Research Evaluation, volume 15, number 2, August 2006, pages 000–000, Beech Tree Publishing, 10 Watford Close, Guildford, Surrey GU1 2EP, England

Financial fluxes

Methodologies for the analysis of research funding and expenditure: from input to positioning indicators

Benedetto Lepori

This paper discusses the status of indicators concerning research funding and expenditure and proposes some pathways for further developments. First, I discuss in depth the design of the R&D statistics based on the Frascati manual and its limitations concerning analytical categories, data availability and quality. Further I argue that, to answer to specific policy questions concerning the allocation of funds, the development of a new generation of indicators is needed — so-called positioning indicators — focusing on the analysis of financial fluxes between research funders, intermediaries and performers, and I present some recent results of comparative European work in this direction. Finally, I draw some general methodological lessons on the nature of these indicators and on the procedure for their production, discussing key aspects such as reproducibility, quality validation, simplicity, contingency and transparency.

Benedetto Lepori is in the Research Service, Università della Svizzera italiana, via Lambertenghi 10a, CH-6904 Lugano, Switzerland; Email: blep[illegible]@unisi.ch; Tel: +41 58 666 46 14/7; Fax: +41 58 666 46 19.

A draft version of this paper was been presented to the Workshop on S&T Production in Europe: Status and New Perspectives, Lisbon, 22–23 September 2005.

THIS PAPER DISCUSSES the actual situation of indicators concerning research funding and expenditure, constituting a major component of the so-called input indicators for science, technology and innovation (Luwel 2004) and proposes some pathways for further developments. Namely, data and indicators concerning funding and expenditure on research activities seem to be a rather puzzling case. From one side this is the oldest and the most successful domain of S&T statistics where, thanks to the work of the Organization for Economic Cooperation and Development (OECD), data have been systematically produced since the end of the 1960s using a common methodology established in the Frascati manual (OECD, 2002). These data and some basic indicators, like the gross expenditure on R&D (GERD), are readily available online through the OECD databases and, thus, are frequently used by policymakers and researchers for their work.

However, at the same time, a careful analysis shows some limitations of these data, concerning their quality and comparability across countries, but also the information they contain and the degree of disaggregation (Godin, 2005a; Jacobsson and Rickne, 2004); some analysts went so far as to affirm that “although the OECD figures remain the conventional benchmark for international comparisons of higher education R&D, it is apparent that they are of increasingly limited utility for policy purposes” (Irvine *et al*, 1990: 5).

In this context, the aim of this paper is twofold. First, I will examine carefully both the methodological foundations and the practices of the official R&D

statistics based on the Frascati manual[1] and put it in the context of its original design and historical development since the 1960s; this in order to understand more precisely the kind of questions that can reasonably be answered using R&D statistics and the care needed to interpret these data correctly. Second, I will discuss some possible methodologies to produce other indicators on research funding that enable answering the policy and research questions that emerged in the last decades, namely the analysis of the repartition of funding between different instruments, a more detailed assessment of project funding and, finally, tools to examine funding and expenditure on research activities in the higher education sector. It will be seen that these methodologies rely on a careful (and mostly *ad hoc*) combination of different data sources, including also (but not only) R&D statistics.

I will argue that these are examples of a methodological shift in the production of S&T indicators from a set of indicators based on an input/output framework — hence aimed essentially to measure efficiency of research either through productivity indicators or econometric analysis — to so-called *positioning indicators*, aiming to position actors of national innovation systems and to identify linkages and fluxes between them (Godin, 2005a; Barré, 2006).

The paper is based on work of the author in the analysis of public research funding in the Swiss case (Lepori, 2006) and on collective work in the EU-funded network of excellence PRIME, particularly in the European Network of Indicator Producers (ENIP) project, where a mapping exercise of S&T data and indicators production in the ten participating countries was realized (see Esterle and Theves, 2005).[2]

R&D statistics

The domain of S&T indicators we are considering here concerns the financial resources devoted to research activities and, more precisely, research expenditure by performers (higher education organizations, public research institutes, private laboratories) as well research funds from other organizations and, especially, from public authorities. This is an important part of the so-called input indicators in S&T, alongside the measure of human resources devoted to research; the two domains are of course linked since a large part of research expenditure is for salaries.

Of course input indicators are not stand-alone, but have generally been developed to combine them with different types of output indicators — especially publication indicators (Van Raan, 2004) — or with economic indicators to develop measures of performance or productivity of research units and of whole countries, as well as measures of the impact of R&D on economic growth (Luwel, 2004).

This concerned essentially two aggregation levels, namely individual research organizations (for an overview see Bonaccorsi and Daraio, 2004) and whole national economies, either through simple productivity indicators (Barré, 2001) or more sophisticated econometric approaches trying to disentangle the contribution of R&D investments to economic growth (Mairesse and Sasenou, 1991).

Historical studies show that these issues were one of the major rationales leading to the development of the Frascati manual — along with the wish of national states to compare R&D efforts — and the underlying production function model largely inspired the design of the Frascati system (Godin, 2005b). This occurred even though in the 1960s very few output indicators for S&T were available (as readily acknowledged by the first edition of the manual). Despite the large body of literature available today on output indicators (see Moed *et al*, 2004 for an overview), the last edition of the manual still includes a short section on output indicators as well a rather comprehensive appendix on other S&T indicators, also covering such classical output indicators as bibliometrics and patent statistics (OECD, 2002: 200–211).

The origins and the Frascati manual

After early attempts before the Second World War, research expenditure data were first systematically produced in the USA by the National Science Foundation since the beginning of the 1950 (Godin, 2002): NSF data series on national expenditure for R&D go back to 1953 (NSF, 2004; Brown *et al*, 2004). At the beginning of the 1960s, the OECD entered in this field and in 1962 produced a methodological manual on how to measure R&D expenditure, the so-called Frascati manual (OECD, 2002). The manual proposed a set of basic definitions, including the definition of research and development and a classification of R&D types, categories and performance sectors, as well as a set of methodologies for measuring R&D expenditure including the set-up of specific surveys.

The manual has undergone five major revisions, which basically broadened its scope with the inclusion of social sciences and humanities in the 1970s and a wider inclusion of R&D in the service sector in the 1980s and 1990s; moreover, some methodological aspects were more precisely defined (for a detailed account of the revision history see the annex to the fifth edition of the manual: OECD, 2002: 151ff). However, it is important to notice that the general construction of the system and the main definitions and classification schemes have not been fundamentally modified since the first edition.

Since the principles and the content of R&D statistics based on Frascati are well known among S&T indicators specialists (see Luwel, 2004 for a concise presentation), I will limit myself to discussing some relevant issues for this paper, based particularly on the results of the ENIP project. These include an assessment of the successes and of the emerging shortcomings of this system and two possible explanations, that is the original design of the system — linked to the kind of questions it should answer — from one side, its early institutionalization and, thus, the difficulty of adapting it to a changing policy environment on the other side.

A success story …

R&D statistics has to be considered as *the* great success in S&T data and indicators production; in fact there is no other domain where data are produced systematically since the end of the 1960s in all OECD countries following a common methodology and mostly relying on specific surveys, rather than on the elaboration of data produced for other purposes. Thus, all ENIP countries except Hungary began R&D surveys around the end of the 1960s and the beginning of the 1970s. After the end of the communist regimes, Eastern European countries quickly adopted this methodology, as in the case of Hungary (Inzelt, 2005). Also, we notice that the use of Frascati manual — alongside of that of the Canberra manual on S&T human resources (OECD, 1995a) and of the Oslo manual on innovation (OECD, 1991) — has been implemented as the reference for R&D statistics by the European Parliament (European Commission, 2004) and thus has become the official methodological base for European Union R&D statistics; since this decision contains also a very detailed breakdown of variables to be produced, it becomes now very difficult for member states to depart from the OECD definitions and methodologies.

As Benoit Godin (2005a) notices, there are some reasons for this success: first, the manual was proposed by a supranational organization rather than by an individual state. Second, the OECD was very cautious in introducing it and from the onset discussed it with representatives of the member states; actually, even in its fifth edition, the title of the manual is still "*proposed* standard practice" and thus it is, at least in theory, left to the individual states to decide whether to adopt it or not. Third, the first edition of the manual came at a moment where very few countries collected S&T data and had specialized services for this, and thus it was much easier to introduce the new practices.

I would like to add some other reasons. First, the Frascati manual offers a very complete set of definitions and methodologies to produce coherent R&D statistics and the OECD succeeded in the following years in addressing in more detail some critical methodological issues, such as the measurement of R&D expenditure in higher education, some delimitation issues concerning private R&D, etc. In this sense, the quality and precision of the manual are really remarkable and most of the relevant questions for R&D statistics are discussed in great detail.

Second, from the onset, R&D statistics have been closely linked to a relevant policy question; that is, to compare the nation effort of each country in R&D and to try to link it with its level of economic development. Thus, already in the 1960s, the OECD used directly the first R&D statistics for international comparisons and for the national science policy reviews which started in 1962. Also, even at national level, R&D statistics were created to answer a policy need, as in the case of Switzerland at the end of 1960 (Lepori, 2006a). Bringing these numbers into the political arena meant thus also establishing them as the official measure of national research expenditure.

Finally, the OECD made a large effort in diffusing R&D statistics and, especially, putting them in a format suitable for international comparisons. This exercise involved reducing the complexity of national data to a set of basic indicators, especially the GERD/GDP ratio as the main benchmark for comparing countries (Godin, 2005a) and the matrix of R&D flows between funding and performance sectors, and producing them for most member countries in the same format. These data were regularly published in the annexes to the science and technology indicators reports and, from the 1990s, made available on CD-ROM and later on the Internet. Standardization and availability thus played a central role in promoting the use of R&D statistics.

Thus, the OECD went further than defining a methodology for collecting R&D data through the Frascati manual; namely, the OECD largely succeeded also in establishing how these data should be analyzed, which are the most important indicators to be used for international comparisons and how these indicators should be produced from national data series.

… but emerging shortcomings

However, among specialists of S&T indicators, the limitations of R&D statistical data are well known (see Godin, 2005a; Irvine *et al*, 1990). In reviewing them, I will distinguish between two main categories: data quality problems, owing to differences in coverage and quality of the collected data across countries; and limitations due to the lack of categories

Important differences between countries exist concerning the application of the definition and the methodology proposed by the Frascati manual and in the quality and coverage of the collected data

and classifications that would be needed for policy analysis The latter are thus, to a large extent, structural limitations of the Frascati system.

Despite the efforts of the OECD, important differences between countries exist concerning the application of the definition and the methodology proposed by the Frascati manual and in the quality and coverage of the collected data. These include, for example, differences in the definitions of performance sectors (especially concerning the coverage of the higher education sector), periodicity and coverage of the surveys, and missing data replaced by estimates by the national statistical office. To document these differences, the OECD developed a set of standard footnotes for its statistical publications and started in the 1980s a *Sources and Methods* series containing methodological notes provided by member states with their R&D data (OECD, 2000b); as Godin spells out, “a brief scan of the metadata is enough to shake one’s confidence in the reliability of international R&D statistics” (Godin, 2005a: 177).

The quality of data in two areas is considered to be particularly poor, namely, general university funds in higher education, and private sector R&D expenditure.

First, the estimate of the part of the general budget of the higher education organizations spent for R&D activities — the so-called general university funds — poses particular problems, since most universities do not have an accounting system distinguishing between research and education expenditure. Since the GUF accounts in most OECD countries for 50% to 70% of the public-sector R&D expenditure, their estimate is essential for the calculation of the public-sector research expenditure (considered as the sum of the expenditure of the higher education (HE) and government sectors). The methodology proposed by the Frascati manual is based on surveys of the use of time by staff (OECD, 2002: 158). Since problems emerged from the beginning, a detailed methodological annex was produced in the 1980s and introduced in the Frascati manual (OECD, 1995b; 2002: 158ff).

According to the ENIP survey, the application of this methodology varies greatly between countries. In a few countries, such as Switzerland and Norway, surveys are performed regularly on most of the staff and thus quite detailed data are available; however, in other countries the calculation is based on old surveys (Spain, 1983) or, even, it is not much more than guesswork done by the national statistical office. Thus, in France the share of R&D in the salary costs of HE personnel is a flat rate of 50%, irrespective of the university or discipline, while in Italy there is some differentiation according to the category of personnel and in Germany specific keys for scientific domain are used (OECD, 2000a); these keys are based on the result of old surveys and on an “educated guess” of S&T and, in some cases, there seems to have been some political pressure to modify them (for example, to increase to HERD national total).

Moreover, the whole foundation of this methodology is to some extent problematic: at least in some countries and in some scientific domains, research and teaching are so closely connected (concerning both individual activity and organizational setting) that the whole notion of separating them appears to be somewhat artificial; thus, there is some risk that the *subjective* evaluation by people of their use of time in a whole year period reflects also their perception of their own status and role in the university, rather than their actual use of time (Teichler, 1996; Jongbloed and Salerno, 2004). Also, time surveys of individuals require large resources and this makes it difficult to implement them, especially in large countries. Only a widespread introduction of an accounting system separating research and education costs from the beginning would really improve the situation.

Finally, work performed in the PRIME-AQUAMETH project showed large differences in the coverage of higher education expenditure, concerning investments and capital costs (depending also on different legal status and accounting systems), perimeter of the university (for example, university hospitals and student services) and the inclusion of PhD students (Bonaccorsi *et al*, 2006). The most problematic issues are accounting for capital costs and the separation between research and healthcare in clinical medicine (OECD, 2001); both concern large shares of higher education expenditure and thus could readily lead to false international comparisons.

The second major domain of concern is private R&D expenditure data. Besides accounting for most of the R&D expenditure in OECD countries, these data are particularly sensitive since private research is thought to be a key for economic development. In the ENIP survey, the opinion was shared that these data are nowadays of little use for policy analysis, except for very aggregated comparisons. Three problems stand out: quality of surveys and of data; access to micro-data; and treatment of internationalization and multinational companies.

Surveys in the private sector always posed particular difficulties, since the effort to survey all business enterprises would be too large; different strategies across countries in defining the sample have been used as the Frascati manual doesn’t give very clear guidelines (OECD, 2002; 127ff). Moreover, since R&D is not a separated cost centre in most enterprises’ accounting systems, R&D data have to be estimated from definitions and examples accompanying the survey; this can lead to large differences according to the wording of the definitions (and their translations in national languages), to examples given and also to the person filling in the form. As a result, R&D expenditure measured in an R&D survey differs significantly from that measured in a Community innovation survey (Godin, 2005a: 151–152). Similar problems have emerged in the recent assessment of the business R&D survey in the USA (Brown *et al*, 2004).

An additional difficulty is that, for confidentiality requirements, micro-data are usually not available even for research purposes; according to the ENIP results, even in large countries published data are disaggregated only by industrial sector, region and some size classes, and data are usually not disaggregated at the same time by all these categories to avoid easy identification of major R&D performers.

This limits the value for analysis and makes impossible cross-checking with other data (eg business reports). Finally, in the context of increasing internationalization of research activities, the whole treatment of multinational companies is based on breaking them down to national subunits and measuring national expenditure of subsidiaries (OECD, 2002: 61). In the mid-1990s, the OECD modified slightly the structure of the business R&D survey, by including a distinction between national companies and subsidiaries of companies located in a foreign country and by including a question on R&D expenditure of subsidiaries abroad (Godin, 2004; OECD, 2004). However, a look to the MSTI database shows that, actually, for most countries only one or two data exist for the period 1995–2004. Even these data show the relevance of the phenomenon, since for countries such as France or Germany R&D expenditure by affiliates abroad are one fifth of total national firms' R&D expenditure, while for Switzerland expenditure abroad by affiliates is actually larger than domestic expenditure. The newly published handbook on economic globalization indicators (OECD, 2005) deals at length with the internationalization of R&D and technology and proposes a set of indicators to be collected through a specific survey. However, it remains to be seen whether member countries are ready to collect these data and to what extent they are consistent with those from the R&D surveys.

A further difficulty concerns time series. Even if in all ENIP countries (except Hungary) R&D expenditure data had begun to be collected at the beginning of the 1970s and even if the OECD MSTI database went back to 1981, most ENIP correspondents signalled major breaks in series at the beginning of the 1990s, due to important changes in methodology and definitions. For instance, in the Swiss case, higher education statistics were completely revised at the beginning of the 1990s and later series are not comparable with the preceding years. Major breaks in series have been signalled in the ENIP reports for France in 1992, Italy in 1997 and Norway in 1993–1995. Even if these breaks are documented in the MSTI database, reconstructing coherent time series proves to be a painful task, especially where national statistical offices did not deliver very detailed information on change to the OECD. In the Swiss case, I showed that it is actually possible to connect the two series, but that this requires a painful work of correction and estimation, going back to original documents to assess how the data have been collected (Lepori, 2006).

This limitation is particularly relevant for research policy and economic analysis, since in most cases only sufficiently long time series allow the detection of quantitatively structural changes in research policy and in the research system. My hypothesis is that this reflects largely the political use of these data and the related limited interest for long-term series (Godin, 2005a).

The limitations of the original design …

It is worth noticing that most of these problems emerged in the 1970s, were discussed at length at the OECD and led in many cases to more detailed guidelines in the successive revisions of the Frascati manual. In my view, their persistence despite these efforts is a sign that at least some problems are to a large extent of structural nature, depending on how the system was originally designed and on the way it was successively institutionalized in the statistical and political system.

In fact, even if today we identify R&D statistics with the measure of research funding and expenditure in general, we should remember that the Frascati manual was essentially developed to answer to a single question: to measure the national effort in research and to compare it between countries. This had two major rationales: first, for policymakers, to assess if their country was spending enough for research and to catch up with the USA. Thus, already the introduction of the first edition of the Frascati manual set the level of 3% of GDP for GERD — according to the NSF statistics the level of the USA at that time — as a target for all member countries (Godin, 2005a). Second, with the emergence of the economy of innovation in the 1960s, a comparable measure of research expenditure was required to look for quantitative correlations with the economic development.

This implied two major choices in the design; that is, focusing on input rather than on fluxes and producing national aggregates rather than data on sectors or on individual organizations.

First, the whole of R&D statistics has been constructed to produce a *national aggregate* for R&D expenditure and this necessitated finding common measures of them across sectors (implying also the use of monetary units). This for example necessitated delineating a method for separating education and research activities in universities, which however has little meaning in the functioning of these organizations and has been unnecessary for a sectoral analysis. Also, it was impossible to design a cross-sector classification of R&D activities since it is most likely that R&D means different things in different organizational contexts: thus, higher education R&D is classified according to the disciplinary organization of universities (main fields of science and technology), while private R&D is classified on product fields using the standard classification of economic sectors. The use of these generic categories, which do not match the economic reality of the organizations, is a major reason for the poor quality of the surveys.

Second, the whole of R&D statistics is concerned with measuring financial inputs of performers (R&D expenditure), while it devotes limited attention to data on R&D funding, which are basically constructed by aggregating performer's declarations. Moreover, the tracking of the funding fluxes between funders and performers, and the importance of different channels, such as general funds and projects funds, as well as the role of intermediary agencies such as research councils are largely out of the scope of the Frascati manual. Yet, in the analysis of research policies it is widely acknowledged that the composition of funds and their allocation mechanisms are as important as their volume in determining the functioning of a research system (Millar and Senker, 2000; Braun, 2003) and most of the recent economic literature has concentrated on these issues (Geuna, 2001; Geuna and Martin, 2003).

Actually, in some countries, such as Switzerland and Norway, statistical offices check R&D data from performers with the funding agencies and, usually, a rather good consistency is found for aggregated data (see OECD, 2002: 119–120), but concerning individual items differences can be larger. Later, the OECD proposed a methodology to measure the research content of state appropriations — the so-called government budget appropriations or outlays for R&D (GBAORD; OECD, 2002: 137ff) — but this methodology proved to be difficult to apply and the data are of little use, since appropriations are divided by 'generic' socio-economic objectives and not by funding agency and instruments.

… and the inertia of a successful system

It is my opinion that many of the actual shortcomings of R&D statistics are to some extent a consequence of their original design, but also of the way they were institutionalized both at the OECD and in the member countries, which made it difficult to implement main changes and to follow new requests, especially from research policy analysts

First, from the outset, the aim of Frascati was to build a R&D statistical system, based on precise definitions and methodology, which would allow the production of coherent and comparable data sets across countries; this meant also designing and implementing special surveys, since achieving this objective would have been impossible using secondary data. Of course, establishing and running the system in all OECD countries was an extraordinary success; but it implied also a rigidity, since changes in definitions and methodologies translated in most cases into modifications of the questionnaires and data collection procedures to be coordinated across all member countries. Then only incremental changes — mostly clarification of definitions and methodology, as well as some additions to the questionnaires — became possible. For instance, the ongoing revision of the fields of science classification has been essentially limited to adding some more subdivisions at the two-digit level, an exercise which is not really useful since in most countries only one-digit-level data are available. Moreover, the experience shows that many countries are not ready to invest more resources in R&D statistics and, actually, in most cases the data collected are less detailed than the Frascati manual recommends.

Second, in many countries R&D data and indicator production was from the beginning integrated in national statistical offices. Among the ten ENIP countries, only in two cases (Norway and Portugal) are R&D statistics produced by specialized units (Esterle and Theves, 2005). This is completely different from bibliometric data and indicators, which are normally produced by specialized units, in many cases also active in research in the field and more closely linked to the academic community (eg CWTS in the Netherlands, ISI-Fraunhofer in Germany, OST in France). Even when many R&D statisticians were highly educated and experienced in this domain, this organizational setting made the communication between researchers in the field and data producers more difficult. Thus statisticians concentrated rather on improving datasets in the existing framework, while researchers tried to use existing R&D data — taken for granted — to answer to new questions, leading to rather unsatisfactory outcomes. Also, proximity to the state meant that political interests and needs for legitimating public policies had strong impact on R&D statistics, thus limiting the scope for new developments (Godin, 2005a).

Moreover, the leading role of the OECD in the introduction of the Frascati manual and its development, notably through the NESTI committee, played an essential role in diffusing the Frascati methodology to the member countries, but later this consensus-based system constrained its evolution since every change required a long process of consultation with national representatives. At the same time, since the OECD made direct use of these data for policy analysis both at the national level (national reviews of S&T policies) and for international comparisons and benchmarking (eg through the STI outlook series), member states have been well aware of the political implications of changes in these indicators and thus highly attentive to them.

A strategy for new indicators

The preceding discussion shows both the strengths and the limitations of R&D statistics and the fact that these are to a large extent related to their original design and form of institutionalization, rather than to weaknesses in implementation. Moreover, it appears that official R&D statistics have become a system so complex and so deeply institutionalized, both in the functioning of the statistical offices and in the political discussion, that it is unrealistic to call for major changes at least in the short and medium term. This doesn't mean that efforts to improve the system and, in particular, the quality of collected data in some domains should be abandoned but, from the perspective of policy analysis, it is unlikely that this alone will provide the data needed to answer the issues dealt with in the next sections.

This is why I argue in this section for the development, alongside and complementary to R&D statistics, of a set of indicators on research funding and expenditure allowing answers to specific policy questions, such as the repartition of funds between funding channels or the role and composition of the so-called project funds.

The approach used to build these indicators is in a sense much less rigorous than that proposed in the Frascati manual. Namely, we build *ad hoc* (and thus contingent) indicators to answer specific questions, making use of all kinds of existing data rather than trying to develop a coherent set of financial statistics on R&D funding and expenditure as in the Frascati manual. The advantage is the possibility of exploiting and recombining existing secondary sources — based of course on some understanding of the phenomena to be observed — rather than having recourse to systematic surveys.

ENIP mapping of S&T indicator production in European countries (Esterle and Theves, 2005) shows that in all countries there exists a wealth of data on research funding and expenditure:

- Higher education financial data, either from national sources for countries having a developed national higher education statistics, or from accounts of individual universities (eg collected by the rectors' conference, as in Italy).
- Data from state budgets or outlays.
- Data from national reports on public research funding published by ministries, as in Germany or Austria, where they contain a very detailed breakdown of public funds (BMBF, 2004).
- Data on project funding from ministries and from funding agencies' yearly reports (available in most ENIP countries).

A final source of data is *national* R&D statistics. In fact, much of the information contained in national data from R&D sources gets lost when they are put in the standard format of the OECD databases. This includes many details based on national categories, but also most of the information on coverage and data quality. The author's experience in the Swiss case showed too that, despite the OECD effort to document methods, a substantial part of the information on methodology and data quality is either on internal documents of the statistical offices or is informal knowledge of the S&T statisticians.

In the following, I will discuss two examples of how this work could be done and comment on methodological problems which emerged in this work. These are the construction of structural indicators for public funding of research and comparative analysis of project funding. I will then come back in the last section of the paper to some basic issues underlying the approach proposed here.

Indicators of the structure of public funding

The first exercise I propose aims at developing some indicators concerning the *structure* of public funding integrating the data available in the R&D statistics. As is well known, R&D statistics provide basically a breakdown of public funding according to the main performing sectors, based on five sectors: government, higher education, private, private non-profit and abroad. The Frascati manual provides quite precise delimitation of these sectors.

Based on these data, it is possible to compare the countries according to some indicators such as the share of the government and the higher education sectors in public research and the share of public funds directed to the private profit sector.

However, these indicators overlook a basic distinction concerning research funding between general funds, which are attributed globally to research performers for their institutional tasks (leaving the repartition of funds to internal decision-making processes) and projects funds, being attributed to research teams to perform specific tasks (normally over a limited time period). The simplest structural model of public research funding is presented in Figure 1 (Millar and Senker 2000). This model depicts fairly well the structure of public funding of research in most OECD countries, except in countries where public research and higher education are so closely linked that a separation becomes very difficult (the main case being France).

If we fill in Figure 1 with ratios indicating the percentage of the flows, we get a very simple structural diagram that can be used either to compare countries in the same year or to analyze the evolution over time of the structure of public research funding going beyond existing OECD indicators. The strategy I suggest is a minimal one, using as much as possible R&D statistics, but integrating them with some other sources and, possibly, estimates where data are not available or we know that they are not reliable. Thus, in the ENIP action on public funding we have

R&D statistics provide basically a breakdown of public funding according to the main performing sectors, based on five sectors: government, higher education, private, private non-profit and abroad

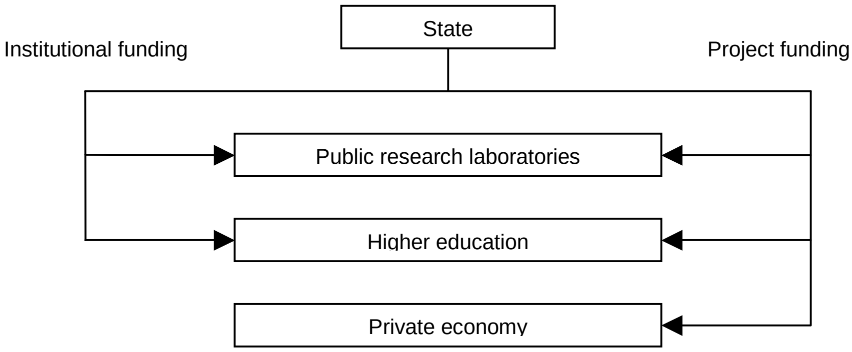

Figure 0. Structural indicators of public funding

developed a methodology to estimate from the funding agency data project funding, information which is non-existent in R&D statistics.

The second major issue concerns higher education R&D expenditure, where data are less reliable and cover in many cases only few years. Since in most countries data on the total higher education expenditure can be retrieved from different sources (Bonaccorsi *et al*, 2006), I suggest going back to these data and then introducing only at a second stage the coefficient for the share of R&D expenditure, using where available results from R&D sources. The sensitivity of the indicators to different coefficients can then be tested. This approach is particularly useful for time series, since it becomes possible to test different assumptions concerning the evolution of the coefficient over time.

With these techniques, I showed that in the Swiss case it is possible to produce these figures for the whole period 1970–2002 and this actually leads to the interesting result that little changes happened in project funding over this period (Lepori, 2006). Similar diagrams have been produced for Italy (Potì and Reale, 2005) and for Austria (Dinges, 2006). An extension of the scheme to the specific case of France, including 'human resources funding' through CNRS, is in development.

We notice that this model could be improved in a second dimension; that is, the classification of public research organizations. The division proposed by the Frascati manual consisted simply in separating the higher education sector from the whole public sector, as defined in the system of national accounts. It appears from research policy studies that this classification is not well suited to catch some relevant distinctions:

- The existence of associated laboratories between public research institutes and universities, especially in countries like Portugal and France. These laboratories are a structural feature of the public research system in these countries and thus splitting up their research expenditure for the sake for comparability makes little sense.
- The diversity of the statute and mission of government sectors, including for example inhouse research in ministries (closely linked to their political tasks), mission-oriented public laboratories but also public laboratories for academic research, as in most MPG institutes in Germany or national research facilities managed by research councils in the UK.
- The diversity in the higher education sector, for example between organizations conferring PhD degrees and the higher education second sector (so-called universities of applied sciences); at least in countries such as Finland, Germany or Switzerland these organizations are becoming a relevant actor in public research.
- The treatment of international research organizations, which cannot be found anywhere in the R&D statistics. Some of them should be assimilated into national research laboratories performing fundamental research (eg CERN), while others act rather as funding agencies redistributing most of the money they receive from national governments (eg the European Space Agency).

A more refined division could comprise some additional categories that are found to be relevant for the organization of national research systems. Thus I suggest that, rather than starting from a predefined classification and trying to fit the reality to it, it could be more feasible to built the relevant categories from the bottom up. Since, by the methodology of the R&D surveys, R&D expenditure data in the public sector are already disaggregated by performer at the level of individual institution, it should be feasible in many countries to reclassify existing data, thus having a more comparable perspective on the structure of public research systems. National data at a disaggregated level are in this respect much more useful than the aggregated OECD figures.

Analysis of project funding

Project funding — defined as money attributed to a group or an individual to perform a research activity limited in scope, budget and time — is a category of central concern for the analysis of research policies. Not only does project funding account for between one quarter and one third of public funding of research, but also it has been a choice instrument for steering research activities towards political and economical goals. For these reasons, assessing differences between countries in the amount and composition of project funds, as well as their evolution over time, would be quite important (Lepori *et al*, 2005).

Yet, the state of data in R&D statistics is particularly poor in this domain. We notice that, at least for the higher education sector, the Frascati manual deals at length with this issue (OECD, 2002: 166–169), but the manual itself already recognized that third-party funds might be poorly recorded in university accounts. In the recent OECD project on steering and funding of public research institutions only four countries provided some data on project funding and only for a four-year span (OECD, 2003: 85). Moreover, even if these data existed, they would be of limited value since they are not disaggregated by funding agency (eg between intermediaries and ministries) and type of instruments.

In the framework of ENIP a small group of countries (Austria, Italy, Portugal, Spain and Switzerland) tested a different approach based on the use of data from the funding agencies. Basically, the whole methodology is based on the identification of a list of project instruments for research activities, such as funds from research councils, programs managed by the research ministry and European research programs. Three national reports, on Austria, Switzerland and Italy, have been published (Dinges, 2006; Lepori, 2005; Potì and Reale, 2005). A first comparative analysis displays profound differences in the orientation of project funding in these countries — concerning the managing agencies, the portfolio of instrument and the beneficiaries (Lepori *et al*, 2005).

There a number of methodological lessons to be drawn from this work.

1. The *choice of the standpoint* has been essential both for the feasibility of the exercise and for its political relevance. Funding agencies data are more or less available in all countries, even if collecting them and correcting for some national specificity might require some time. Moreover, if we want to study national policies, it is relevant to examine the instruments as declared by the funding agencies rather than their actual use by researchers (which could be, to some extent, different). Collecting project data from performers is much more complex, as is well known from R&D statistics.
2. The approach is also based on some *underlying assumptions* built on our knowledge of the national research system. Thus, we assume that, even if borderline cases might exist, it is possible to identify *for each country individually* a list of instruments constituting the bulk of project funding of research activities and that this list is sufficiently short to be manageable. Moreover, we accept also that overall most of the funds distributed through these instruments is effectively used for research activities and, thus, that there is some correspondence with R&D statistics. Both assumptions can be of course verified where more detailed data are available (as in the Swiss case).
3. Rather than give general definitions, we prefer to rely on a *set of criteria* for inclusion or exclusion of instruments and to classify them; criteria to be applied by national correspondents on the basis of their knowledge of national systems. It is only the combination of data and expert knowledge that allows producing meaningful indicators, by addressing delimitation issues which would be impossible to deal in a general framework.
4. Finally, *transparency and flexibility* have been central for the whole exercise. The basis is a list of instruments, for which data are gathered individually and separately from original sources. All kinds of aggregation are then based on this list. Thus, it is possible to reclassify instruments or to compare subsets of them — for example, only instruments managed by research councils — or to add new instruments if further analysis makes it advisable. Since we acknowledge that some of the choices or categories we are using could be revised in the future, the design of data provides purposefully for more flexibility than R&D statistics.

Conclusions

This paper has largely an explorative character, to indicate some examples of methods to produce new indicators concerning R&D funding and expenditure and to review ongoing work in the PRIME project. At least in the Swiss case, I could show that these methods can be successfully employed for analysis of research policy and that they lead to results that go beyond what is possible using official R&D statistics only (Lepori, 2004, 2005). Also, preliminary results of PRIME projects concerning project funding (PRIME-funding activity) and higher education funding (AQUAMETH and CHINC projects) are encouraging in that these methods can also be employed for cross-country comparisons (Bonaccorsi *et al*, 2006; Lepori *et al*, 2005, 2006).

Actually in general terms the described indicators are good examples of what Rémi Barré called *positioning indicators*, which consider a national innovation system to be made of differentiated, autonomous and strategic agents. Thus, its characterization cannot be derived from the analysis of its individual components, nor by the measurement of aggregates built from them, but one needs to describe the actors of the systems in terms of their interactions — linkages as well as complementarities, competition and cooperation — and to figure out the types and categories they belong to; in sum to characterize their “positioning” in the system (Barré, 2006).

To conclude the paper, it is useful to summarize some general features of the approach proposed here.

1. Instead of developing a general statistical system on research funding and expenditure, this is an attempt to develop *ad hoc* indicators to answer to specific questions and this, actually, distinguishes this work from classical indicators produced by statistical offices (Barré 2006). This approach has some advantages: first, we are free to choose the definitions and the methodology best suited to the kind of question posed and, also, to the real availability of data; second, it becomes possible to do more careful work in checking and validation, since there are much less data to be produced each time. The drawbacks are the limited reusability and scalability of the work. Coherency is also not completely ensured since the definitions can be different from case to case, at least in the strict statistical sense. Thus most of the work in described projects has been devoted to designing procedures that guarantee some degree of comparability between the produced indicators and allows reproducing them (both for future years and for other countries). This is an essential step to go from case-by-case development of descriptors to indicators, which can be compared and, to some extent, reused.
2. The intelligence on the structure of research policy and research systems is crucial: only if definitions, categories and indicators match our knowledge about the functioning of policies and research activities, can we hope to produce indicators relevant for policy analysis. For instance, the whole exercise on project funding was based on its understanding — developed by research policy analysts — as a distinct way of funding research, with its organizational forms, rules and allocation mechanisms. Also, the decision whether or not to separate research and education expenditure in the higher education sector should depend on our understanding of how universities function and whether, in their real life, the two activities are really separated. This intelligence becomes even more important when data are lacking or of poor quality and need to be corrected or estimated. One should always recall that indicators are descriptors of a complex reality and, thus, that theoretically informed simplification is as important for their production as precision in the measurement.
3. One has to accept that the same data might have different meanings for different countries since the institutional structures are not the same. By the way, national differences in the organization of research policy and research activities are a crucial difference between the European Union and the USA that affects all the work on S&T indicators and makes it more difficult to develop meaningful European-wide indicators. Abstracting data from their national contexts entails the risk of drawing largely incorrect conclusions, even if the underlying data are correctly measured. This is also the reason why the projects discussed have been organized as networks of national correspondents knowing very well not only their data, but also their national systems. Thus positioning indicators are always contingent and linked to a specific use context.
4. Finally, transparency is essential in S&T indicators production. If we accept that our indicators are only very approximate representations of reality and that in many cases even the quality of the underlying data is questionable, then a careful documentation of data and treatment methodologies becomes essential, not only towards the S&T specialists, but also towards the users. Concerning critical aspects, such as the estimate of the share of research expenditure in universities, we could even decide to leave this as a free parameter that the users could modify to test different hypothesis and to perform sensitivity analysis. Actually, this approach turns out to be an advantage, since it becomes perfectly acceptable to make *ad hoc* choices and estimates when needed, provided that they are justified by reasonable arguments and are reported so that future analysts could also test different assumptions.

This research-driven approach, largely based on the *ad hoc* elaboration of data sets case by case, is by no means in contradiction with the more systematic and, in a statistical sense, rigorous approach promoted by the OECD and by the national statistical services. In reality, the experiments proposed here would be much more difficult without the existence of the R&D statistics and, if they are to be really successful and widely used, the new indicators concerning research funding should go into the future through some kind of systematization and institutionalization process as the R&D statistics in the 1960s.

Acknowledgements

This paper owes very much on the work and experiences in different project of the EU-funded network of excellence PRIME, especially to the European Network of Indicator Producers project and to its funding activity and to the AQUAMETH project, as well as to the EU-funded contract on changing incomes and their impact on university research (CHINCH). The author wishes particularly to thank for their valuable comments and support Rémi Barré, Laurence Esterle, Benoît Godin, Stig Slipersaeter and Jean Theves. I wish also to thank the Science and Technology Division of the OECD for granting access to internal documents, including the *Sources and Methods* database.

Notes

1. In this paper, I will coherently use the wording ‘R&D statistics’ to refer to data collected according to the Frascati manual and based on its definition of research and development (OECD, 2002: 30). I will use the more generic terms ‘research funding’ and ‘research expenditure’ to refer to other financial data,

which are not necessarily based on an analytic definition of research activities.
2. Italy, Israel, Switzerland, France, Germany, Hungary, Spain, Austria, Portugal, Norway.

References

Barré R 2001. Sense and nonsense of S&T productivity indicators, *Science and Public Policy*, **28**(4), 259–266.

Barré R 2006. Towards a European STI Indicators Platform (ESTIP), position paper to the Second PRIME Network of Excellence Annual Conference.

BMBF 2004. *Bundesbericht Forschung 2004*, Berlin.

Bonaccorsi, A and C Daraio 2004. Econometric approaches to the analysis of productivity of R&D systems. In *Handbook of Quantitative Science and Technology Research*, eds. H Moed, W Glänzel and U Schmoch, pp. 51–74. Dordrecht: Kluwer.

Bonaccorsi, A and C Daraio eds. 2006. *Universities as Strategic Units*. Aldershot: Edward Elgar, forthcoming.

Bonaccorsi, A, C Daraio and B Lepori 2006. Indicators for the analysis of higher education systems: some methodological reflections. In *Universities as Strategic Units*, eds. A Bonaccorsi and C Daraio. Aldershot: Edward Elgar, forthcoming.

Braun D 2003. Lasting tensions in research policy-making – a delegation problem. *Science and Public Policy*, **30**(5), 309–321.

Brown, L *et al* 2004. *Measuring Research and Development Expenditures in the US Economy*. Washington, National Academies Press.

Dinges M 2006. *Project Funding. Country Report: Austria*. Paris: European Network of Indicators Producers.

Esterle, L and J Theves 2005. Analysis of the different European systems for producing indicators, paper presented at the Workshop on S&T Indicators Production, Lisbon, 22–23 September 2005.

EC, European Commission 2004. *Commission Regulation (EC) No 753/2004 of 22 April 2004 implementing Decision No 1608/2003/EC of the European Parliament and of the Council as regards statistics on science and technology*. Brussels: EC.

Geuna, A 2001. The changing rationale for European university research funding: are there negative unintended consequences?. *Journal of Economic Issues*, **35**(3), 607–632.

Geuna, A and B Martin 2003. University research evaluation and funding: an international comparison. *Minerva*, **41**, 277–304.

Godin, B 2004. Globalizing S&T indicators: how statisticians responded to the political agenda on globalization. *Project on the History and Sociology of S&T Statistics*, Working Paper 27.

Godin, B 2005a. *Measurement and Statistics on Science and Technology: 1920 to the Present*. London: Routledge.

Godin, B 2005b. Science, accounting and statistics: the input-output framework. *Project on the History and Sociology of S&T Statistics*, Working Paper 31.

Inzelt, A 2005. *National Systems of S&T Data/Indicators Production: Country Report on Hungary*. Paris.

Irvine, J, B Martin and P Isard 1990. *Investing in the Future: An International Comparison of Government Funding of Academic and Related Research*. Aldershot: Edward Elgar.

Jacobsson, S and A Rickne 2004. How large is the Swedish "academic" sector really? A critical analysis of the use of science and technology indicators. *Research Policy*, **33**, 1355–1372.

Jongbloed, B and C Salerno 2004. *Kosten per Student. Methodologie, schattingen en een internationale vergelijking*. CHEPS report, Twente.

Lepori, B 2004. xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

Lepori, B 2005. *Project Funding. Country Report: Switzerland*. Paris: European Network of Indicators Producers.

Lepori, B 2006a. *La politique de la recherche suisse: Acteurs, institutions et développement historique*. Bern: Haupt, forthcoming.

Lepori, B 2006b. Public research funding and research policy: a long-term analysis for the Swiss case. *Science and Public Policy*, **33**(4).

Lepori, B, M Benninghoff, B Jongbloed, C Salerno and S Slipersaeter 2006. Changing models and patterns of higher education funding: some empirical evidence. In *Universities as Strategic Units*, eds. A Bonaccorsi and C Daraio.Aldershot: Edward Elgar, forthcoming.

Lepori, B, M Dinges, B Potì and E Reale 2005. Public project funding of research activities. a comparative analysis in six European countries, paper to presented at the Workshop on S&T Indicators Production, Lisbon, 22–23 September 2005.

Luwel, 2004. The use of input data in the performance analysis of R&D systems. In *Handbook of Quantitative Science and Technology Research*, eds. H Moed, W Glänzel and U Schmoch, pp. 315–338. Dordrecht: Kluwer.

Mairesse, J and M Sasenou 1991. R&D and productivity: a survey of econometric studies at the firm level. *Science-Technology Industry Review*, **8**, 9–43.

Millar, J and J Senker 2000. *International Approaches to Research Policy and Funding: University Research Policy in Different National Contexts*. Brighton: SPRU.

Moed H, W Glänzel and U Schmoch 2004. *Handbook of Quantitative Science and Technology Research*. Dordrecht: Kluwer.

NSF, National Science Foundation 2004. *Science and Engineering Indicators 2004*. Washington.

OECD, Organization for Economic Cooperation and Development 1991. *OECD Proposed Guidelines for Collecting and Interpreting Technological Innovation Data (Oslo Manual)*. Paris: OECD.

OECD 1995. *Manual on the Measurement of Human Resources in Science and Technology (Canberra Manual)*. Paris: OECD.

OECD 1995a. *The Measurement of University R&D: Principal Problems of International Comparability*. Paris: OECD.

OECD 2000a. *Measuring R&D in the Higher Education Sector: Methods Used in the OECDEU Member Countries (Revised with 2000 Data)*, DSTI/EAS/STP/NESTI(97)2.

OECD 2000b. *R&D Statistics: Sources and Methods*. Paris: OECD.

OECD 2001. *Measuring Expenditure on Health-Related R&D*. Paris: OECD.

OECD 2002. *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development*. Paris: OECD.

OECD 2003. *Governance of Public Research. Towards Better Practices*. Paris: OECD.

OECD 2004. *The Internationalisation of Industrial R&D: Policy Issues and Measurements Problems*. Paris: OECD.

OECD 2005. *OECD Handbook on Economic Globalisation Indicators*. Paris: OECD.

Potì, B, and E Reale 2005. *Project Funding. Country report: Italy*. Paris: European Network of Indicators Producers.

Spain, 1983. xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

Teichler, U 1996. The conditions of the academic profession: an international comparative analysis of the academic profession in Western Europe, Japan and the USA. In *Inside Academia: New Challenges for the Academic Profession*, eds. P A M Maassen and F A van Vught. Enschede, Netherlands: Center for Higher Education Policies.

Van Raan, A. 2004. Measuring science. In *Handbook of Quantitative Science and Technology Research*, eds. H Moed, W Glänzel and U Schmoch, pp. 19–50. Dordrecht: Kluwer.