to take account of their wider social benefits and to know what the situation would have been in the absence of public support. The latter is a particularly important hurdle. Evaluations of government support are also complicated by the fact that it may take time, even many years, for the benefits to appear, but judgments regarding the use of the public money cannot always wait that

long. Finally, wider socioeconomic benefits and other potential spillovers that are hard to identify and measure are possible. Thus, many challenges must be overcome to achieve effective ex post evaluations of public programs that support research.

Selection bias is another issue. One way to overcome it might be to compare firms that have received funding with similar firms that have not. However, if the decision on funding evaluates the quality of proposals correctly, these would also have been the most likely to succeed in the absence of funding. Therefore, this approach does not necessarily get around the selection bias problem. Identifying factors that determine the probability of selection but not the probability of a successful outcome would make ex post evaluations less biased.

Another difficulty when evaluating the impact of public support for R&D is to identify and take account of potential knowledge spillovers. These may include both economic and socioeconomic benefits, especially when the recipient of public support produces innovations that are used by economic actors not included in the support program. In the absence of data with which to test this hypothesis, the impact of public support programs may be underestimated. In developing countries, these spillover mechanisms may be relatively less important.

A central issue in the evaluation of public support is the “additionality” of public funding, that is, the extent to which public support leads to a higher overall level of R&D expenditure than would otherwise have occurred. A so-called crowding-in effect may also occur if public support enables firms to carry out projects they would otherwise have been unable to finance. At the same time, a “crowding-out” effect may occur if firms that receive public support reduce the amount of funding they would have invested themselves; in which case, public support does not bring about additional R&D.

Tax incentives for R&D may be less likely to result in increased crowding-out effects than direct subsidies, because they operate by reducing the marginal cost of R&D rather than as a potential substitute for funding raised elsewhere, for example, on capital markets. Nonetheless, their impact on real resources may be relatively small, at least in the short term, as they may also help increase the prices of inputs in fixed supply, such as the wages of skilled researchers. They may also distort private sector project decisions if the design of the tax credit gives firms an incentive to undertake projects with a particular payback period.

Empirical Studies on the Impact of Fiscal Incentives. The literature on the impact of government subsidies, tax incentives, and public research programs has been reviewed in many studies. These suggest little consensus on the effectiveness of such instruments. All studies, however, emphasize the sensitivity of the conclusions to the control variables included in empirical assessments and the

level of aggregation of the data set used. For example, evidence of crowding-out effects is more common in firm-level studies than in studies at higher levels of aggregation. It may also be the case that complementarities between publicly financed and privately financed R&D are due to the effect of the former on the input prices of the resources used by the latter. While Guellec and van Pottelsberghe (2000) find evidence of a positive net overall effect from public funding on the growth of privately financed R&D, some forms of funding are found to have a positive effect while others have a negative effect. Finally, some studies have found that the effects of government funding vary with firm size, although results again differ.

Jaumotte and Pain (2005a) have studied cross-country differences in business sector R&D and patenting and shown the importance of initial conditions on the effects of subsidies on innovative activities. They find a small positive effect on R&D from higher direct subsidies, especially when the share of corporate profit is small. In this case, the availability of public funding can help alleviate potential financial constraints. However, at other times, higher subsidies reduce innovative activity. The authors also find that more generous tax relief for R&D has a positive impact on the amounts of both R&D and patenting, with the impact often greater than that of additional direct funding. However, these results are sensitive to the exact specification of the regressions. Using sectoral data from the CIS, they find a significant positive correlation between public funding and the shares of innovator firms and of turnover accounted for by new products.

Jaumotte and Pain (2005b) suggest that tax incentives may be effective, at least in some circumstances, but they fail to show that the social gains from such programs outweigh the associated compliance and administrative costs, although the wider spillover effects of higher R&D on productivity growth raise the likelihood that they do. They also note a higher probability of research duplication when the support takes the form of tax relief rather than grants. Furthermore, new and small firms may be at a relative disadvantage if tax incentives are the only type of support, since they may have relatively little taxable income.

Even if tax relief for R&D is effective, other issues still need to be considered. As for direct grants and subsidies, a complete evaluation would also need to take into account the budgetary costs for the public sector. These need to be balanced by offsetting changes in other fiscal instruments (for any given overall budget balance target), which will also have economic effects. Even if fiscal instruments are effective, the wider question is whether the gains from supporting innovation are greater than the potential gains from supporting other activities or the (deadweight) costs of raising the necessary revenues (Jaumotte and Pain 2005a).

Finally, little is known about whether fiscal incentives for R&D have additional effects arising from their impact on the international location decisions

of research-intensive multinational firms. If tax relief affects location decisions, countries that do not offer it may be at a disadvantage (Poot and others 2003). The extent to which the benefits from cross-border knowledge spillovers require local research capabilities also matters; these considerations may imply a stronger argument for tax relief in smaller countries.

Public R&D

The increased emphasis on evidence-based policy making means a greater need to understand and measure the impact of public sector R&D, notably to ascertain whether public spending on R&D is efficient and whether it contributes to the achievement of social and economic objectives. Public R&D is also increasingly used to address global challenges, such as climate change and the environment. However, it remains difficult to determine and measure the various benefits of R&D investment for society. Furthermore, because the benefits of public R&D can take some time to materialize, especially for basic rather than applied, research, it is difficult to determine the appropriate time for measuring the impact of public R&D and for identifying and quantifying its socioeconomic benefits.

The impacts of public R&D investment have been assessed using econometric analysis and case studies. However, the techniques used and the underlying assumptions determine, in part, the results. Particular challenges include establishing causality, capturing spillovers (international, sectoral, interdisciplinary), the unknown and varying time lags between the investment and the outcome, identifying the main actors and appropriate indicators of outcomes, and the evaluation of results. To some extent, these difficulties reflect the public good nature of public R&D investment and public knowledge more broadly: that is, the fact that it is not depleted when shared and it is difficult to exclude others from its use.

To date, few microeconomic studies address the impacts of public R&D on private sector productivity, and their results are not very conclusive. However, studies of the impact on private sector R&D have demonstrated strong returns to private investment and strong spillover effects that generate substantial economic benefits. Jaumotte and Pain (2005a) find evidence suggesting that research in the nonbusiness sector is an important component of innovation, both directly, as reflected in patenting, and indirectly, through its wider effects on private sector research. Even though an expansion in public sector research can help push up the wage costs of business sector researchers, this effect is more than offset by the positive impact on their efficiency.

The extent of collaboration between business and public research organizations, as proxied by the share of nonbusiness R&D expenditure financed by industry, has increased over time in almost all OECD countries. The work of Jaumotte and Pain also suggests that higher funding shares by the business sector provide an additional stimulus to private sector innovation, in addition

to the direct effects from higher R&D spending in the nonbusiness sector. Data from the CIS also show that collaboration between the public and the private sector increases the share of turnover from new products. These aggregate findings need to be complemented by more detailed analyses of specific programs and different forms of research collaboration to gain a closer understanding of some of the mechanisms at work.

Recently developed indicators provide a means of assessing not only the economic but also the social impacts of public investment in R&D. They link government budget appropriations or outlays for R&D, which classify public budget figures according to socioeconomic objectives, to other data sources. They can help show the contribution of public money to achieving national socioeconomic objectives (OECD 2008d). The next step in assessing the impact of public R&D will be to link data on public R&D budgets by socioeconomic objectives to other data sources, such as scientific publications and patents. Definitions and practices will need to be better harmonized before the contribution of public R&D to socioeconomic objectives can be more fully understood.

Public support for R&D can be channeled in a variety of ways. OECD (2006) distinguishes four levels of evaluation of publicly funded research: (a) institutes and groups, including research departments, teams, laboratories; (b) institutions and operators, including public research organizations and research councils; (c) programs and procedures; and (d) research and innovation systems. Box 7.4 identifies a set of emerging cross-cutting issues.

Research Institutes and Groups. Evaluation tends to be carried out in accordance with one of two common models: the one-off model (such as the Max Planck *Gesellschaft's* approach to creating new groups and using committees of peers) and the periodic recurrent model (such as France's INSERM, which periodically reviews bottom-up proposals from research groups). However, there is an increasing shift toward the latter (OECD 2006), with evaluation evolving in two directions: embedding evaluation within overall strategic exercises (for example, the bottom-up strategic plans of the Spanish national research agency CSIC are reviewed by thematic panels) and taking a more transversal approach to the allocation of core grants (like the German Helmholtz Association's competitive process based on program-oriented funding with interdisciplinary programs evaluated by review panels).

National or subnational governments' evaluations of university research and research institutes have also changed to improve the allocation of core grants at the national level (for example, the United Kingdom's Research Assessment Exercise, a disciplinary peer review exercise, which has inspired similar models in Hong Kong, China, and in New Zealand) and to search for critical mass and excellence, with public funding increasingly concentrated in a few institutes or centers (such as the U.S. National Science Foundation

Box 7.4 Emerging Cross-Cutting Issues in the Evaluation of Publicly Funded Research

Changes in evaluation practices have been driven by four major trends: tighter public governance, competition for research funding, increased focus on the interfaces between fields of research and the economy and society, and increased political acceptance and integration of evaluation outcomes owing to better evaluation methods and tools. The analysis of the impact of these changes on evaluation of publicly funded research points to five issues:

- Clarification of the differences and interactions among indicators, benchmarking, and evaluation is needed.
- The increasing tendency toward internationalized peer review may lead to a rapid normalization of criteria for funding and evaluating research, at the risk of losing specific aspects of local settings.
- The object of evaluation needs to be situated in its proper context (scope, timing) to avoid project fallacy problems (see box 7.3).
- The impact of evaluation depends on the context in which it is implemented and on whether it is a one-time event or an institutionalized part of a regular policy process. In addition, context matters, in particular whether the interests of key stakeholders are aligned with evaluation goals. Operating below the political level can be useful for getting results accepted, timing and matching the decision cycle are important, and the evaluation has to be relevant and robust to be credible.
- The success of an evaluation can be measured by its effects, including the intended and unintended consequences of the public intervention as well as that of the evaluation.

Source: OECD 2006.

Engineering Research Centers, the U.K. Research Councils, Canada's Networks of Centers of Excellence, Australia's Cooperative Research Centers, Sweden's competence centers, and the Dutch Top Technology Institutes).

Research Councils and Public Research Organizations. Research councils and public research organizations can be differentiated according to their functions in the research system and the type of research they carry out. National research councils mainly *fund* research, while public research organizations *perform* research. Hybrids both fund and carry out research. Some focus on basic research, while others are industry oriented. These institutions' assessments of the impact of public R&D have tended to be relatively successful at quantifying impacts. The methodologies used in impact assessments include surveys, input-output analysis, a combination of top-down (contribution of funding to productivity growth) and bottom-up (return to funding through main benefits' transmission channels) approaches, and simulations on computable general equilibrium models.

Research Programs. Research programs are one of the main instruments used by developed countries to implement research and innovation policies. They may fund basic or more applied research in general or in a specific sectoral context, with or without a commercial objective. Two of the most important research programs in terms of resources are the European Union's Framework Programme and the U.S. Advanced Technology Program (ATP). The nature and scope of the research carried out under these two programs are very different.

The EU Research and Technological Development (RTD) Framework Programme (FP) is the main multi-annual R&D funding program in Europe. The FP7—that is, the seventh four-year framework program—is more ambitious than the previous programs as it brings all research-related EU initiatives together under a common umbrella. It has a budget of €53.2 billion and runs from 2007 to 2013. It funds both basic and applied research and aims at enhancing the research capacities and results of all stakeholders: private companies, individual researchers, universities, public research institutions, and foreign actors. Impact assessment tends to rely on econometric modeling. The FP7 also uses an *ex ante* or prospective calculation of the impacts of expenditure generated by a general equilibrium model using impact scenarios drawn up by the European Commission.

In the United States, the ATP, which started in 1990, aims to accelerate the development of innovative technologies for broad national benefit through partnerships with the private sector. It provides cost-shared funding to industry to speed up the development and dissemination of challenging, high-risk emerging technologies that can yield promising commercial possibilities and widespread benefits. It was designed specifically to help U.S. firms translate inventions in universities or national and corporate laboratories into new products, processes, and services able to compete in rapidly changing world markets. Between 1990 and September 2004, the ATP held 44 funding competitions and provided US\$2.2 billion in grants, complementing the US\$2.1 billion provided by industry. Impact assessment of these awards is carried out by the Economic Assessment Office (EAO), which tracks the progress of funded projects for several years after ATP funding ends and identifies the benefits, both direct and indirect, that ATP award recipients deliver. Direct benefits are achieved when technology development and commercialization are accelerated, leading to private returns and market spillovers. Indirect benefits are considered to include publications, conference presentations, patents, and other means of disseminating knowledge.

The EAO uses a variety of methods to measure the investments of the ATP, including surveys, detailed case studies, cost-benefit analysis, statistical analysis, comparison of *ex post* benefits with *ex ante* expected benefits, the tracking of knowledge created and disseminated through patents, and informed judgments. Because the evaluation of emerging technologies is a

relatively new field, existing tools often have to be modified or new ones developed. The ATP also relies on the Business Reporting System (BRS), a data collection tool for tracking the progress of its portfolio of projects and individual participants, from the ex ante baseline to the end of the project and beyond. Progress is assessed against business plans, projected economic goals, and other ATP criteria.

Research Systems. Two recent trends in the evaluation of research systems are the application of detailed evaluation tools to the research or innovation system as a whole to answer particular policy questions, and country reviews of national innovation systems and policies. Examples are the review of the Finnish innovation support system by Georghiou and others (2003), the evaluation of Japan's First and Second Basic Plans, the indicators developed for the U.K. government's 10-year investment framework for science and technology, and the U.S. Government Performance and Results Act and Program Assessment Rating Tool (OECD 2006).

Field Experiments

Field experiments and pilot studies of policies and programs are undertaken to evaluate how they actually work and how their effects might differ from expectations. This assessment is essential for ensuring that only the most effective microprograms are scaled up to national or international levels.

In econometric studies of policy interventions, a particular problem is to know what would have happened to the "treated" group (that is, the group subject to the intervention in question) in the absence of the intervention. A credible impact evaluation has to address this issue. The work of Duflo and her collaborators (Duflo 2004, 2006; Banerjee and Duflo 2008) has recently popularized randomized evaluations as a possible way to address this problem.

Randomized evaluations are intended to help overcome various types of selection bias when measuring the impact of a program or policy intervention by randomly allocating individuals to a "treatment" group of individuals who benefit from the program and a "comparison" group of individuals who do not. For the method to be effective, the random selection of both groups is essential. The outcomes are then compared across the treated and the comparison groups. This approach can be used to test not only the overall effectiveness of a particular program but also the effectiveness of different parts of the program, as some parts may be especially effective while others are not. Duflo (2004) argues that randomized evaluations can be used in many different contexts, provided that the programs have clearly defined objectives and are targeted to individuals or local communities. Programs that affect all individuals or communities as a whole are not suitable, because it is not possible to define a random group that is not subject to the program.

Field experiments and randomized evaluations by Duflo and others have already provided some important insights. Experimental evidence has confirmed that individuals respond to incentives and will try to pervert them if they can do so at little cost. Experiments have also highlighted features that are important in the design of incentive schemes, as opposed to their effect. For example, people are more responsive to an immediate reward, even if it is small, than to a longer-term reward. More research is needed to examine the role and impact of delayed rewards.

The extent to which people learn from each other is another important issue and one with clear implications for innovation and technology diffusion. In developing countries, the impact of learning on technology adoption has been examined in the context of agriculture: in this case, an experiment identifies how social learning affects the development of a technology within a group of farmers and follows its subsequent adoption by them and the members of their network (this treatment group is selected because it faces common unobserved shocks). Such experiments can be designed to examine specific questions or mechanisms and the conditions in which they might work.

Field experiments can be used to test the predictions of theories, and randomized program evaluations can be used to test the effectiveness of more complex policies, including the combination of a variety of policy levers that have not necessarily been tested or even implemented. Ideally, the results of field experiments and the underlying theories would also inform the design of “combination” policies, so that the two approaches are both policy relevant and complementary. Field experiments need theory to derive specific testable implications and to give a general idea of the interesting questions. Field experiments also make it possible to test empirical predictions. Scaling up by generalizing from experiments is the next step to consider. Well-designed program evaluations are, in effect, international public goods. They provide information to other countries about what might work and what might not. As a result, they are very important for international agencies that seek to introduce related programs in different countries.

Several reasons explain why the results of a well-executed experiment may not always be generalizable (Duflo 2004; Banerjee and Duflo 2008). First, the experiment may affect the treated or the comparison sample, for example, if the provision of inputs temporarily increases morale among beneficiaries, thereby improving performance (the so-called Hawthorne effect). While this factor would bias randomized evaluations, it would also bias other types of evaluation, including econometric techniques such as fixed-effect or difference-in-differences estimates. The two groups may also be temporarily affected by their participation in the experiment (the “John Henry effect”). However, such effects are less likely to be present for large-scale evaluations if the time span is long enough or if the program is large.

Finally, it is never possible to replicate a project identically: circumstances vary, and ideas need to be adapted to local contexts.

Duflo's work also points to a number of important recommendations for the design and implementation of evaluations. These include reducing the number of evaluations; conducting credible evaluations in key areas, combined with randomized evaluation in other areas as opportunities occur; and establishing dedicated evaluation units in (international) organizations. She also argues that it is as important to publish negative results as it is to publish positive results and calls for institutions to ensure that negative results are also systematically disseminated, as is already the case for medical trials.

Innovation Policy Reviews

The importance of policy evaluation is twofold. First, it is important to learn from experience which policies and programs work and which do not and under what circumstances. These lessons are especially important as circumstances may change rapidly, and, as new forms of innovation emerge, innovation policies need to reflect these developments. Second, policy evaluation is essential as a guide for public spending on R&D and for resource allocation more generally. Many types of evaluation are possible, the quality varies widely across projects and countries, and feedback of the results into policy making is not always sufficient. A lack of transparency is also a relatively common hindrance.

The National Level

The evaluations of national policies are largely based on the experience of the OECD. They began some 50 years ago in the 1960s, primarily with science policy, and have shifted gradually toward review of broader innovation systems and policies.

The OECD's Innovation Policy Reviews. The OECD's reviews of national innovation policies provide a comprehensive assessment of the innovation system of individual countries.⁵ They focus on the role of government and provide recommendations on improving policies that affect innovation performance. Each review identifies good practices that might be useful in other countries. An evaluation of a country's innovation policies and systems is prepared and then peer reviewed by an OECD committee of government officials and national experts in the field of innovation policy.

The reviews undertaken to date provide some important insights into the efficacy of innovation policies and the use that can be made of them. They look not only at government policies for stimulating innovation directly but also at broader factors, such as the overall governance of the innovation system and

the extent to which the changing nature of innovation affects the linkages between innovation and economic performance and, hence, policy priorities. The reviews, in conjunction with associated work by the OECD and other international organizations, also draw attention to some important issues for developing economies. These include the strategies that can be adopted to build up innovation capabilities, policies to move up the value chain, and the importance of history and path dependence on former industrial experience or economic regime. Box 7.5 presents some of the key initial findings of the OECD reviews on market and governance arrangements for innovation and policy instruments and the policy mix.

Reviews of the overall innovation system are an important complement to studies of individual programs and policies. Not only do they provide a broader perspective on the activities of governments, but also they make it possible to assess the overall coherence of the policies adopted to support innovation. These may include policies that affect innovation indirectly, such as competition policy and the openness of the economy to international trade, investment, and migration. Such economy-wide factors are an important part of the innovation system.

Strategies to Build Innovation Capabilities. Countries may develop innovative capabilities as part of their catch-up strategy, with a range of positive effects. Developing domestic human resources and other forms of scientific capabilities increases a country's attractiveness as a location for foreign investment and enhances its absorptive capacity, raising the extent to which domestic companies and institutions can take advantage of spillovers and technology transfer from inward foreign investment. At the same time, this process enables the country to diversify its activities and reduce its dependence on any one activity or sector. It will also enable it to link more closely with the activities of globalized economies and tap into new markets.

The innovation systems of developing and emerging economies share certain weaknesses, including a lack of skilled human resources, inadequate innovation capabilities in business firms, and poor coordination among industry, universities, and public research organizations. These weaknesses need to be addressed in innovation policies, the implementation of which will require good policy governance. The OECD's innovation policy reviews of Chile, China, Korea, Mexico, and South Africa and provide examples of effective policy and governance reforms aimed at developing new areas of comparative advantage. Korea's experience, for example, illustrates the importance of significant stocks of science and technology capabilities for implementing imitation strategies, for moving up the value chain, and for speeding up the catching-up process (OECD 2009). China illustrates the benefits of large-scale investment in science and technology infrastructure, including human resources for science and technology (OECD 2008b).

Box 7.5 National Innovation Strategies: Lessons from OECD Country Reviews

Market and Governance Arrangements for Innovation

Improving framework conditions: Lack of competition acts as a barrier to innovation in many countries, but there is too little awareness of the role of competition policy in fostering innovation.

Policy coordination and participatory governance: These are important for ensuring effective policy coordination and effective participation by stakeholders.

Leadership: Involvement of the highest level of government is needed for securing policy attention and commitment.

Commitment: It is important to ensure that public funding for innovation is not “crowded out” by short-term demands.

Stability and predictability of institutions and policy delivery: While innovations in the policy framework are sometimes necessary, frequent changes tend to be counterproductive.

Evidence-based policy making: It is important to make effective use of reviews and evaluations. However, policy learning is easily disrupted and often difficult to institutionalize.

Steering and funding of public research organizations (PROs): The role of PROs needs to be redefined and their connection to the business sector improved to enhance their contribution to the overall performance of the innovation system.

Policy Mixes and Instruments

Striking a balance in policy instruments: Balance is important for stimulating business innovation, as policies are often introduced along several dimensions. Some are top-down, especially when the need for a change in direction is clear, but others are bottom-up. Some aim at improving economy-wide capabilities, such as policies to reduce financial barriers to investment in innovation, while others have specific policy objectives, such as tax credits for R&D.

Building capabilities: This can be done, for example, by reducing financial barriers to investment in innovation.

Direct and indirect support measures: Ideally, these two types of tax incentives should be applied in a complementary way to make the best use of their respective advantages, but this is not always the case.

Bottom-up and top-down approaches: These approaches should be complementary. Bottom-up approaches should be used for standard types of innovation projects and for gathering information and inducing self-organization in new areas, for example, by competitive calls. Top-down approaches should be used for changes in policy directions.

Different types and combinations of support: This support can include individual project-based support, ad hoc support, consortium-based, and longer-term support. Consortia are also useful for triggering behavioral change, such as cooperation between different types of actors.

Competition for funding: The shift toward competitive funding has provided powerful incentives for PROs and universities, while safeguarding a degree of stability and maintaining capabilities.

Source: Adapted from Guinet and Keenan 2008.

Long-term reform is not easy, although the potential gains may be large if it is successfully carried out. Becoming locked into wrong technologies or infrastructure is a potential danger, as is the danger that special interest groups will capture the project. For that reason, the governance of innovation projects matters.

Moving Up the Value Chain and Diversification. Another area of interest for developing countries is the use of innovation policy to move up the value chain. As globalization intensifies competitive pressures, many countries respond by trying to diversify their economies and move up the value chain. For example, Mexico and Hungary have an important manufacturing base, driven by inward foreign investment, which largely functions as an export platform (to the United States and the European Union, respectively). Hungary produces and exports many medium- and high-technology goods, in spite of the relatively low R&D intensity of the country's firms. This fact indicates that national innovation policy should aim for better integration of the foreign-owned sector into the national innovation system, including universities and public research organizations, and for improved absorptive capacities of domestic small and medium enterprises.

The economies of some countries are highly specialized, which may present a risk in the long term. The economies of Chile, Mexico, and Norway, for example, are largely based on natural resources, and Luxembourg is dominated by its financial sector. Diversification may also be desirable for countries with a small domestic market or a remote geographical location, such as New Zealand. National innovation policies can help meet such challenges. Norway has seized opportunities for knowledge-intensive activities in and around the oil and gas sector (OECD 2008c). Chile and New Zealand are adopting measures to aid a shift toward more innovation-based growth strategies (OECD 2007a, 2007b). Developing human resources is crucial to any strategy for innovation-based growth. For example, Chile, where the lack of skilled human resources constitutes a significant bottleneck, is adopting measures to raise educational standards to international levels, among others. In New Zealand, more emphasis needs to be put on improving framework conditions and stimulating market-led innovation throughout the economy, including by stimulating entrepreneurship and developing management, marketing, and distribution skills.

History and Path Dependence. History and path dependence are a significant issue for developing countries. Existing institutions have norms and routines that are reflected in their day-to-day operations. While such features provide stability and can thus be a positive factor, they may also result in inertia and prevent institutional reform. All countries face such risks, irrespective of their degree of development or the state of their innovation policies and institutions. If the nature of innovation and the dominant technologies change, as

they do at present, this fact needs to be reflected in the innovation policies and strategies pursued.⁶

Reported Evaluation Practices. In Norway and Switzerland, for example, evaluation is common and attempts to follow internationally accepted best practice. Programs are often legally required to undergo ex ante and ex post evaluation, as well as ongoing monitoring during implementation. In Switzerland, domestic and foreign experts are involved in the evaluation process and often contribute to the development of evaluation methodologies. In Norway, evaluation is actively promoted by key agencies such as Innovation Norway and the Research Council of Norway.

The importance of evaluation has been recognized in many other countries, such as Chile and Korea, and efforts are underway to catch up with practices elsewhere. Korea has recently introduced a large-scale evaluation system, involving a combination of program evaluations and meta-evaluations (that is, those with strict performance targets). Interim evaluation results have been used to modify the resources made available to particular programs. Other countries have as yet made relatively little effective use of evaluations of innovation policies and programs.

It is very important for the results and findings of evaluation exercises to be used in subsequent evidence-based policy making. However, the differences in the extent to which policy making is evidence based are significant. Effective evidence-based policy making is not easy to implement, as it requires resources and substantial expertise, as well as clearly defined objectives. Outcomes also need to be measurable, which again may require investment in new resources and institutions.

How Can Developing Countries Use These Reviews? The OECD country reviews of innovation policies cover a variety of countries at different stages of development and innovation performance. As the reviews evaluate different policies and programs, their implementation, and the governance of the innovation system and also identify best practices and make recommendations, countries can learn about what works and what does not and under what circumstances. These results can inform the design and implementation of innovation policy in developing countries if they are adapted to local characteristics. However, policy measures that are effective in one country may be ineffective or inappropriate in another, depending, for example, on institutional factors, industry specialization, and size. A country's innovation performance depends not only on its performance in each element of the national innovation system but also on how the elements interact. OECD analysis suggests that no single combination of elements is successful: what matters is the cohesiveness of the system for innovation performance and how well the country performs in each of the main dimensions.⁷

Recommendations from the OECD reviews of particular importance for developing countries include improving framework conditions for innovation (that is, fostering competitive and open product, labor, and capital markets); implementing and enforcing intellectual property legislation; ensuring a supply of appropriately skilled people by improving access to higher education (including vocational training); improving incentives for firms to invest in training; building capacity in small and medium enterprises; and promoting and supporting entrepreneurship.

UNCTAD also carries out science, technology and innovation policy reviews specifically designed to help developing countries identify and adjust their policies and institutions to support technological transformation, capacity-building, and enterprise innovation. At the time of writing, reviews had been completed for Colombia, Iran, and Jamaica.

Regional Level

New microlevel work coordinated by the OECD has looked at the regional dimension of innovation, including firm location and linkages (OECD 2008a). Linkages among geographic areas and between firms result from the flow and transfer of intellectual assets and knowledge spillovers, which often require proximity. Evidence points to significant variations in the inventive performance of regions, with a high concentration in certain regions of continental Europe, North America, and Japan. The development of inventive activities in countries tends to take place in a small number of regions, and highly inventive regions usually cluster together. This spatial dependence has increased over time. Moreover, the inventive performance of regions is directly influenced by the availability of human capital and R&D expenditure. Cross-country differences point to the importance of national innovation systems and of linkages within firms across regions, as the most inventive regions have relatively more multiregional companies among their innovative firms.

Governments are increasingly aware that the regional dimension of innovation matters for strategies that use innovation to promote growth. The OECD is carrying out work on regional innovation to help policy makers from different backgrounds at both national and regional levels. Objectives include strengthening the evidence base for policy making, improving the use of resources in different regional contexts, ensuring coherence between innovation and other policy objectives, and assessing the impact of policies at the regional and national level. Current work includes an ongoing series of reviews of innovation in regions from national and regional perspectives (Italy, Mexico, and United Kingdom) and an analysis of innovation indicators using the OECD Regional Database's innovation data set. Regional policy initiatives are evaluated in OECD (2007c).

The origin of national and EU programs to support clusters and regional specialization can be found in regional, S&T, and industrial policies. Several

programs that originated in S&T policy specifically support large-scale collaborative R&D projects to stimulate the most promising technology sectors in regions in which key institutions, researchers, and firms are concentrated. However, the evaluation of these approaches is often inadequate, especially since not all programs are evaluated and tools to measure impacts are often lacking. As a result, it is difficult to assess whether these programs are appropriate, realistic, and flexible enough to achieve their goals. The stated goals of cluster and regional specialization programs are often vague or broad (OECD 2007c), complicating the identification of appropriate participants, duration, targets, budgets, and funding. Cluster policies may also lack the private sector engagement on which their long-term effectiveness depends.

Overall, there are three main issues for policy and program design, based on practices across OECD countries: the degree to which the programs are appropriate, realistic, and flexible enough to achieve their goals; policy coherence within and across different levels of government; and the importance of private sector engagement to the ultimate outcomes.

Conclusions

It is clear from the above discussion that evaluation practices of innovation systems and programs are yet embryonic in emerging and developing countries. However, it is also clear that methods, surveys, and reviews adapted to their needs are being increasingly elaborated and implemented. More and more countries are using them and providing pioneering examples that can inspire the whole community. Several stand out as particularly important:

- *The development of “macro” benchmarking methods and indicators.* These take due account of emerging and developing countries’ particular situations, do not measure innovation performance or capabilities exclusively with R&D-related indicators, and allow an accurate appraisal of improvements over years.
- *The implementation of innovation surveys.* These surveys capture evolutions of fundamental importance for emerging and developing economies, such as the diffusion of new and basic technologies and improving productivity, welfare, or the environment. The use of the newest tracking methods, such as geographic information systems, should be strongly encouraged.
- *The systematic evaluation of policy programs.* These take into consideration intangible developments, such as network and competence building, that always precede visible technical or economic achievements. They also make use of innovative approaches to examining program implementation and impact, such as field experiments and tests.
- *The use of national policy assessments.* These assessments use standard approaches such as international peer reviews that involve foreign expertise,

including experts from the developed world and the donor community, to stimulate mutual learning processes.

- *New types of indicators.* Finally, it is of the utmost importance to develop new types of indicators that go beyond the usual quantitative measurement of economic growth to integrate more qualitative dimensions. Significant attempts in this direction are summarized in box 7.6.

Box 7.6 Beyond GDP: Alternative Measures and Indicators of Economic and Social Progress

Gross domestic product is widely used by economists and the public at large to gauge the health and welfare of a nation. However, “if ever there was a controversial icon from the statistics world, GDP is it. It measures income, but not equality, it measures growth, but not destruction, and it ignores values like social cohesion and the environment” (OECD 2005a). Challenges to the use of GDP as a standard measure of comparison between countries reached a new high after the recent global economic crisis and the rise in consciousness over climate change (for example, GDP treats loss of ecosystem services as a benefit instead of a cost). Some “alternative” measures that attempt to include the social dimension already exist, although GDP is often used as a basis. These include, among others, the UN Human Development Index (HDI) and the Bhutan Gross National Happiness Index (GNH).

The most widely used alternative measure is the HDI, which is a composite index that combines normalized measures for three dimensions: (a) life expectancy at birth, as an index of population health and longevity; (b) knowledge and education, as measured by the adult literacy rate (with two-thirds weighting) and the combined primary, secondary, and tertiary gross enrollment ratio (with one-third weighting); and (c) standard of living, as measured by the natural logarithm of GDP per capita at purchasing power parity.

For its part, the GNH is an attempt to define quality of life in more holistic and psychological terms than GDP, and is used by Bhutan in its development strategy. The concept of GNH is based on the premise that the true development of human society takes place when material and spiritual development occur side by side to complement and reinforce each other. The four pillars of GNH are the promotion of sustainable development, preservation and promotion of cultural values, conservation of the natural environment, and establishment of good governance.

Reflecting the general dissatisfaction with GDP as a measure, some initiatives have also proposed a revision of the measure itself. In February 2008, for example, the French president, Nicolas Sarkozy, asked a commission of esteemed economists and statisticians to look more closely at GDP, identify its limits as an indicator of economic performance and social progress, consider what additional information may be required to produce more relevant indicator(s), and assess the feasibility of alternative measurement tools. The underpinnings of this initiative were clear: “What we measure affects what we do. If we have the wrong metrics, we will strive for the wrong things. In the quest to increase GDP, we may end up with a society in which most citizens have become worse off” (Stiglitz 2009). The commission has recently

continued

Box 7.6 continued

submitted its report, which gives a series of recommendations for improvements to the existing measure of GDP, as well as the possible construction of new indexes that will better measure social well-being and sustainability of growth. Some recommendations include looking at income and consumption rather than at production, as these are more likely to reflect material living standards; emphasizing the household perspective, as household income has often been quite different—and much lower—than GDP growth; giving more prominence to the distribution of wealth in measurement of economic progress; broadening measurement to nonmarket activities; and increasing measurement of sustainability.

Source: Author, based on Stiglitz 2009; Stiglitz, Sen, and Fitoussi 2009.

At a time when the world community is experiencing major challenges in coping with a deep economic slowdown and a mounting environmental crisis, it is more important than ever to develop methods that allow a better allocation of resources. Demonstrating by rigorous methods what works and what does not work in the field of innovation is paramount, since innovation is the basic factor of economic growth and more generally for adaptation to social challenges.

Notes

1. The WEF defines competitiveness as the collection of factors, policies, and institutions that determines the level of productivity of a country and that, therefore, determines the level of prosperity that can be attained by an economy.
2. The 113 variables included in the GCI are grouped into 12 pillars, each of which reflects one aspect of competitiveness. The 12 pillars are institutions, infrastructure, macroeconomic stability, health and primary education, higher education and training, goods market efficiency, labor market efficiency, financial market sophistication, technological readiness, market size, business sophistication, and innovation.
3. To enable a proper comparison across the indexes, the indexes had to be adjusted so that each index ranks a common set of countries. In other words, for this exercise, any country that was omitted from any one index was dropped from the remaining three. This process resulted in a set of 122 countries common to all four indexes.
4. The United States does not carry out the equivalent of a “community innovation survey.” Instead, it conducted a pilot survey in 1994, a survey on innovation in the information technology sector in 2001, and, more recently, a business R&D and innovation survey in 2008.
5. See <http://www.oecd.org/sti/innovation/reviews>.
6. See also the work carried out under the OECD Innovation Strategy: <http://www.oecd.org/innovation/strategy>.
7. To assess national innovation performance within the context of national innovation systems, the OECD recommends analysis of the following 10 areas: demand; human resources; finance; physical inputs; access to science, technology, and business best practice; ability and propensity of firms to innovate; effectiveness of market processes; networks, collaboration, and clusters; institutions and infrastructure; and business environment (OECD 2005b).

References and Other Resources

- Archibugi, Daniele, and Alberto Coco. 2004. "A New Indicator of Technological Capabilities for Developed and Developing Countries." CEIS Research Paper 15 (44) from Tor Vergata University, Centre for Economic and International Studies, Rome.
- Banerjee, A. V., and E. Duflo. 2008. "The Experimental Approach to Development Economics." NBER Working Paper 14467, National Bureau of Economic Research, Cambridge, MA.
- Duflo, E. 2004. "Scaling up and Evaluation." Paper prepared for the Annual World Bank Conference on Development Economics, "Accelerating Development," May 21–23, 2003, Bangalore, India.
- . 2006. "Field Experiments in Development Economics." Paper prepared for the World Congress of the Econometric Society, January, Boston, MA.
- European Commission. 2007. *European Innovation Scoreboard 2007: Interactive Benchmarking*. <http://www.proinno-europe.eu/>.
- Georgioui, L., K. Smith, O. Toivanen, and P. Yla Anttila. 2003. *Evaluation of the Finnish Innovation Support System*. Helsinki: Ministry of Trade and Industry.
- Guellec, D., and B. van Pottelsberghe. 2000. "The Impact of Public R&D Expenditure on Business R&D." STI Working Paper 2000/4, OECD, Paris.
- Guinet, J., and M. Keenan. 2008. "National Innovation Strategies: Lessons from OECD Country-Specific Work." Presentation for the OECD–World Bank conference, "Innovation and Growth," Paris, November 18–19, 2008. <http://www.oecd.org/dataoecd/60/17/41709674.pdf>.
- ITIF (Information Technology and Innovation Foundation). 2009. *Benchmarking EU and US Innovation and Competitiveness*. Washington DC: ITIF.
- Jaumotte, F., and N. Pain. 2005a. "Innovation Policies and Innovation in the Business Sector." Economics Department Working Paper 459, OECD, Paris.
- . 2005b. "An Overview of Public Policies to Support Innovation." Economics Department Working Paper 456, OECD, Paris.
- OECD (Organisation for Economic Co-operation and Development). 2002. *Proposed Standard Practice on Research and Experimental Development: Frascati Manual*. Paris: OECD.
- . 2005a. *Guidelines for Collecting and Interpreting Innovation Data: Oslo Manual*. Paris: OECD.
- . 2005b. *Innovation Policy and Performance: A Cross-Country Comparison*. Paris: OECD.
- . 2006. *Science, Technology and Industry Outlook 2006*. Paris: OECD.
- . 2007a. *OECD Reviews of Innovation Policy: Chile*. Paris: OECD.
- . 2007b. *OECD Reviews of Innovation Policy: New Zealand*. Paris: OECD.
- . 2007c. *OECD Reviews of Regional Innovation: Competitive Regional Clusters: National Policy Approaches*. Paris: OECD.
- . 2007d. *Science, Technology and Industry Scoreboard*. Paris: OECD.
- . 2008a. *Creating Value from Intellectual Assets: Synthesis Report*. Background report for the meeting of the OECD Council at Ministerial Level. Paris: OECD.
- . 2008b. *OECD Review of Innovation Policy: China*. Paris: OECD.
- . 2008c. *OECD Review of Innovation Policy: Norway*. Paris: OECD.
- . 2008d. *Science, Technology and Industry Outlook 2008*. Paris: OECD.
- . 2009. *OECD Review of Innovation Policy: Korea*. Paris: OECD.
- Poot, T., P. den Hertog, T. Grosfeld, and E. Brouwer. 2003. "Evaluation of a Major Dutch Tax Credit Scheme (WBSO) Aimed at Promoting R&D." Presentation at the FTEVAL conference, "The Evaluation of Government Funded R&D," Vienna, Austria, May 15–16.
- Ravallion, M. 2009. "Evaluation in the Practice of Development." *World Bank Research Observer* 24 (1): 29–53.

- Stiglitz, Joseph E. 2009. "Towards a Better Measure of Well-Being." Available on FT.com (Sept. 13, 2009).
- Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress. http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.
- UNCTAD (UN Conference on Trade and Development). 2005. *World Investment Report 2005*. Geneva: UNCTAD.
- UNDP (United Nations Development Programme). 2001. *Human Development Report 2001: Making New Technologies Work for Human Development*. New York: UNDP Report Office.
- Wagner, Caroline S., Edwin Horlings, and Arindam Dutta. 2004. "Can Science and Technology Capacity Be Measured?" Unpublished manuscript. <http://www.google.com/search?q=Can+Science+and+Technology+Capacity+Be+Measured&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a>.
- WEF (World Economic Forum). 2008. *Global Competitiveness Report 2008–2009*. Geneva: WEF. <http://www.weforum.org/pdf/GCR08/GCR08.pdf>.