Annals of Science and Technology Policy

Measuring Science, Technology, and Innovation: A Review

Suggested Citation: Bronwyn H. Hall and Adam B. Jaffe (2018), "Measuring Science, Technology, and Innovation: A Review", Annals of Science and Technology Policy: Vol. 2, No. 1, pp 1–74. DOI: 10.1561/110.00000005.

Bronwyn H. Hall
University of California at Berkeley, USA and University of Maastricht, The Netherlands;
bhhall@econ.berkeley.edu

Adam B. Jaffe
Brandeis University, USA;
ajaffe@brandeis.edu

now
the essence of knowledge
Boston — Delft

Contents

Measuring Science, Technology, and Innovation: A Review

Bronwyn H. Hall[1] and Adam B. Jaffe[2]

[1] *University of California at Berkeley, USA and University of Maastricht, The Netherlands; bhhall@econ.berkeley.edu*
[2] *Brandeis University, USA; ajaffe@brandeis.edu*

ABSTRACT

The measurement of scientific, technological, and innovative activities (STI) in the economy is an increasing challenge faced by statistical agencies around the world. In this review, we survey the current state of the art. We discuss the concept of indicators, their quality and use, and present a schematic model of the STI system that can help us identify gaps in the set of indicators commonly in use. We then review the developments in STI measurement that have taken place in the rest of the world, particularly the widespread use of innovation surveys. The monograph concludes with a discussion of the measurement gaps and issues in the U.S., which we identify as innovation (especially in the service sector), non-R&D investment related to innovation, data timeliness, data linkages, measurement related to public policy goals, and the sources of capital for innovation.

Bronwyn H. Hall and Adam B. Jaffe (2018), "Measuring Science, Technology, and Innovation: A Review", Annals of Science and Technology Policy: Vol. 2, No. 1, pp 1–74. DOI: 10.1561/110.00000005.

1

What is an indicator?

An "indicator" is a set of facts or observations that tells us something meaningful about the underlying phenomenon of interest, in this case science, technology and innovation (henceforth STI). In order to evaluate whether a particular set of facts is a good indicator, or to determine whether there are indicators we should have that we do not, we need to consider in some detail what makes a particular set of facts *meaningful*. In this section we address the important dimensions that affect the meaning and usefulness of an indicator.

The foundation of an indicator is a set of data. The word "data" is a plural of the Latin *datum* meaning "given." We frequently think of data in the form of numbers, but conceptually data can be either quantitative or qualitative. Either way, formal analysis of data is predicated on the idea that the data are, indeed, given, meaning that they are generated by the world itself rather than created by the people who want to use them. In reality, what we call "data" in most contexts are numbers or qualitative observations that are usually collected by human beings, and the nature of this collection process combines with the "givens" of the world to determine what are recorded as "data."

For many purposes – particularly their use by policy makers and other "lay" people concerned about science, technology and innovation – presentation of data without further processing does not constitute a meaningful indicator. All of the indicators published by NSF are constructed in some way from underlying data. So any recommendation about an indicator is, in a sense, a compound recommendation: (1) what data should be collected, and (2) how should those data be summarized in a published indicator or indicators?[1]

1.1 The relationship between indicators and a framework for analysis of STI

In order to determine whether an indicator is meaningful, we need to assess both the data that are used, and the manner in which those data are summarized. Carrying out this assessment requires specification of the underlying concept we are trying to understand, and the relationship between this concept and the process that generates the data. For example, we tabulate how many academic degrees of various kinds are granted each year. We don't really care about degrees per se, but we care in some way about the knowledge and skills of the population, and we believe that the number of degrees granted and the fields in which they are granted is in some way informative about the accumulation of knowledge and skills. We measure income in various ways. For some purposes, we might care about income itself, but more often we are using income as an indicator of well-being or happiness or success. Sometimes the proxy nature of our measurements is explicit; in other cases it is implicit. But either way, we cannot assess how meaningful an indicator is without thinking about its relationship to the concept or concepts that we really care about.

But where does the statement "the number of degrees granted is in some way informative about the accumulation of knowledge and skill"

[1] Data collected by the government are also used in disaggregated form (i.e. without creation of summary indicators) by social scientists and others interested in studying the underlying phenomena. We return below to the importance of the inter-relationship between the collection of data in support of indicators, and the use of data for research.

come from? And why do we think that the population's accumulation of knowledge and skill is itself something we care about? The determination of both what aspects of the real world matter, and what measurements are illuminating with respect to these aspects will be based on some kind of framework, either explicit or implicit. A framework is an abstract representation of the world, typically focusing on one or a small set of aspects (e.g. science, technology and innovation). Such a framework can be constructed and described in a variety of ways: economists like to use symbols and equations; anthropologists tend to use words; systems engineers frequently use diagrams. Whatever the mode of presentation, the framework is designed to capture the essence of the underlying complex reality. Even if such a framework is not fully articulated, it will underlie the choice of data collected and the interpretation placed on the resulting indicators. In this report, we make the framework we are using explicit, because we believe that leads to greater clarity.

Having a framework for analysis is also important for the question of how data and indicators can inform public and private decision-making. In order to use data to inform decisions, we again need an understanding of what those data tell us about the workings of the STI system. Either explicitly or implicitly, we have some kind of framework in which those data relate to the instruments or the goals of our decisions. Again, we believe that being explicit about the framework leads to greater clarity in understanding how the data or indicators relate to the goals and instruments of policy.

Note that the relationship between any framework for analysis and data is an iterative one. We need some kind of framework to organize ideas and to know what data we should be looking for, and to support interpretations of the data that we see. At the same time, social scientists and others will use the data to test various models, by studying the extent to which the predictions of their models are borne out in the data. This process may suggest changes to the framework used and also to the data that is needed for analysis. We return below to the implications of this two-way relationship between the data and the model.

In Section 2, we summarize the framework for analysis of the STI system that is most widely used by economists. In this framework, for example, the level of knowledge and skills of the population is important for several reasons: it affects the amount of goods and services that can be produced per capita, and it also affects the rate at which *new* knowledge is produced, which affects the improvement over time in the amount of goods and services that can be produced. In Section 3, we take the next step, and consider how the specific data currently collected by the U.S. government, and indicators constructed from those data, relate to the important concepts within the framework. Section 4 expands this discussion by reviewing data collected and indicators published by other countries and international organizations. Section 5 then turns to the issues of policy, and draws explicit connections between important policy questions and indicators, using the framework of Section 2. Section 6 then builds on these discussions to highlight gaps and issues with the existing indicators. Section 7 provides concluding comments.

1.2 Uses of STI indicators

In addition to understanding the data underlying an indicator, how those data were processed, and the relationship of the data and its processing to a framework for analysis of the STI system, evaluating indicators also requires an understanding of the purposes for which they are used.

Performance assessment and benchmarking

Some indicators serve as performance measures that give an assessment of whether the STI system or some component thereof is doing better or worse over time, and better or worse than some comparison group (e.g. other countries). For this purpose, the indicator may stand for some aspect of the system that is intrinsically valued (e.g. income), or it may stand for an aspect of the system that, within the framework, is understood to have important impacts on aspects that are valued. For example, past research has suggested that, all else equal, a greater intensity of investment in new knowledge will lead to higher rates of productivity growth and income growth. For this reason, one might focus on the R&D/GDP ratio over time or across countries as a benchmark of innovative activity. But in doing so, one should always keep in mind the role of knowledge investment in the framework, and the relationship between the framework concept of knowledge investment, and the R&D data that are actually collected. In this example, it is clear that spending on R&D may not capture all of the investments that are directed towards creating new knowledge (and may capture some investment that is not thus directed). In addition, the share of innovation-related investment captured by an R&D measure may be changing over time so that the measure is potentially misleading taken on its own. The framework also makes clear that there is no particular level for this ratio (such as 3%) that is optimal. Finally, even though the relationship between R&D investment and productivity growth is well established, it is still true that R&D investment is an indicator of innovative activity, not an indicator of innovation, which is the concept we think is more closely associated with the growth outcomes of interest.

Informing public policy decisions

An important function of STI indicators is to provide an informed basis for public policy decisions. But of course policy is not intended to affect the indicators, it is intended to affect the underlying concepts of interest. So to determine if a suite of indicators is well-suited to inform public policy, one needs to identify the goals of public policy, and to use the framework for analysis of STI to understand how those goals relate to data that can be collected. We discuss these issues in Section 5.

Informing private sector decision-making

Firms and individuals in the for-profit sector also use STI indicators to make business decisions, and not-for-profit organizations (e.g. universities) use them to make decisions in pursuit of their missions. It is unclear to what extent these parties needs and desires are different from those of policymakers, and, if they do differ significantly, to what extent these distinct needs are considered in the decisions made by government statistical agencies about the STI indicators system.

Facilitating social science research

As noted above, social scientists use data to test the implications of models, and thereby refine the models. Hence their interest in the indicators endeavor is more in the collection and availability of data than in indicators *per se*. But because model testing and validation is so important, we will comment in the concluding section on how the needs of social scientists might be considered in the context of recommendations about indicators.

1.3 Issues of data collection and indicator construction

As noted above, though formal statistical theory treats data as "given," in reality the potentially messy process by which the data are created and compiled is sometimes important in evaluating indicators. In this section we identify a few basic issues.

Data dimensions

Data collected to construct indicators may span time, space and institutional categories. Construction of summaries to be reported as indicators can then be aggregated along one or more of these dimensions, depending on the purpose of the indicators. Reporting of indicator values in a time series is frequently of interest; this raises issues of consistency of interpretation of the summary values based on data collected at different points in time. Comparisons across different geographic areas, or across different categories of institutions (e.g. small firms versus large firms), are also of interest. When comparing data from the same point in time, but different institutions or geographic areas, there is an issue of comparability that is analogous to the consistency issue when comparing over time. For analytical purposes, it is often useful to be able to construct a *panel*, in which the indicator of interest varies across both time and another dimension such as geography or institution type.[2] The need for comparability can come into conflict with the need to

[2]For micro-data analyses, it is often valuable to have access to longitudinal data, in which data from the same individuals or institutions is collected at multiple points in time.

redesign surveys and data collection methods in response to a changing landscape, especially in the case of indicators describing innovation.

How the data were generated

Some data collected for use in STI indicators are generated by companies or other organizations in the normal course of their business for their own internal use; some are also generated by these organizations for other external reporting requirements (e.g. financial reporting required by Generally Accepted Accounting Practices). Other data are generated by organizations specifically in response to government requests tied to STI data collection. The advantage of using data otherwise generated in the course of business is that it may be less burdensome for the entities involved. Also, the intrinsic importance of the data elements to the organizations may lead to greater care and consistency in construction, compared to data that are generated only because NSF asks for them. On the other hand, precisely because the data are created for other purposes, they may be less ideally suited to the intended purpose than data that can be specified with their intended STI indicator use in mind.

How the data are collected

For data that are generated for other purposes, there is also a potential distinction between data that are collected by explicit survey requests to the entities that created the data, and data that can be collected passively, i.e. without the active assistance of the generating entities. Examples of passive collection include mining of data from reports filed by companies with agencies such as the SEC or IRS, and "scraping" of data from websites. The advantage of passive collection is minimization of burden on the generating entities. In some cases, there may also be an advantage inherent in the absence of an opportunity for the entities that generate the data to manipulate strategically what is reported. On the other hand, when passively collecting data generated for other purposes, it may be difficult to understand fully what the data really mean and difficult to control the sampling frame to ensure representativenes.

Public versus private data collection

When we speak of indicators, we tend to think of data collected and published by the government. But social science researchers can and do collect data themselves. What is the appropriate mix of government and non-government collection of data and publication of indicators? Data that are of wide potential use have an important public-good attribute. But researchers studying STI can apply for public funding of their research, thus solving the public good aspect of the problem without having the government be the party that collects and publishes the data and indicators. The government does have particular advantages as data collector, where the data in question are related to other data that the government collects for statutory purposes (e.g. census data), and/or the data are considered proprietary and therefore unlikely to be provided by private agents without government mandate. But, by definition, this mandatory data collection principle can only be applied to a limited number of data elements. Recognizing the scope for individual investigators to collect a variety of information beyond what is collected by the government greatly increases the potential scope of indicators available to public and private decision-makers. This is particularly true with respect to data that can be collected passively, as discussed above, because in such a case the requirement of mandatory compliance that the government can bring to its surveys is not necessary. Thus, in considering the possibilities that may be created for additional indicators by passive data collection, we should not assume that these data have to be collected and published by the government. There is considerable scope for individual researchers, probably with public research funding, to collect and publish such data.

1.4 Data quality

Before moving on to the framework for analysis of indicators within the STI system, we pause briefly to review measurement issues and how they affect the quality of data. This section draws heavily on Griliches (1986), which can be consulted for more detail.

Griliches categorizes issues of data quality as falling into three categories: extent, reliability and validity. Extent refers to the scale and scope of the data: for how many years has it been collected, how many different data items are collected, and how broad is the coverage in terms of regions or types of institutions. With respect to extent, more is better, in a fairly straightforward way.

Reliability refers to the inherent reproducibility of the data collection process itself. In effect, it asks, if the data were collected on different occasions or in different places, and nothing real about the world were different in the two times or places, how close would the different data be to each other? They won't typically be the same, because each measurement includes random sampling error. Reliability is a measure of the signal-to-noise ratio, i.e., the fraction of the variance in the data that is systematic rather than random.

Validity refers to the extent to which the data are generated and collected in a manner that makes them correspond closely to the underlying concept that we care about. Griliches subdivides validity into considerations of relevance and representativeness. Relevance is, in essence, the question of the extent to which what we are measuring is closely related to what we care about. Representativeness arises in any context in which we collect only a data sample, i.e. we do not collect data from all relevant individuals or units. It refers to the extent to which the different kinds of units had the same probability of contributing to the data (or if the probabilities differed, they did so in a systematic way that we can account for).

To illustrate these concepts, consider the NSF industrial R&D survey. It is very extensive, because it goes back many decades, and covers all firms and all parts of the country. It is reliable, in part because it is so extensive, and in part because the definitions and procedures are well established and understood. But how valid it is really is a question of what underlying STI concept you wish to use it to measure. Its relevance is potentially questionable, because it looks only at formal R&D activities, which may not be the right thing to look at if what we care about is overall knowledge investment. In the past (prior to redesigns in 1992 and 1995) its representativeness may have been questionable, because the sampling frames failed to adequately sample small and medium-sized firms and firms in the service sector.[3] These observations don't mean the data are useless; they simply make the point that the validity of a set of data can only be judged in the context of the framework for analysis of STI and the role to be played by those data within the framework.

As noted above, there is variation in the extent to which the potential gap between the measured item and the model concept is explicitly acknowledged. Virtually everyone who uses patent statistics to measure the rate of invention or innovation notes explicitly that patents are only a "proxy" for the underlying concept. But the issue of proxies is really more one of degree than of kind: for virtually every indicator we use we must acknowledge some degree of potential distance between the measurement and the concept it stands for. Rather than viewing "proxy" as a mild epithet that applies to some indicators but not others, it makes more sense to consider, in all cases, the extent to which a particular indicator might deviate from the underlying concept for which it stands.

It is also worthwhile to consider how the choice of indicators and method of data collection affects their reliability and validity. As noted above, there may be significant benefits in terms of government resources and burden on private parties to expanding the use of data that is passively collected through methods such as web scraping, instead of government surveys. With statistically designed surveys, however, we have both a high degree of reliability, and, perhaps more important, a clear understanding of the level of reliability and possible sources of unreliability. With passive methods, it may be much harder to assess reliability, precisely because the relationship between the captured data and the universe of underlying activity is not known.

Finally, collection of data for indicators has to be sensitive to a variety of different ways in which the collection itself may affect behavior and thereby change the data. The "Hawthorne effect," whereby the mere act of observation may cause people to behave differently, has been understood for some time. In this neutral form, it is not necessarily

[3] See the NSF website for information on the changes in sampling during the 1990s. http://www.nsf.gov/statistics/nsf02312/sectb.htm

assumed that one can predict the nature or direction of the change in behavior that results from observation. A more complex and dangerous problem arises from the combination of data that are imperfect proxies for an underlying behavior, with the use of those data to evaluate and reward the people who generate the data.

For example, it is entirely reasonable to use numbers of published papers and citations as a proxy for the generation of new knowledge. We know that the relationship between the proxy and the underlying phenomenon is imperfect, but as long as the relationship is stable across both time and context, it can be useful indicator. Even if there are differences in the relationship in different times or different contexts, it may still be a useful indicator if we know what those differences are, or can use statistical methods incorporating additional information to correct for them. But if we increase our reliance on such data for the purpose of evaluating and rewarding individuals or organizations, they will then have an incentive to generate more papers *relative to* the rate at which they are generating new knowledge. This can easily generate changes over time in the relationship between the data and the concept of interest, and the extent of these changes may vary across institutional contexts. These effects then make the proxy indicator less informative about the underlying phenomenon of interest.

There is no total solution to this problem, because, as noted above, virtually all indicators are to some degree proxies that may diverge from the underlying concept of interest. But some may be less subject to this kind of endogenous distortion than others. It is also possible that this consideration should weigh in favor of passive data collection rather than surveys, to the extent that people are more likely to respond to the possible incentives created by data collection if they themselves are active participants in its collection.

2

A model of the STI system

In this section we outline a stylized model of the way in which science, technology, and innovation contribute to economic welfare. A science, technology, and innovation (STI) system consists of a number of components linked by the knowledge and resources that flow among them. These components obviously include governments, government research laboratories and extension services, the intellectual property (IP) system, higher education and research institutions, venture capital, and industrial research laboratories. Less obviously, they also include individual actors in any arena that are engaged in improving the efficiency of their production, introducing new ideas and new products, or even making the effort to adopt a new technology or method of organization. When the system works well, the interaction of these institutions and individuals produces welfare-enhancing economic growth via the introduction of new and improved processes, products, and services.

For a comparative study of the operation of different national innovation systems around the world, see Nelson (1993). In that volume, the authors emphasize the importance of interaction among the components of such a system in producing good outcomes. To take a simple example, high quality educational and research institutions are less effective in generating economic growth within a country if firms are not capable of making use of either the graduates or the research output they generate, or these graduates find that employment in secure government research jobs is preferable to working for firms. Another example would be the failure or inability of potential entrepreneurs to put good ideas into practice due to lack of a venture capital system or other form of financing.

The task at hand here is to consider the production of indicators that might help in monitoring such a system and designing policies that enhance its functioning if necessary. Because the innovation system is embedded in the entire economy and touches almost all of it and because it is inherently an intertemporal system, the overall framework used to structure analysis is generally an economic growth model. This model expresses aggregate economic output as a function of ordinary inputs such as capital and labor, and adds to these inputs stocks of intangible knowledge assets, which are assumed to improve the production of output above and beyond the level produced by the other inputs. The use of stocks is intended to capture the idea that knowledge generated in one period is useful in many subsequent periods. These knowledge assets are used to increase the amount of goods and services that can be produced by any given amount of human labor and other inputs. They also may be used to produce desirables such as health, clean environment and national security, which are largely unpriced and may not be adequately included in aggregate output, something we will return to later in our discussion. In addition, to some extent the stock of knowledge may also produce unpriced undesirables. For an example, see Rogoff (2017) on the food manufacturing and marketing industry, where there may have been a negative impact of technical change in that industry on obesity and health, that outweighs the advantages of increased efficiency and lower cost of production.[1]

A stylized version of such a growth model can be written like this:

$$Y \sim C^{\alpha} L^{\beta} K^{\gamma}$$

[1] A related literature questions the implicit assumption that more innovation is always better. E.g., David (2012) and Soete (2012).

or in growth rate form as:

$$g_Y = \alpha g_C + \beta g_L + \gamma g_K + e$$

In this formulation, Y is output, C is ordinary physical capital, L is labor input, K is a measure of knowledge assets, the gs are the corresponding growth rates, and e is any output growth that cannot be explained by these inputs. In various forms, this model has been used for over 50 years to describe and account for the sources of economic growth (Abramovitz, 1956; Solow, 1957; Jorgenson and Griliches, 1967). However, it is important to note that use of this model requires estimates of the weights α, β, etc. For growth accounting purposes, it is assumed that the capital inputs are supplied competitively. The shares can then be estimated as the nominal shares of nominal income received by the various factors, which creates some difficulties detailed below when knowledge capital is included separately. In contrast, most micro-econometric productivity work estimates the weights α, β rather than computing them from income shares, an approach that requires weaker assumptions on the nature of competition. The result is that growth accounting measures the contribution of the knowledge assets using input costs, whereas the econometric approach measures their contribution using output produced.[2]

Recently the U.S. system of national accounts has begun to incorporate one aspect of knowledge stock, the knowledge produced by R&D investment, both public and private, into a so-called R&D "satellite" account (Carson, 1994). This is feasible because R&D data have been collected over a long time period (from 1953) and therefore stocks can be constructed. Fraumeni and Okubo (2005) use the resulting system of national accounts (specifically the National Income and Product Accounts or NIPA) to estimate the contribution of R&D to economic growth over a forty-year period 1961-2000. They measure the income

[2] In practice a number of issues arise due to the lack of firm-level price deflators and the contribution of innovation spending to quality change and market power and therefore pricing. These issues always are there in growth accounting, but become more important when evaluating the role of innovative activity in production, because two of the goals of innovative activity by firms are quality improvement and the creation of market power via product differentiation.

share of R&D by assuming both a rate of return and a depreciation rate, and perform sensitivity analysis by varying these rates. They found that returns to R&D capital accounted for 7 to 11 percent of real GDP growth, with a range of 4 to 15 percent under their most extreme assumptions. In addition, treating R&D as an investment raises the national savings rate by two percentage points from 19 to 21 percent and has a small positive effect on the measurement of GDP growth because the creation of R&D capital is added to GDP, and R&D has grown faster over time than GDP.

Using the growth accounting framework, indicators can be used to measure both the level of important stocks in this system and the rate of important flows. Both stocks and flows can be measured at different points in time, and the exercise performed at different levels of aggregation, that is, for different kinds of organizations such as business or government, different regions, or different industrial sectors. Possible investment flows that might be considered as indicators include the following:

- The resources that are being expended toward the goal of new knowledge creation, which is largely R&D, but may include other expenditures such as training in new processes or design, engineering and marketing expense associated with new products.
- The resources expended on new capital equipment that is associated with the introduction of new processes and methods of organizing production (broadly defined to include the production of services and nonprofit/government outputs).
- The resources expended in educating scientists and engineers.
- The rate at which advanced degrees in S&E are received.
- The rate at which new knowledge is being generated, such as measures of patents applied for or publications.
- The rate at which knowledge is being transferred between or among different regions or organizations, sometimes measured by citations.

- The rate at which new products or new methods are being incorporated into production of goods and services.

Most of these flows can be transformed into stocks, with varying rates of depreciation, and incorporated into the growth accounting framework described earlier. In section 4 of this report, we describe how the currently available indicators might fit into this system.

2.1 Limitations of growth accounting

Growth accounting is a useful schematic for organizing one's thoughts about the sources of economic growth, but it has a number of limitations, especially in the analysis of intangible knowledge and innovation investments. We highlight three: 1) the assumption of normal rates of return; 2) the omission of unpriced output; and 3) that fact it obscures the underlying functioning of the STI system, by focusing only on the input expenditures and the final output in the form of economic growth.

Normal rates of return

The share weights in the growth accounting framework are computed by multiplying an estimate of the annual return to a knowledge asset times a measure of the knowledge stock and dividing by total nominal output. Assuming a stream of investments $R_t, R_{t-1}, \ldots$ and using a declining balance formula with deprecation rate δ to compute the knowledge stock, we have the following equation for the stock:

$$K_t = (1 - \delta_t)\, K_{t-1} + R_t$$

The share weight for its contribution to growth is the following:

$$\gamma_t = \frac{(r_t + \delta_t)\, K_t}{p_t Y_t}$$

where pY is nominal output and r_t is the net rate of return to the knowledge investment. These equations depend heavily on the assumed depreciation rate δ that may vary over time, and on the assumption of a normal rate of return. In the absence of markets for knowledge or innovation assets, we have relatively little information on their values and what information we do have suggests that both depreciation and return vary greatly over time (Hall, 2005). In addition, there is the obvious fact that returns to innovation at the industry or economy-wide level, unlike returns to ordinary tangible assets, are not tied down by the discipline of the market, because of the existence of knowledge spillovers.

Unpriced outputs

A second limitation of growth accounting is related to the fundamental question of what we should be measuring as output: GDP or some other concept such as Net Economic Welfare (NEW) or Gross National Happiness (GNH). In the case of innovation, the problem of output measurement centers on the lack of good price data for a set of innovation outputs that include improvements in health from investment in the healthcare sector, national defense, and improvements in environmental quality.

With respect to healthcare, the US Bureau of Economic Analysis is currently engaged in a project to develop a satellite account for health spending that involves the following tasks: 1) creating a set of health expenditure statistics that is common to BEA and the Center for Medicare and Medicaid Services (CMS), and 2) producing a comprehensive set of health care sector accounts for health care income, expenditure, and product.[3] In turn, this involves developing medical care price and real output measures that break out the delivery of health care from increases in the price of that care, rather than simply capture output of the sector via input spending. To the extent that this project is successful, it should improve our ability to track the contribution of spending on health-related innovation to improvements in healthcare.

Popp *et al.* (2010) review the economic research on the impact of innovation in the environmental area, including energy. Much of this research focuses on measurable output, such as pollution levels, or on the inducement of innovative and R&D activity in this area via legislative standards. Another growing area of research is dynamic modeling of

[3] For further information about this effort, see http://www.bea.gov/national/health_care_satellite_account.htm (2012).

the implications of R&D directed at climate change for the economy (e.g., Acemoglu *et al.*, 2012; David *et al.*, 2011). None of this work provides much guidance for translating the intermediate outputs such as reductions in pollution or reduced risk from global warming into measures of economic welfare. But it does suggest that in some areas of innovation, the measurement of direct effects like the reduction in effluents, air and water quality, and so forth, may be relevant.

Evaluating the contribution of innovation to national defense and some kinds of environmental improvement or avoidance of environmental damage is considerably more challenging. This is an area where tracking spending seems the best we can do. In the case of national defense, a long history of analysis in this area points to important spillovers to non-military applications from defense-related R&D investment (Mowery, 2010). Tracking these effects would involve tracking military R&D spending by area, which is already done to some extent. Additional data that might be useful would be entry, growth, and non-military sales shares for firms that contract with the Defense Department to develop new technologies, products, and services. Some evidence of this kinds exists for small firms via evaluation of the SBIR set-asides (Wessner, 2009).

Inside the STI system

The growth model identifies the relevant inputs and outputs, but leaves the mechanisms that connect them unspecified. In the words of Rosenberg (1982), it is a "black box." The inputs are expenditures to increase knowledge stocks and related capital equipment, as well as the quality and education of the labor force. The outputs are knowledge stocks in the subsequent period and output growth. But a number of intermediate inputs and outputs are clearly worth tracking. First, there is the human capital embodied in the labor force, which is partly the creation of the education and job training systems in the economy. Second, the knowledge stocks created by innovation investments are used to produce new and improved products and processes, which in turn lead to the growth of output. Third, the various actors engaged in innovation and knowledge creation interact in ways that add to output and we would like to measure this also, in order to understand the mechanisms that encourage such interaction.

Thus ideally we would like to have indicators for inputs, intermediates, and outputs. To put it another way, we are interested in indicators of innovative activity, innovation, and the impact of innovation on economic welfare. As a general rule, inputs are easier to measure than outputs, mostly because some entity is paying for them so we have a way to aggregate them (using dollar expenditure). On the other hand, the intermediate and final outputs of innovation are often either untraded, or is qualitative, so it is more difficult to measure and to aggregate. That is one reason why we sometimes rely on productivity analysis, which allows the use of national income measures such as GDP, or sales in the case of firms, adjusting these measures for the level of other inputs.

A final observation is that once one goes within the innovation system (unpacks the black box), it quickly becomes clear that there is feedback from later stages to earlier. That is, the so-called "linear model" of innovation is not quite an accurate picture. Various authors, led by Rosenberg (1982), have emphasized this feature of the process.[4] Nevertheless, from a policy perspective, it is natural to think of the model in this way, because it is primarily the inputs (in the form of spending on innovation, basic research, higher education, and the like) that are subject to policy intervention. In addition, it cannot be emphasized enough that innovative output is subject to great deal of uncertainty that is beyond the control of policy or planners. The implication is that although output or success indicators are of interest, they will not turn out to be much of a guide for detailed planning.

[4] See also Kline and Rosenberg (1986). Balconi *et al.* (2010) present a thoughtful defense of the linear model. Among other things, they point out that in many cases basic science has served as an essential basis for subsequent innovations.

3

Overview of existing U.S. indicators

The state of the existing indicators for the U.S. economy can be seen in the biannual volume published by NSF, entitled Science and Engineering Indicators. Table 3.1 summarizes the indicators available in the latest volume of this publication, dated 2012, and covering data through 2010 in most cases. Although the table suggests an apparent wealth of innovation measures, many of the numbers in it are drawn from administrative records collected for another purpose, and may not be ideal for the measurement of innovative activity. Nevertheless, these statistics provide a very useful starting point. In addition, many of the measures we do have, such as various types of R&D spending and patents granted, are specialized towards the manufacturing sector and technological innovation. Thus they provide valuable, but incomplete coverage of the full range of innovative activities. Other measures, such as value added in various sectors, are bottom line measures influenced by many other factors in addition to innovative activity.

Table 3.1: NSF Innovation Indicators.

Broad area	*Area*	*Indicators available*	*Figure 1 Reference*	*Comment*
Human capital	Elementary and secondary math and science education	Various scores, including PISA; highest skill area;	$2 \rightarrow 4$	Quality of human capital creation
		% taught by BA in math or science;	2	Quality of teaching
		% passing APT and % going on to college;	$2 \rightarrow 4$	Quality of human capital creation
		Salaries and qualifications of math & science teachers;	$1 \rightarrow 2$; 2	Funding of education; quality of teaching
		State standards in math & science	2	Quality of teaching
	Higher education in science and engineering	S&E degrees awarded by type and field;	$2 \rightarrow 4$	Magnitude of human capital creation
		Enrollment in higher education by field and nationality;	$2 \rightarrow 4$	Magnitude of human capital creation
		Degrees granted by private for-profit institutions;		
		Field of major intentions;	$2 \rightarrow 4$	Magnitude of human capital creation
		Sources of support and amount of debt;	$2 \rightarrow 4$	Composition of human capital creation
		Foreign enrollment at all levels;	$1 \rightarrow 2$; 2	Financing of human capital creation
		Postdocs by field and nationality;	2	Composition of human capital creation
		Spending on tertiary education	$2 \rightarrow 4$	Magnitude of human capital creation
			$1 \rightarrow 2$; 2	Financing of human capital creation
	S&E labor force	Occupation of S&E degree holders;	$4 \rightarrow 7$; 7	Deployment of human capital in industry
		BLS projections of occupational employment 10 yrs out;	Exp. $4 \rightarrow 7$	Deployment of human capital in industry
		S&E in work-related training;	$7 \rightarrow 4$	Human capital creation within industry
		Patenting by S&Es by field and industry sector;	$7 \rightarrow 3$	Knowledge creation in industry
		Federal S&E employment;	1; 2	Deployment of human capital in government
		US residents with foreign highest degrees	4	Composition of human capital

Table 3.1: *continued*

Broad area	*Area*	*Indicators available*	*Figure 1 Reference*	*Comment*
		US R&D spending by sector, source of funds, state;	$1 \rightarrow 2, 5, 7; 7 \rightarrow 5,6$	Funding of knowledge creation
		Basic, applied, and dev. R&D by sector, source of funds;	ditto	Composition of knowledge creation
		Federal R&D budget authority or obligations by function & character of work, agency, location;	$1 \rightarrow 2, 5, 7$	Federal funding of knowledge creation
		Federal R&E tax credit claims by NAICS;		
R&D spending	R&D spending	R&D spending by foreign affiliates in US by NAICS & country;	$1 \rightarrow 7$	Federal subsidy of industry research
		R&D spending abroad by US firms by NAICS & region;	7	Source of industry funding of knowledge creation
		Contract R&D in US by NAICS;	7	Location of industry-funded knowledge creation
		US trade in R&D and testing services;	7	Composition of industry funding of knowledge creation
		Tech alliances by type, technology, country	7	Location of industry-funded knowledge creation
		SBIR and STTR awards by type & agency;	5; 7	Nature of partnerships
		ATP projects, TIP, Mfg extension information	$1 \rightarrow 7$	Federal funding of industry knowledge creations

Table 3.1: *continued*

Broad area	*Area*	*Indicators available*	*Figure 1 Reference*	*Comment*
R&D spending	Academic R&D	Academic R&D for basic, applied, development;	$2 \rightarrow 3$	Composition of knowledge creation
		Support for academic R&D by sector;	$1 \rightarrow 2$; $7 \rightarrow 6$	Funding of academic knowledge creation
		Federal support to acad. R&D by agency & character of work, field;	$1 \rightarrow 2$; $1 \rightarrow 5$	Federal funding of academic knowledge creation
		Academic R&D by field and source;	$1 \rightarrow 2$; $7 \rightarrow 6$	Funding of academic knowledge creation
		Faculty in S&E by field;	2	Composition of academic sector
		S&E doctorates in academia by field, type of work, nationality, type of support, collaboration	$2 \rightarrow 4$	Magnitude of university human capital formation
		S&E publication data by field and region;	$2 \rightarrow 4$; $7 \rightarrow 4$	Knowledge creation
		Internationally co-authored pubs by selected country pairs;	$2 \rightarrow 4$; $7 \rightarrow 4$	International composition of knowledge formation
		US utility patent awards by owner type;	$7 \rightarrow 3$; $2 \rightarrow 3$	Knowledge creation
		Academic patent licensing;	$2 \rightarrow 3 \rightarrow 7$	Academic-source knowledge used in industry
		Citations of papers in US patents by sector and field	$3 \rightarrow 7$	Use of knowledge in industry

Table 3.1: *continued*

Broad area	*Area*	*Indicators available*	*Figure 1 Reference*	*Comment*
Commercialization	Industry, technology & globalization	VA of K-intensive, ICT, high-tech & other industries by country;	$7 \rightarrow 8$	Commercialization of technology
		VA of education, health, business, financial services by country;		
		Exports & imports of high-tech goods by country and broad sector;	$7 \rightarrow 8$	Commercialization of technology
		US trade in high-tech goods by country and sector;	$7 \rightarrow 8$	Commercialization of technology
		Domestic & foreign VA and employment by US MNEs by sector;	$7 \rightarrow 8$	International spillovers
		US FDI abroad and foreign FDI in US, by industry	3	International spillovers
		US receipts & royalty payments from foreign affiliates;	$7 \rightarrow 3$	Knowledge creation by industry
		US patent apps and grants by country, technology;	$7 \rightarrow 8$	Innovation
		US trademark apps and grants by country, technology;	$7 \rightarrow 3$	Knowledge creation by industry
		Triadic patent families by country;	$9 \rightarrow 7$	
		US VC investment information		

Table 3.1: *continued*

Broad area	*Area*	*Indicators available*	*Figure 1 Reference*	*Comment*
Attitudes	Public attitudes to S&E	Primary sources of information about news & science;	n.a	Environment for innovation
		Public interest in selected issues;	n.a	
		Visits to science & cultural institutions;	n.a	
		Correct answers to various science questions;	n.a	
		Assessment of govt spending on R&D and quality of science education in US;	n.a	
		Attitudes towards nanotechnology, stem cell, animal research, by country;	n.a	
Regional	State-level information	Average math & science scores;	$2 \rightarrow 4$	Quality of human capital formation
		S&E doctorate holders, employed and otherwise;	4	Human capital stock
		See above for R&D information		

Source: US NSF STI Indicators (2010) with authors' additions to columns 4 and 5.

Figure 3.1 presents a schematic summary of how resource inputs, knowledge outputs, and ultimate impacts flow in the economy, providing a framework for assessing which aspects of the system are captured well by existing indicators and which are captured less well. The blue rectangles represent the main actors in the knowledge economy-government, the education sector including public research organizations, various industries, and the finance sector. The red hexagrams represent the two main classes of research and innovation collaboration, between industry and university, and between public (governmental) and private organizations. The yellow ovals are the two intangible assets of the knowledge economy, which are difficult to measure: the knowledge base (largely codified and/or not embodied in individuals) and human capital (largely tacit and/or embodied in individuals). The green arrows represent flows of resources (funding) from one sector to another, and the dotted arrows represent intermediate knowledge outputs that flow between and among the sectors. The heavy dark arrow at the bottom indicates that the for-profit sector, using both internally and externally generated knowledge, generates commercial innovation, which is then manifested in output growth, productivity improvement, and enhancement of individual welfare.

The Figure is designed to capture, in a stylized way, the important investment mechanisms and the important impact pathways of the endogenous growth model described in the previous section. We see, for example, that both the education sector and the industry sector contribute to the creation of new knowledge and new human capital, and these knowledge stocks in turn feed back into both the education and industry sectors to foster cumulative knowledge growth. Within the industry sector, we have shown two (of many) industries, to represent that knowledge spillovers flow between industries, as well as between firms in a given industry (illustrated within the right-hand industry in the diagram). Also within industries lie R&D labs, which are largely funded from within industry, but also receive public funds. These labs are a source (but not the only source) of knowledge creation within the industrial sectors. Finally, the government is a major funder of all of the knowledge creation sectors.

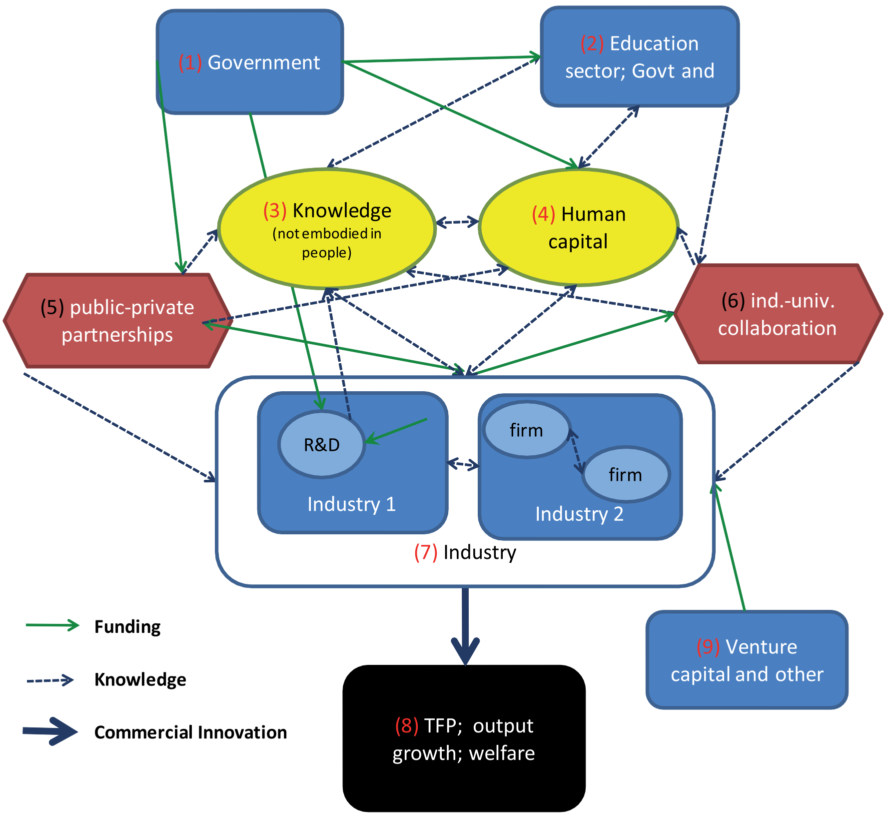

Figure 3.1: Schematic Overview of the STI System.

The last two columns of Table 3.1 show the connections between the STI Indicators currently published by NSF and the stocks and flows in Figure 3.1, using the numbers that appear in each polygon in Figure 3.1 to make the connections So, for example, in the first row of the table, an indicator related to elementary and secondary test scores is labeled " $2 \to 4$," indicating that it corresponds to the arrow connecting the rectangle (2) to the oval (4) in Figure 3.1.

Perusal of Table 3.1 and Figure 3.1 together indicates that our existing indicators are quite complete in their coverage of the resource flows that support the generation of new knowledge, containing considerable information about the magnitude and distribution of those flows. There are also multiple indicators that correspond to knowledge and human capital outputs, although these measures are universally proxies that are related to the underlying concepts with substantial measurement error (e.g., degrees as a measure for the human capital of graduates; papers as a measure of new scientific knowledge; patents as a measure of new technical knowledge). The measures of innovation (as opposed to innovative activities) are much less complete, and arguably more "distant" from the underlying concepts. For example, the errors of both over-inclusion and under-inclusion in using new trademark registrations as a proxy for innovation are probably even greater than the corresponding errors in using patents as a proxy for new inventions. And a similar observation applies to using high-tech value added as an indicator for increased output that can be attributed to innovation.

We will return to a discussion of the potential issues and gaps in the existing NSF Indicators, after a discussion in the next section of analogous data collection and publication efforts in other countries.

4

International context

Almost every developed and mid-level developing country in the world has been concerned to some extent with the problem of measuring innovation. Broadly speaking, most have arrived at some combination of R&D spending, tertiary education, patenting activities, publication activities, research funding availability, and success in exporting and growth. In addition, many countries have made use of the innovation surveys pioneered in Europe to assess the innovative activities of their own firms. See Table 4.1 for a list of the countries that have fielded one or more innovation surveys.

Two aspects of these international efforts are important for the U.S. innovation indicator enterprise. First, as a latecomer in some areas of measurement, we can learn from the experience of others. Second, the fact that many countries are constructing "innovation scoreboards" means that there is a demand for international comparability of at least some of these measures, in order to facilitate various kinds of benchmarking. However, exact comparability is difficult to achieve due to differences in the ways questions are interpreted, the structure of industrial firms, and differences in the cultural and institutional contexts.

Table 4.1: Innovation Surveys.

Region	*Countries*	*Website*
EU27	Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg Malta, Netherlands, Poland, Portugal, Romania, Slovenia, Slovakia, Spain, Sweden, UK	http://www.proinno-europe.eu/metrics
Rest of Europe	Norway, Switzerland	http://www.ssb.no/innov_en/ http://www.kof.ethz.ch/en/surveys/structural-surveys/
Latin America	Argentina, Brazil, Chile, Colombia, Cuba, Ecuador, Mexico, Panama, Peru, Trinidad and Tobago, Uruguay	http://www.ricyt.org
Asia	Japan, Korea, Malaysia	http://www.nistep.go.jp/en http://www.kistep.re.kr/en/c3/sub2.jsp
Pacific	Australia, New Zealand	
Other	Canada, South Africa	http://www.statcan.gc.ca/concepts/index-eng.htm

Both DIW and NESTA (discussed below) have introduced several measures that might be used to capture some of the cultural differences across countries. These measures include some that are already collected by the NSF, such as attitudes to the development of technology despite its potential risks, and evidence on the public's interest in new inventions and technologies. Additional measures that might be useful are those for buyer sophistication with respect to innovation, the level of computer and internet skills, social capital and trust. A particularly salient measure for the development of entrepreneurship is the measure of the fear of failure, which is frequently identified as an important reason for lower entrepreneurship rates in Europe, but might also be useful for comparing regions within the United States.

Table 4.2 below gives an idea of the innovation indicators that are being measured by the European Union. Some of them reflect particularities of the European institutional setting. A couple of examples: 1) The focus on measuring climate change mitigation innovation is due to a strong policy stance of the EU with respect to green technology.[1] 2) The age restriction on doctoral graduates is probably because it is extremely unusual in Europe for anyone to go back to school for a doctoral degree at older ages, unlike the situation in the U.S. However, in general similar indicators can be or already have been constructed for the United States, with the notable exception of those based on the innovation survey questions, at least to date. In the following subsection we discuss what has been learned from others' experience with these surveys, and what the U.S. is doing in this area.

[1] E.g., see the Europe 2020 agenda for growth. http://ec.europa.eu/europe2020/index_en.htm

Table 4.2: European Union Innovation Indicators.

Input or output	*Broad area*	*Indicator*	*Available in US?*
Input indicators (human capital)	Human capital	New doctorate grads per 1000 pop aged 25-34	New PhDs
		Share of pop aged 30-34 having completed tertiary educations	Yes
		Share of youth aged 20-24 having attained at least upper secondary level education	Yes?
	Research system	Intl. scientific co-publications per million pop.	In principle
		Scientific publications among the top 10% most cited publications worldwide as % of total scientific publications of the country	In principle
		Non-EU doctorate students per million pop.	Non-US
Input indicators (industry)	Finance	Public R&D spending as share of GDP	Yes
		VC (early stage, expansion and replacement) as share of GDP	?
	Firm investment	Business R&D spending as a share of GDP	Yes
		Non-R&D innovation spending as a share of turnover	No
	Linages and Entrepreneurship	SMEs innovating in-house as a share of SMEs	Not at present
		Innovative SMEs collaborating with others as a share of SMEs	No
		Public-private co-publications per million pop.	In principle

Table 4.2: *continued*

Input or output	*Broad area*	*Indicator*	*Available in US?*
Intermediate output indicators	Knowledge assets	PCT patent applications per billion GDP	Yes
		PCT patent applications related to societal challenges per billion GDP (climate change mitigation; health)	In principle
		Community trademarks per billion GDP	US trademarks
		Community designs per billion GDP	US trademarks
Output indicators	Innovators	SMEs (>10 employees) introducing product or process innovations as share of SMEs	Not at present
		SMEs (>10 employees) introducing marketing or organizational innovations as share of SMEs	Not at present
		High-growth enterprises as share of all enterprises (???)	??
	Economic effects	Employment in knowledge-intensive activities (manufacturing & services) as a share of total employment	Yes
		Medium and high-tech manufacturing exports as a share of total product exports	Yes
		Knowledge-intensive services exports as a share of total service exports	possibly
		Sales of new to market and new to firm innovations as a share of turnover	Not at present
		License and patent revenues from abroad as a share of GDP	Yes

Source: European Union (2012, 2012) and authors additions.

4.1 Innovation surveys

The challenge of measuring innovation broadly has been met at least partially during the past 20 or so years by the introduction of direct surveys of business firms in many countries. Although earlier pioneering surveys of various kinds exist, this activity really took off with the publication of the first edition of the Oslo Manual in 1992 (third edition, Tanaka *et al.*, 2005), which had guidelines for the definition of various kinds of innovation and for the collection of innovation-related data. For example, several non-R&D kinds of innovative expenditure were identified in the manual: the later phases of development and testing that are not included in R&D, capital expenditures related to the introduction of new processes, marketing expenditures related to new products, certain kinds of employee training, expenditures on design and technical specifications, etc.

The Oslo Manual defines innovation as follows:

> "An innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations."[2]

In spite of the apparent clarity of this definition, measuring innovation in a form that is useful for statistical analysis has proved challenging. The central problem is that no two innovations are alike. Some innovations (e. g., the invention of the telephone or perhaps the telegraph) create a whole new market sector whereas others are useful but trivial, and there is a wide range in between. In general we can say that smaller innovations are more numerous than game-changing ones. As an illustration of this phenomenon, consider the data of Acs and Audretsch (1990), who collected a comprehensive list of innovations introduced by business firms during the year 1982. Over 85 per cent of the innovations they identified were modest improvements to existing products, and none created entire new markets. Fewer than 2 per cent

[2] *Oslo Manual* (Tanaka *et al.*, 2005), third edition, p. 46.

were considered even the first of its type on the market in existing market categories.[3]

The NSF has recently redesigned the industrial R&D survey and renamed it the Business R&D and Innovation Survey (US-NSF, 2009). In doing the redesign, they added a pilot question on innovation similar to those in the European CIS (Community Innovation Survey) in 2008:

> Did your company introduce any of the following during the three-year period, 2006 to 2008?
>
> 1. New or significantly improved goods (excluding the simple resale of new goods purchased from others and changes of a solely aesthetic nature).
> 2. New or significantly improved services.
> 3. New or significantly improved methods of manufacturing or producing goods or services.
> 4. New or significantly improved logistics, delivery, or distribution methods for your inputs, goods, or services.
> 5. New or significantly improved support activities for your processes, such as maintenance systems or operations for purchasing, accounting, or computing.

Unfortunately, for a number of reasons, including survey design, coverage, and nonresponse, the results from the pilot question cannot really be compared with those from other countries (Hall, 2011; Jankowski, 2012). In response to the problems they encountered with this first trial, the NSF redesigned the survey in the following year, moving the innovation question to the front of the survey and adding several other questions, but the data for 2009 and 2010 have not yet been released so that we are unable to use them in this report. Preliminary numbers in Jankowski (forthcoming, 2013) suggest that the U.S. innovation rates

[3] Note that by using the 1982 date, Acs and Audretsch did miss two major innovations: the IBM personal computer and Microsoft DOS, both of which were introduced in 1981 and which arguably meet the definition of "created entire new market".

are still substantially lower than those for European countries, but issues of firm size and sector coverage comparability still remain.

The innovation data produced by the European CIS have been widely used by economists in models of R&D, innovation, and productivity, so we have learned something about their quality (Mairesse and Mohnen, 2010). The main variables used have been the binary indicators for product and process innovation at the firm level. Although they are clearly related to the conduct of R&D in the data (especially in the manufacturing sector), they are also very noisy indicators of the underlying innovation concept.[4] Given a choice between the amount of R&D spending, or a binary innovation indicator to predict productivity growth at the firm level, the estimation models clearly prefer R&D spending as an explanatory variable, when it is available. However, since many innovating firms do not report that they do any R&D, the innovation indicator can still be useful.

A simple thought experiment will demonstrate why a binary innovation indicator has limitations: imagine comparing the answers given by a micro enterprise and a large multinational to the question of whether they have introduced a new product or process during the past three years. Clearly, if the large multinational answered no, we would be worried about its survival, whereas it might be perfectly normal for a small enterprise to innovate, but possibly not at less than three year intervals. But have we learned that the small enterprise is less innovative than the very large one? Not necessarily. In contrast, R&D is a continuous variable that can be transformed into the intensity with which a firm invests in technological innovation.

Alternatively, consider comparing two larger enterprises, both of which introduce a new product, but one of them captures the market, and the second enters a market new to it and fails to get more than small share, while remaining profitable due to its old products. Both will be recorded as innovative, but the result is clearly quite different. The amount they earn from innovation as a share of sales would be a lot more

[4] Hall *et al.* (2013) report that in the UK, the share of service sector innovating firms that do not do any type of R&D (including external, capital equipment for new process, and training) is about twice that in the manufacturing sector.

informative. In the case of product innovation, fortunately the share of sales due to new products has been collected by the CIS (and will be collected in future U.S. surveys) and this has turned out to be a better predictor of productivity than the simple binary variable, especially in the absence of a measure of spending on innovation. It would be desirable to try to collect a similar measure for process innovation.[5] If one cannot obtain estimated cost reduction from this kind of innovation, it might be possible to use measures of investments in new capital and training that are associated with the introduction of new processes.

All this suggests a couple of ways that innovation surveys could be improved to capture a full picture of firm innovative activity. Some of these have already been discussed in the Oslo Manual, but have not been incorporated into most innovation surveys. The first would be a focus on trying to measure the benefits of process or organizational innovation in a more quantitative way. For this purpose, the Oslo Manual suggests first asking whether the innovation led to a reduction in cost, and then asking by how much:

> "These questions can either be asked with respect to average costs or to specific costs, for example changes in the cost of material, energy or labour inputs. Quantitative questions can either ask for an interval estimate of the percentage change in costs, or ask enterprises to choose from a set of predefined categories (e.g. an increase or decrease of less than 5%, 5% to 25%, over 25%). Experience from earlier surveys indicates that enterprises find the latter method easier to answer and thus results in much higher item response rates. The same techniques can also be used to ask about the effect of process innovations on employment, i.e. whether employment increased or decreased, and by how much." (Oslo Manual, third edition, page 111).

A second area where more information than currently available would be helpful is expansion of the information collected on innovation expenditures (question 5.2 on the standard CIS questionnaire). Currently

[5] Peters (2006) reports on such an effort in Germany.

some but not all of the innovation surveys in Europe collect information on the following categories of expenditure:

- In-house R&D - including software development in-house
- External R&D - performed by other enterprises (including other enterprises or subsidiaries within your group) or by public or private research organisations and purchased by your enterprise
- Acquisition of machinery, equipment and software used to produce new or significantly improved products and processes
- Acquisition of external knowledge - Purchase or licensing of patents and non-patented inventions, know-how, and other types of knowledge from other enterprises or organisations for the development of new or significantly improved products and processes
- Training for innovative activities - Internal or external training specifically for the development and/or introduction of new or significantly improved products and processes
- Market introduction of innovations, including market research and launch advertising
- Design - to improve or change the shape or appearance of new or significantly improved goods or services
- Other - Other activities to implement new or significantly improved products and processes such as feasibility studies, testing, routine software development, tooling up, industrial engineering, etc.

In some cases only a yes/no answer is required for the above in the current surveys, but the data would be considerably more useful if actual expenditure data can be collected.

A final suggestion for additional data that might help correct the bias introduced by sectoral variation in the meaning of a new product also comes from the Oslo Manual:

> "In order to take into account the effects of product life on this indicator [the introduction of a new good or service in the past three years], it is suggested that the firm should be asked to give an estimate of the average length of its products' life cycles. This information could be used to weight the percentage shares suggested above. An alternative way of putting this question is to ask how often the firm usually introduces innovations." (Oslo Manual, third edition, page 110).

We have not found any evidence that this question has ever been attempted in an actual innovation survey, so it is hard to evaluate whether the answers would be helpful. However it does see worth considering as an indicator that would help interpretation of the innovation indicators themselves.

4.2 Other innovation indicator efforts

In this section we catalogue a few other efforts toward constructing innovation indicators, by the World Bank, the UK government, and a research institute in Germany, the Deutsche Institut für Wirtschaftsforschung (DIW).

World Bank

KAM is the Knowledge Assessment Methodology developed by the World Bank, which consists of 4 pillars:[6]

- Economic Incentive and Institutional Regime
- Education
- Innovation
- Information and Communications Technologies

[6] http://www.worldbank.org/kam

From these pillars, each of which contains a number of individual measures, they produce two indices: a Knowledge Index (KI) and a Knowledge Economy Index (KEI). The Knowledge Index is based on education, innovation and ICT investment. The Knowledge Economy Index adds the economic incentive and institutional regime pillar to this.

As is clear from the website, these measures are primarily input and institutional measures, describing the framework in which the knowledge economy can grow, and intended as a diagnostic and benchmarking tool for the World Bank client countries. The website provides a range of different "scorecards," including custom combinations of indicators, that can be consulted by policymakers.

UK NESTA (for the UK Dept of Business, Innovation, and Skills)

A study of wider frameworks for innovation (Allman *et al.*, 2011). This study focuses to some extent on indicators that are designed to measure UK underperformance in innovation environment such as fear of failure, entrepreneurial activity, ICT take-up, etc. They divide indicators into five categories and produce a long list of indicators for each one:

- Public research base and linkage to industry
- Demand conditions and competition
- Supply of high quality human resources and finance
- Infrastructure and services in the economy
- Degree of entrepreneurship

Among the detailed lists that they give, these indicators are already present in most collections:

- Gross Expenditure on Research and Development
- Share of GBAORD as per cent of total general government expenditure

- Higher Education R&D as percentage of GDP
- Percentage of BERD financed by abroad
- Percentage of BERD financed by the government
- Percentage of HERD financed by the private sector
- Publications per 100,000 population
- Publications per million dollars
- Average annual citations per HERD Expenditure
- Patents, Invention Disclosures and Licenses
- Sources of Universities Research Income
- Firms with new-to-market product innovations by size
- Patents and trademarks per capita, 2005-07
- Number of scientific & technical articles cited in patents
- Citations to academic patents
- Market Capitalisation of listed companies (percentage of GDP)
- Venture capital investment as a percentage of GDP by early stage and expansion
- Share of employment in KIBS, 2009 54
- Education expenditure and performance
- Human resources in science and technology as a share of labour force (%)
- Costs of firm entry

The indicators below are potentially new. We have grouped them according to the areas about which they provide information.

End-user demand for innovation (consumers):

- Interest in new inventions and technologies
- Consumer Confidence Index
- Buyer sophistication: ability of buyers to understand innovation and use it
- Final consumption expenditure of households: communication share

End-user demand for innovation (firms):

- Government procurement of advanced technology products
- Value of public procurement which is openly advertised as a % of GDP
- Firm-level technology absorption
- Cooperation with clients
- Intensity of local competition

Startups and firm growth:

- Births, Deaths and Active Stock indicator
- Employer enterprise birth and death rates
- Fear of failure rate
- Early-stage entrepreneurial activity distinguishing necessity and opportunity
- High-growth early stage entrepreneurial activity
- Global market penetration by SMEs
- Cost of access to IPR services

Financing:

- Ease of access to loans
- Venture capital availability
- Ease of access to local equity markets

Labor markets:

- Employment in creative sectors as a share of employment
- Individuals' level of computer and internet skills
- Life-long learning
- Workforce development indicators (firm-level OJT, etc.)
- Choice, discretion and creativity at work (survey)
- Share of firms with co-operation agreements with government or higher education

Diffusion of new technologies

- Business use of mobile internet
- 3G (and higher) coverage
- 3G or higher cellular mobile adoption
- e-Intensity Index (Boston Consulting Group – weighted sum of a nation's supply of Internet infrastructure and the demand for Internet services)
- Business use of social networking

The above lists are very comprehensive and it is not at all clear that it is even feasible to collect many of the measures, or that all of them would add information to what we already know. However, some of them have the potential to be useful, especially those focused on demand side considerations and diffusion, neither of which are especially well-measured in our current set of indicators.

Deutsche Institut für Wirtschaftsforschung (DIW)

Belitz *et al.* (2011) describe an effort by the DIW in which they computed a composite indicator for 13 EU countries, Japan, Korea, the US, and Canada and do some sensitivity analysis. Unfortunately they are not very specific about which measures they use and how exactly the variables are measured. The subindicator components are the following:

- Education
  - Costs
  - Output in terms of grads
  - PISA scores
  - Lifelong learning
- R&D
  - Spending
  - Patents
  - Publications & citations
- Financing
  - General financing conditions
  - Financing of startups
  - Public R&D support
  - Tax policy
- Demand
  - GDP per capita
  - Domestic demand for innovative products
  - Measures of buyer's sophistication, firm level technology absorption and government procurement of advanced technology products

- Networking
  - degree of inter-company networking
  - alliances with suppliers and customers
  - R&D cooperation, esp. internationally
- Implementation
  - Research intensive VA and employment, balance of payments
  - Valuations of Transport and Energy systems, Network Readiness and E-Readiness
- Competition
  - Product Market Regulation (PMR) index of the OECD
  - Measures of competition and corruption fighting
- Societal innovation climate
  - World Value Survey Data for openness to new technologies and formation of social capital
  - Eurobarometer Data for trust measures and concerns about science and technology.

5

STI indicators for STI policy

As noted above, indicators are desired to some extent merely for the purpose of a scorecard or benchmarking of the level of activity in different areas. And it is important for the advance of social science that the data underlying STI indicators be available. But consistent with the idea of "The Science of Science and Innovation Policy," it is also desirable for the indicators to be as helpful as possible in informing major public policy decisions. In this section, we survey briefly how indicators can (and often cannot) meaningfully inform policy choices.

Overall level of public investment in R&D. Implicitly, the Congress and the President are continuously deciding what overall level of resources to invest in new knowledge creation through the R&D process. Ideally, this would be informed by data showing the marginal rate of return on these investments. But marginal rates of return are very difficult to measure. Economists and others have made estimates of the average rate of return to R&D investments (Hall *et al.*, 2010). Within the model, the marginal rate of return declines with the intensity of R&D investment (R&D/GDP) other things equal, so a high average rate of return is a necessary but not sufficient condition to justify increased investment.

In the absence of explicit information, R&D intensity measures do provide some implicit evidence on the rate of return. Economic models typically presume that there are diminishing returns to increased R&D expenditure, so that the rate of return to R&D will fall as R&D/GDP rises. This means that if today's U.S. R&D/GDP ratio is lower than at another point in time, we may be able to infer that the rate of return in the U.S. today is higher than it was at that point of time, assuming that nothing else has changed. The same argument applies when comparing R&D intensities across countries, although it is even more difficult to assume that other things are equal in that case. Thus if we have some reason to believe that the investment level was right at some point in time, then we might be able to infer that the implied high rate of return in the U.S. today justifies a higher level of investment (and vice versa if today's U.S R&D intensity is higher than at some other time or place). However, given all the uncertainties, it would probably be better to attempt to measure the return to R&D spending in this case.

Overall level of public investment in education and training. The issues with respect to the optimal level of investment in education and training are analogous to those related to R&D. We would, ideally, like to have measures of the rate of return; measures of the current ratio of investment to GDP may provide indirect evidence on the rate of return, at least relative to other times or places. In addition, public policy may view having an educated public as a desirable end in itself, over and above any return it may provide in terms of innovation and technology. If so, then data on years of schooling and degrees awarded are useful policy indicators independent of their indirect implications for the economic rate of return.

Education and training also take many forms and occur in many different contexts. We have better data on what occurs in formal educational institutions than we have on training that occurs on the job, or is otherwise provided by firms without recourse to formal educational institutions.

Allocation of both of above by scientific/technical area or area of ultimate application. Even more than the overall determination of public investment, the government must continuously decide the allocation of public resources for R&D and education/training across scientific and technical fields, and across areas of application. Again, within the model the most relevant information for these decisions would be the marginal rates of return. And again, these are hard to measure, and measurements of average rates of return are incomplete as indicators of marginal rates. In addition, there are substantial spillovers across scientific fields (e.g., the importance of computer science for DNA analysis) so that localized rates of return may not capture the true importance of some fields.

The relevance of investment intensity measures as indirect indications of marginal rates of return is more complex in the context of allocation across fields or sectors. If the inherent technological opportunity is greater in a given sector, then its marginal returns are higher at any given level of investment. Thus it is possible, for example, that our much higher level of public investment in research in health sciences than in other fields represents an implicit belief that technological opportunity, and hence marginal returns, are higher in that area than in others. On the other hand, no other country in the world devotes such a large share of its public research investment to health sciences. Unless the variation of technological opportunity across fields is different in different countries, comparative benchmarking on sectoral allocations may provide indirect evidence on rates of return. As noted above, however, this is a particularly problematic sector due to the difficulty of measuring output properly and the fact that health improvements are not completely capture by national income accounts.

Allocation of federal R&D and training resources by types of institutions (e.g. intramural versus extramural or universities versus firms). Allocation of public resources across different kinds of institutions raises the same issue of relative rates of return as allocation across sectors. In addition, different kinds of institutions play different roles in the STI system. Hence indicators reflecting intermediate outputs of the research process, and flows of knowledge within the system, might be informative about imbalances within the system. It would also be useful to construct and publicize more detailed statistics on the demand for S&T skills in certain areas, including starting salaries, in a timely manner.

Science and technology policy choices other than spending. Many government policy choices explicitly or implicitly affect the STI system, including R&D subsidies (and other tax policies), intellectual property rules, and mechanisms for the transmittal of funds (e.g. basic research grants, contract research, prizes, etc.). It is not clear that indicators, as we normally think of them, shed light on the relative efficacy of different policy choices of this kind. But the data collected as the basis for indicators can also be used by social scientists to study the relative effectiveness of different mechanisms. In fact, these data are essential for this purpose.

Immigration policy (as applied to scientific/technical workers). Indicators related to the number of number of scientific and technical workers, combined with the level of investment in research, may be useful for informing the nature and extent of visa programs to allow more technically trained immigrants to work in the U.S.

Indicators for use by university administrators or firm managers? Firm managers and university administrators face many of the same choices as governments: how much to spend and what to spend it on. Many of them rely to some extent on benchmarking, that is, observing the spending patterns of their immediate competitors. Therefore the same kinds of data as described above can be useful, preferably broken down by sector and by geography.

6

Issues and Gaps in existing U. S. indicators

In this section we identify a few broad areas where consideration of the STI framework relative to the existing indicators, and learning from data collection efforts in other countries, suggest areas for possible expansion and improvement of U.S. indicators. In the Concluding Section, we move to a broader examination of STI indicators relative to the Key Issues and Questions identified by the Panel in their Interim Report.

6.1 Innovation measures

As discussed above, the Europeans have pioneered the collection of data on innovation through the implementation of the Oslo manual. The new BRDIS surveys already administered for 2009 and 2010 should give us more information on this activity. Unlike the 2008 survey, these surveys began with a page of questions about innovative activity, including the following:

1. Whether the firm introduced
   (a) new or significantly improved goods
   (b) new or significantly improved services

   (c) new or significantly improved methods of production
   (d) new or significantly improved logistics, delivery, or distribution methods
   (e) new or significantly improved support activities
2. Was any of the above new to the market or only new to the firm.[1]
3. Give percentages of sales due to 1) goods & services new to the market; 2) good & services new to the firm only; 3) goods & services that are not new. Must add to 100%.

As discussed above, existing analyses of data from the European surveys suggests that yes/no answers to the innovation question in (1) create very coarse indicators that provide very weak measurements of the impact of innovation on such things as productivity. Question 3 is more promising in this regard, but it provides quantitative impact data only for the innovations corresponding to questions 1a and 1b. This survey will not collect any information about the cost-savings and employment changes from the kinds of innovations covered by 1c, 1d, & 1e. Given the importance of these innovative activities in service-related sectors and non-R&D innovation, future work might want to explore improving the measurement of the benefits of process and organizational innovation. As discussed below, it would also be very useful to add a version of question 5.2 from the CIS (non-R&D expenditures on innovative activities) to the BRDIS survey.

6.2 Innovation in the service sector

There is no doubt that the service sector (broadly defined) is an increasingly important part of most developed economies: in the U.S., sectors producing services now account for 69% of private non-agricultural employment and 69% of value added in GDP as compared to 9% and 17%

[1] The way this question is posed might have been confusing if the firm answered yes to more than one item in (1).

respectively for the goods-producing sectors including manufacturing.[2] However, historically US data collection efforts related to innovative activity have focused on manufacturing, even though the service sector has been more important in economic terms than manufacturing for some time.[3] To some extent, the focus on manufacturing reflects a traditional science and technology view of innovative activity, one that centers on organized R&D laboratories and research employees. In the service sector, much of the effort at knowledge creation and innovation does not occur in organized research labs, and is undertaken by employees who are not categorized as researchers, making it harder to measure innovation effort.[4]

As services have become more important to the economy, and their innovative activity more obvious, it has become clear that more effort needs to be focused on data collection in services, where spending on R&D is a less useful measure of innovative investment. One can get a rough idea of the importance of innovation in the service sector from the results of the CIS surveys in the European Union.[5] In 2005, 39% of firms in the manufacturing sector reported that they had introduced either a product or process innovation (or both). For the service sector as a whole,

---

[2] *Economic Report of the President* (2011), Tables B-12 and B-46. The service-producing sectors includes utilities, wholesale trade, retail trade, transportation and warehousing, information, finance, insurance, real estate, rental, and leasing, professional and business services, educational services, health care, and social assistance, arts, entertainment, recreation, accommodation, and food services, and other services, except government.

[3] E.g., see the historical NSF statistics on R&D spending by sector, which reflect the bias toward manufacturing in the sampling frame during earlier years. In 1992, the non-manufacturing sector was reported as responsible for 25 per cent of R&D spending, whereas in 1980, only 4 per cent. These numbers are unlikely to accurate reflect the actual increase in spending among the non-manufacturing firms, due to changes in the sampling frame used for the RD-1 survey. For more information, see http://www.nsf.gov/statistics/s2194/conten3a.htm

[4] This issue is not limited to the service sector. Even in manufacturing, some research effort occurs outside of laboratories, perhaps increasingly so and these activities are harder to measure.

[5] Unfortunately published results from the NSF for the 2008 BRDIS survey (US-NSF, 2010) are not complete enough to report exactly similar numbers for the United States, although it does appear that while innovation rates are lower in services than in manufacturing, in some KIBS they are much higher (e.g., software).

the number was 34%, whereas for knowledge-intensive business sectors (KIBS) it was 52%.[6] These numbers clearly suggest the importance of capturing this activity by broadening of data collection efforts. One could argue that these numbers would be even higher in the United States, which has been a leader in ICT-related process reorganization and in the production of information services.

Thinking about important innovations in the service sector during the past few decades, we can see that many of them have been driven by the availability of networked computer and communication technologies: e.g., online reservation, ticketing, and load management systems in airlines, the growth of the logistics industry (Federal Express, DHL, UPS, and so forth); management of inventory systems that span regions or countries; financial innovation of various kinds. Thus past history suggests that innovation in this sector is frequently process-oriented, where changes in process may lead to new and improved products without explicit spending on R&D, and that one useful measure might be the level of investment in new ICT hardware and software. This view is supported by the UK data cited earlier (Hughes and Mina, 2012), where the acquisition of innovation-related machinery, equipment, and software in the service sector ranges anywhere from 30 to 90 percent of innovation expenditure, depending on the sector.

Data from the service sector can also be helpful in informing us about the pricing of certain kinds of research output. In particular, the R&D services sector sells its output on the market to other industrial sectors, as does the computing software sector. Given the difficulty of constructing price deflators for intangible outputs like R&D and software, data on production in these sectors may be useful (see Corrado *et al.*, 2011 for a new attempt to construct an R&D deflator).

[6] Manufacturing is defined as usual (NACE sector D). The service sector excluding KIBS covers Wholesale trade, Transport, storage & communications, Financial intermediation, Real estate, renting and business activities (excluding those in KIBS). KIBS includes Computer and related activities, R&D services, Architectural and engineering activities, and Technical testing and services.

6.3 Non-R&D investment that fosters innovation

In addition to expanding data collection on innovative activity and innovation, it would also be useful to have more information on the inputs to the innovation process that do not fall under the rubric of R&D and hence are not captured systematically by existing surveys in the U.S. These include:

- Acquisition of machinery, equipment and software used to produce new or significantly improved products and processes
- Training for innovative activities - Internal or external training specifically for the development and/or introduction of new or significantly improved products and processes
- Design - to improve or change the shape or appearance of new or significantly improved goods or services

The overall significance of these non-R&D forms of STI investment is difficult to assess, but they are clearly important, at least in some sectors. Hughes and Mina (2012, Figure 3) present data on the distribution of these forms of investment across sectors for the UK. Overall the two largest categories of spending in their figure are internal R&D and capital equipment. The share of innovation expenditure that is internal and external R&D is above 50 per cent only in computing, technical testing and analysis, motion picture and video production, R&D services, and the manufacture of electrical and optical equipment. In many other sectors, including financial intermediation, telecommunications, retail trade and repair, construction, mining, and post & courier activities, the bulk of innovation spending is on the acquisition of machinery, equipment, and software related to innovation. So if we ask what we are missing when we only collect R&D spending data, clearly the most important omission is the purchase of (possibly R&D-intensive) goods from other firms that are used for innovation in the purchasing firms.

As can be seen in Table One, our data related to the accumulation of human capital all relate to education and training in the education sector. But scientists, engineers and other workers also accumulate human capital on the job. Current indicators do include data on scientists and engineers engaged in work-related training, but these data are not connected to the innovation process. This will be particularly important with respect to the introduction and diffusion of new technologies, which may require knowledge and skills that current workers do not possess.

Non-R&D related innovation includes a great deal of design-related activity. In fact, judging from the numbers in Hughes and Mina (2012), design expenditures are relatively more important in several manufacturing industries as in the service sector. One way to try to measure design innovation might be to look at design patenting. At the USPTO, design patent grants were 5.7 per cent of all patent grants in 1963, rising to 8.7 per cent in 2011. Thus in spite of the enormous increase in utility patents during the same period (utility patents grew at 8 per cent per annum during the same period), design patenting grew even faster at 13.7 per cent per annum. From these number we can at least conclude that protecting designs with IPR has increased in importance. Unfortunately, an extensive search of the economics literature has found only one study on the use of design patents and firm performance or innovation.[7] This study, by Bascavusoglu-Moreau and Tether (2011), looks at registered design rights and firm performance (measured as labor productivity) in the UK, finding a positive association prior to 2000, but not after. Unfortunately their analysis is confounded by the fact that there were substantial changes to the benefits and costs of design registration in the UK with the introduction of European community design registration in the early 2000s and by their inability to control for the use of other (related) IP rights. Nevertheless, the study is a useful start to this kind of analysis.

6.4 Other issues

Timeliness The usefulness of indicators data is limited by the timeliness with which they are made available. As noted above, data from the 2009 and 2010 BRDIS were not available as of early 2012, although as

[7] There is a literature on design patenting, but it is primarily legal and historical in nature.

of the time of this revision (May 2012) they are becoming available in aggregate form. Results from the 2008 Survey of Doctorate Recipients were released only in late 2011. Since these data were collected before the acceleration of the financial crisis in the fall of 2008, as of now we have no information about how the Great Recession affected the job market for S&T degree recipients. Whatever the constraints within the statistical agencies that produce these delays, they greatly limit the value of the information for informing policy.

Linkages In order to gain meaningful understanding of the underlying processes, it is necessary in many cases to link data from different sources. For example, in order to understand how human capital created at universities contributes to productivity growth of firms, we need to be able to connect the individuals who are granted S&T degrees with their subsequent employers, and the performance of those firms. To go a step further, and analyze the return on the public investment that was made in those degree holders through research and training grants, we need to link the individual degree recipient to the specific grants that supported her education. Making these linkages, and exploiting them to test hypotheses and measure returns, is the province of social science researchers rather than the statistical agencies themselves. But the researchers need the underlying data to be captured in a way that identifies the entities involved in a way that permits subsequent linkage. We have at least one example where a linkage between the Survey of Earned Doctorates and the firms they went to was performed and the researcher was then prevented from publishing work based on this linkage due to NSF concerns about confidentiality. Such a match can be quite informative about firm needs and the supply if highly trained personnel and it ought to be possible to use the data statistically without revealing individual information. The Bureau of the Census has a great deal of experience with this kind of research now, via the federal Research Data Centers.

For looking at the sources and outcomes of federally funded research, the STAR Metrics project, launched by NSF with the cooperation of NIH, DOE and EPA, is an initial effort to capture information about all research grants to universities by these agencies in such a way that the people involved in the research can subsequently be connected to their activities (National Academy of Science, 2011, pp. 72-74).

Measures of knowledge advancement in specific policy-relevant areas The NSF-published indicators do not systematically present data on publication or patent trends in specific areas of public policy interest, such as environment, national security or health. The underlying data exist, however, to prepare such series, based on journal and patent classifications. Hence this is an example of an area where individual researchers, supported by public research funding to the extent the questions are deemed important, can provide the necessary indicator series on the basis of data that are available.

Capital available for financing of technology commercialization The existing indicators provide a wealth of data related to the financing of research. The only data they include on financing of the commercialization stage of innovation is on government programs (e.g. SBIR, ATP), and venture capital funding. This leaves a significant gap with respect to so-called "angel funding," i.e. private funding of startup or other small firms by individuals. A recent survey by the OECD (OECD, 2011) concluded that angel financing is an important source of capital for new and small firms in most countries, but is difficult to quantify because of the heterogeneity of institutions and the consequent difficulty of finding an appropriate target population for data collection.

Exports and imports An important measure of international competitiveness is the trade balance in advanced technology products. But this is difficult to measure when manufacturing is globalized, as the products whose profits accrue to U.S. firms may be manufactured and assembled in more than one country. Ordinarily these firms will use some form of transfer pricing when they move the components from country to country. The NSF S&E Indicators 2010 volume presents a number of high technology trade indicators for the U.S. and the rest of the world that are mostly derived from U.S. Census data. However it is unclear to what extent these data allocate the value addition created by these products correctly, given their multinational origins. The Apple iPod is a notorious example of a product whose market price is considerably in excess of its manufacturing cost (Dedrick *et al.*, 2010). How much of that value creation is allocated to the U.S. trade balance depends heavily on how the prices allocate value added among the various component suppliers and Apple itself.

7

Conclusion

Table 7.1 summarizes the relationship of the existing indicators to the Key Issues and Questions for STI Indicators put forward in the Interim Report. In general, the Table shows that the existing indicators do provide a wealth of information relevant to the key issues, with the possible exception of “Systemic Changes on the Horizon.” If one looks more deeply, however, at the specific questions identified in the Table, it is also clear that the indicators, in and of themselves, do not provide answers to the questions posed.

In our view, indicators and other data will never, in and of themselves, provide answers to questions of this kind. The questions posed are, for the most part, research questions. Answering them requires not just data but modeling and analysis. Indicators and data more generally are necessary inputs to the research process that can provide answers to these questions, but they do not provide the answers without that analysis.

It is therefore worthwhile to ask to what extent gaps in the existing data and indicators constitute important barriers to getting the answers we want. Our view is that, for the most part, the important barriers are not gaps in the data. In the previous section, we did identify several gaps and issues with existing indicators. If these could be addressed, it would facilitate better analysis of the important questions. But at a macro level, we do know the answers to many of the important questions. Investment in R&D is a major driver of productivity growth, and the rate of return to both private and public R&D investments is relatively high. Despite this relationship being clear on average, innovation is a very risky process, so that there is a lot of variance in the results of all innovation efforts. Growth in human capital is a key (perhaps the key) determinant of growth in income, both for individuals and for society as a whole.

Note that all of these questions for which we have pretty good answers are questions about how the world works, not the normative questions of what we should do to improve STI outcomes. These normative questions are much harder to answer. For the most part, the reason why these normative questions are hard to answer has little to do with STI indicators. First, they often involve value judgments about the *relative* value to society of different desirable outcomes, such as better health and a cleaner environment. In addition, some of the most important policy questions relate to the relative effectiveness or efficiency of different policy instruments, e.g. R&D tax credits versus public R&D or grants to universities versus funding government R&D labs. These questions are more answerable in principle than the pure value questions, but the answers require careful, systematic research rather than generic data collection. They could be answered if the Congress and the agencies were interested in systematic program evaluation, and were willing to allocate money for such evaluation. But they will not be answered by different STI indicators.

There is one category of normative questions that does, in principle, relate to indicators, and that is the allocation of public resources across different disciplines or areas of research. Although this contains an element of pure value judgment, it also clearly depends on the rate of return to research in different areas, which could theoretically be observed with ideal indicators. But in reality the incommensurable nature of research outputs in different areas, combined with the highly stochastic nature of research success makes it unlikely that better data is going to provide convincing measures of the differences in rates of return across fields.

Even if better STI Indicators will not provide answers to the "big" questions, there remain important improvements (some already underway) that the Panel can encourage. As discussed in more detail in the previous section, the most important of these include:

- Better coverage of the service sector in R&D and innovation surveys;
- Implementation of innovation surveys, with eventual expansion to include measures of cost savings associated with process innovation;
- Collection of information on investments in equipment and software in support of innovation;
- Collection of information on design efforts;
- Collection of information on training of employees for diffusion and adoption of innovations;
- More timely publication of indicators and availability of micro data to researchers;
- Collection and maintenance of data by grant-making agencies on individual grants and researchers in such a way that they can be linked to other data sources.

Finally, we note again that not all data collection and indicator publication has to be undertaken by the government. Particularly when it comes to experimental or innovative use of passive data collection, individual researchers (typically funded by public research grants) can in many cases collect and publish the data. So long as adequate research funding for such efforts is maintained, this is likely to provide for more extensive and ultimately successful development of new indicators than mandating collection of specific data by the government agencies.[1] Over

[1] Such research funding should also come with the clear mandate to make the data available for subsequent researchers so that the data can be tested and its utility for other uses explored.

time, if particular measures prove useful, it would then be possible to assimilate them into the arsenal of officially collected and published statistics.

Table 7.1: Key Issues and Questions For STI Indicators.

Issue	Current Indicators	Possible Additional Indicators
Growth, Productivity and Jobs What is the contribution of science, technology, and innovation (STI) activity to productivity, employment and growth? What is the relative importance of technological innovation versus non-technological innovation for economic growth? Is the United States falling behind with respect to innovation and what are the effects on socioeconomic outcomes?	Value-added Exports International royalties Foreign direct investment	Innovation revenues Innovation cost savings Service sector innovation
STI Activities What are the drivers of innovation? How influential is R&D for innovation and growth (by sector)? What would constitute a "balance" between the biological and physical sciences? On what basis could that be determined? Does biological science depend on physical science for advancement? How important are the following for advancing innovation: small businesses, large businesses, strategic alliances, technology transfer between universities and firms, academic researchers, government labs and procurement activities, and nonprofit organizations? What are the emerging innovative sectors and what is unique about them?	R&D expenditures by sector and performing organization Degrees Postdocs Education spending Test scores Technology alliances	Capital investment for innovation Design investment

Table 7.1: *continued*

Issue	Current Indicators	Possible Additional Indicators
STI Talent How much knowledge capital does the United States have? How many people, possessing what kind of skills, are needed to achieve a robust STI system? What additional sources of "talent" can best be tapped – especially among immigrants, women, and minorities? How many science and engineering doctorate holders took nontraditional pathways into the science, technology, engineering and mathematics (STEM) workforce? Did this vary by race/ethnicity, gender or the existence of a disability? How important are community colleges in developing human resources for STEM talent? Is the U.S. falling behind in STEM workers? What fields other than science, technology, engineering, and mathematics are important for advances in STI?	Enrollment Degrees Occupations (by country of origin)	Training for innovation adoption and diffusion
Private Investment, Government Investment and Procurement What impact does federal research spending have on innovation and economic health, and over what time frame? How large should the federal research budget be? How should policy makers decide where to put additional research dollars or reallocate existing funding streams – information and communications technology, biotechnology, physical science, nanotechnology, environmental technology, social science, etc.? Does government investment crowd out or energize private investment STI activities? What is the role of entrepreneurship in driving innovation?	Public expenditures broken down in various ways	

Table 7.1: *continued*

Issue	Current Indicators	Possible Additional Indicators
Institutions, Networks and Regulations What impacts are federal research programs having on entrepreneurial activities in science and engineering sectors? Where are the key gaps in the transfer of scientific and technological knowledge that undercut the performance of the STI system? Where is the supposed "valley of death" in innovation? In which industries is the valley of death most prevalent? What part of the process is underfunded for specific sectors? What is the nature and impact of intellectual property protection on scientific and innovation outputs?	Contract R&D ATP, SBIR, STTR Mfg. extension	Birth and death rates for innovative startups Survey evidence on abandoned projects (as in CIS)
Global STI Activities and Outcomes: What can we learn from other countries and what are other countries learning from us? In what technological areas are other countries accelerating? What impact does the international flow of STI have on U.S. economic performance? What is the relative cost of innovation inputs in the U.S. versus other countries? Where are multinational corporations sourcing R&D? What are the institutional differences that affect innovation activities among nations and how are they changing?	International Co-authoring Trade in R&D Royalties Foreign direct investment Foreign enrollment	Improved collection of cross-border R&D investment

Table 7.1: *continued*

Issue	Current Indicators	Possible Additional Indicators
Subnational STI Activities and Outcomes How does innovation activity in a given firm at a given place contribute to that firm's productivity, employment and growth, and perhaps also to these characteristics in the surrounding area? How are those innovation supply chains working within a state? Are firms principally outsourcing new knowledge from customers or from universities?	R&D data Patent data Student scores S&E doctorate holders	Innovation survey data on sources of knowledge
Systemic Changes on the Horizon How will demographic shifts affect the STEM workforce, nationally and internationally? Will it shift the locus of the most highly productive regions? Will global financial crises slow innovation activities or merely shift the locus of activities? When will emerging economies be integrated into the global ecosystem of innovation and what impact will that have on the system? How are public views of science and technology changing over time?	Consumer attitudes	See above

References

Abramovitz, M. (1956). "Resource and output trends in the United States since 1870". *American Economic Review*. 46(2): 5–23.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012). "The environment and directed technical change". *American Economic Review*. 102(1): 131–166.

Acs, Z. J. and D. B. Audretsch (1990). *Innovation and Small Firms*. Cambridge, MA: MIT Press.

Allman, K., J. Edler, L. Georghiou, B. Jones, I. Miles, O. Omidvar, R. Ramlogan, and J. Rigby (2011). *Measuring Wider Framework Conditions for Successful Innovation: A System's Review of UK and International Innovation Data*. London, UK: NESTA Report. URL: https://www.nesta.org.uk/sites/default/files/measuring_wider_framework.pdf.

Balconi, M., S. Brusoni, and L. Orsenigo (2010). "In defence of the linear model: an essay". *Research Policy*. 39: 1–13.

Bascavusoglu-Moreau, E. and B. S. Tether (2011). "Registered designs and business performance - exploring the link". In: *Design Economics.* Chapter 2. London, UK: UK Intellectual Property Office. URL: http://www.ipo.gov.uk/pro-ipresearch/ipresearch-right/ipresearch-right-design.htm.

Belitz, H., M. Clemens, C. von Hirschhausen, J. Schmidt-Ehmcke, A. Werwatz, and P. Zloczysti (2011). "An Indicator for National Systems of Innovation - Methodology and Application to 17 Industrialized Countries". DIW-Berlin Discussion Paper No. 1129.

Carson, C. S. (1994). *A Satellite Account for Research and Development*. November: Survey of Current Business. URL: http://www.bea.gov/scb/account_articles/national/1194od/maintext.htm.

Corrado, C., P. Goodridge, and J. Haskel (2011). "Constructing a Price Deflator for R&D: Calculating the Price of Knowledge Investments as a Residual". Working Paper No. EPWP#11-03. New York, NY: The Conference Board.

David, P. A. (2012). "Introductory comments on Innovation Incentives, Institutions and Economic Growth". In: *The Rate and Direction of Inventive Activity Revisited*. Ed. by J. Lerner and S. Stern. University of Chicago Press.

David, P. A., B. H. Hall, L. Soete, and A. van Zon (2011). *Global Warming and the S&T Research Policy Challenge: Neither the Ghetto nor the Moon*. draft: Stanford University and University of Maastricht.

Dedrick, J., K. L. Kraemer, and G. Linden (2010). "Who profits from innovation in global value chains?: a study of the Ipod and Notebook PCs". *Industrial and Corporate Change*. 19(1): 81–116.

European Union (2012a). *Innovation Scoreboard 2011*. URL: http://www.proinno-europe.eu/metrics.

European Union (2012b). *Key research & innovation indicators and economic analysis to support policy making*. URL: http://ec.europa.eu/research/era/facts/figures/key_figures_en.htm.

Fraumeni, B. M. and S. Okubo (2005). "R&D in the National Income and Product Accounts: A First Look at Its Effect on GDP". In: *Measuring Capital in the New Economy*. Ed. by C. Corrado, J. Haltiwanger, and D. Sichel. NBER conference volume, pages 275-322.

Griliches, Z. (1986). "Data Issues in Econometrics". In: *Handbook of Econometrics*. Vol. III, Chapter 21. Ed. by Z. Griliches and M. Intriligator. Amsterdam: North-Holland.

Hall, B. H. (2005). "Measuring the Returns to R&D: The Depreciation Problem". *Annales d'Economie et de Statistique*. 79/80: 341–381.

Hall, B. H. (2011). "Innovation and productivity". *Nordic Economic Policy Review*. 2: 167–204.

Hall, B. H., C. Helmers, M. Rogers, and V. Sena (2013). "The importance (or not) of patents to UK firms". *Oxford Economic Papers*. 65(3): 603–629.

Hall, B. H., J. Mairesse, and P. Mohnen (2010). "Measuring the Returns to R&D". In: *Handbook of the Economics of Innovation*, Volume II, pages 1032-1080. Amsterdam and New York: Elsevier.

Hughes, A. and A. Mina (2012). *The UK R&D Landscape*. CIHE and UK IRC Report for the Enhancing Value Task Force. URL: http://www.ncub.co.uk/rdlandscape.

Jankowski, J. (2012). "Update on the Business R&D and Innovation Survey". Presentation to the Fourth Meeting of the Panel on Developing STI Indicators for the Future, Washingon, DC, February 7.

Jankowski, J. (2013). *Measuring Innovation with Official Statistics*. In: *Handbook on the Theory and Practice of Program Evaluation*. Ed. by A. N. Link and N. Vonortas. Cheltenham, UK: Edward Elgar.

Jorgenson, D. W. and Z. Griliches (1967). "The explanation of productivity change". *Review of Economic Studies*. 34(3): 249–83.

Kline, S. J. and N. Rosenberg (1986). "An overview on innovation". In: *The Positive Sum Strategy*. Pages 275-305. Ed. by R. Landau and N. Rosenberg. Washington, DC: National Academy Press.

Mairesse, J. and P. Mohnen (2010). "Using innovation surveys for econometric analysis". In: *Handbook of the Economics of Innovation*. Volume II, pages 1129-1155. Ed. by B. H. Hall and N. Rosenberg. Amsterdam and New York: Elsevier.

Mowery, D. C. (2010). "Military R&D and innovation". In: *Handbook of the Economics of Innovation*. Volume II, pages 1220-1258. Ed. by B. H. Hall and N. Rosenberg. Amsterdam and New York: Elsevier.

National Academy of Science (2011). *Measuring the Impacts of Federal Investments in Research: A Workshop Summary* (Report of the Committee on Measuring Economic and Other Returns on Federal Research Investments, Steve Olson and Stephen Merrill, Rapporteurs). Washington, DC: The National Academies Press.

Nelson, R. R., ed. (1993). *National Innovation Systems: A Comparative Analysis*. Oxford University Press.

OECD (2011). *Financing High-Growth Firms: The Role of Angel Investors*. Technology, Industry Directorate's Committee for Industry, Innovation, and Entrepreneurship (CIIE) High-Growth Financing Project.

Peters, B. (2006). "Productivity effects of innovation activities". In: *Innovation and Firm Performance: An Empirical Investigation for German Firms*. University of Würzburg: PhD thesis. Chapter 4.

Popp, D., R. G. Newell, and A. B. Jaffe (2010). "Energy, the environment, and technological change". In: *Handbook of the Economics of Innovation*. Volume II, pages 874-930. Ed. by B. H. Hall and N. Rosenberg. Amsterdam and New York: Elsevier.

Rogoff, K. (2017). "The US is exporting obesity". URL: https://www.theguardian.com/business/2017/dec/04/us-obesity-trump-processed-foods-health.

Rosenberg, N. (1982). *Inside the Black Box*. Cambridge University Press: Cambridge.

Soete, L. (2012). "Maastricht reflections on innovation". Working Paper No. 2012-001. Maastricht, The Netherlands: UNU-MERIT.

Solow, R. M. (1957). "Technical change and the aggregate production function". *Review of Economics and Statistics*. 39(3): 312–320.

Tanaka, N., M. Glaude, and F. Gault (2005). *Oslo Manual: Guidelines for collecting and interpreting innovation data, third edition*. Paris: OECD/OCDE and Eurostat.

US Government (2011). *Economic Report of the President*. Council of Economic Advisors. Washington, DC: Government Printing Office.

US Government (2012). *Health Care Satellite Account*. Bureau of Economic Analysis. URL: http://www.bea.gov/national/health_care_satellite_account.htm.

US Government and National Science Foundation (2009). *Business R&D and Innovation Survey*. URL: http://www.nsf.gov/statistics/srvyindustry/about/brdis/.

US Government and National Science Foundation (2010). "NSF releases new statistics on business innovation". *SRS Info Brief NSF*. (October): 11–300.

Wessner, C. W., ed. (2009). *An Assessment of the SBIR Program at the Department of Defense*. Washington, DC: National Academies Press.

Author Biographies

Bronwyn Hall is Professor of Economics Emerita at the University of California at Berkeley, Visiting Professor at MPI-Munich, and Research Associate at the NBER and IFS-London. She has published numerous articles on the economics and econometrics of technical change and innovation and is the editor of the Handbook of the Economics of Innovation in the Elsevier series. Her research includes analysis of the use of intellectual property systems in developed and developing countries, the valuation of intangible (knowledge) assets, comparative firm level R&D investment and innovation studies, measuring the returns to R&D and innovation, and analysis of technology policies such as R&D subsidies and tax incentives. She has made substantial contributions to applied economic research via the creation of software for econometric estimation and of firm level datasets for the study of innovation, including the widely used NBER dataset for U.S. patents.

Adam B. Jaffe is Senior Research Associate with Motu Economic and Public Policy Research in Wellington, New Zealand and Adjunct Professor with Queensland University of Technology in Brisbane, Australia. He served Motu as Director and Senior Fellow 2013-2017. He was previously the Fred C. Hecht Professor in Economics, Chair of Economics and Dean of the Faculty of Arts and Sciences at Brandeis University in Waltham, Massachusetts, U.S.A. Jaffe is the author of two books – *Patents, Citations and Innovations: A Window on the Knowledge Economy* (with Manuel Trajtenberg, 2002); and *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress and What to Do About It* (with Josh Lerner, 2004). He is also the editor (with Ben Jones, 2015) of *The Changing Frontier: Rethinking Science and Innovation Policy.* In 2017, his paper with Richard Newell and Rob Stavins, "The Induced Innovation Hypothesis and Energy-Saving Technological Change" (1999) won the Association of Environmental and Resource Economists Award for Publication of Enduring Quality. Jaffe's research focuses on technological innovation and technology diffusion, particularly as it affects environmental and energy technologies, and on science and innovation policy.