PART IV

Methodological issues, tools and data platform

12. Indicators for the analysis of higher education systems: some methodological reflections

Andrea Bonaccorsi, Cinzia Daraio and Benedetto Lepori

12.1 INTRODUCTION

University and higher education (HE) activity in general is a multidimensional activity based on a multi-input, multi-output relation in which inputs and outputs are not only qualitatively heterogeneous but sometimes truly incommensurable, the relation between inputs and outputs is not deterministic, the output is lagged but with a non-fixed lag structure, and the relative weight of different types of output is subject to considerable debate and political appreciation.

These methodological and conceptual issues are magnified by the well-known problem of 'data constraint' (Griliches, 1994): some of the most important problems in the economics and policy of science and HE cannot be addressed empirically due to lack of data or poor quality of data or to conceptual problems in defining and measuring suitable indicators.

Until now, what most researchers in HE have done is to choose between aggregate data at national system level provided by statistical offices, or detailed case-study data collected for single individual higher education institutions (HEIs).

An important innovation of the Aquameth project has been the collection of meso-level data – that is, data at the level of whole HEIs – on a part of the European university system (six countries) in a systematic way, by applying broad common definitions of data categories across countries and collecting information already available at national level. The same approach has subsequently been reproduced with minor modifications in the CHINC project (see Chapter 3) on a sample of 100 institutions in 10 countries.

Nevertheless, the Aquameth database which *per se* represents a very important result of the project has to be handled with care. It cannot be used in a 'data mining' way, but its exploitation needs a profound understanding of the meaning of the contained data and of their limitations, due both to conceptual problems and to the data collection procedures. This chapter deals with these kinds of issues with two major aims: to serve as a guide for those interested in further exploiting the database and to point out some major improvements in data which are urgently needed.

The chapter is organized as follows. In Section 2, we introduce the main definitions and data structure developed in the Aquameth project, as well as the data-collection strategy we followed based on national correspondents. In Section 3, we present the availability of data and discuss main quality and comparability problems; this allows us also to design data exploitation strategies which are more robust. In Section 4, we discuss how to interpret the produced indicators in different national higher education systems and taking into account the specificities of individual HEIs. Section 5 concludes with some lessons to be drawn concerning the production of institutional-level indicators and their use in a European context.

12.2 DATA CONSTRUCTION AND COLLECTION

The concept for the definition of data structure in Aquameth and CHINC, as well as for their analysis, has been based on the following main ideas (see Chapters 1 and 2 in this book):

- first, the adoption of the individual HEIs as the main level of analysis;
- second, considering individual HEIs as multi-input and multi-output organizations, where their relationship depends on organizational and production structures which are very different from institution to institution; and
- third, individual HEIs are embedded in an institutional context which determines to some extent their framework conditions (for example, the available funding sources), as well as their rules for functioning; in the European context, differences between *national* systems are particularly relevant (Kyvik, 2004).

Based on the above discussion, a very simple framework for the production of HEI indicators could then be represented in Figure 12.1. From the outset, we notice two relevant features of this framework. First, dimensions which could be, at least in principle, amenable to quantitative indicators (such as financial resources) go alongside other dimensions which are qualitative or can be operationalized with difficulty. This, of course, has important consequences for the methodology for collecting data. Second, there are complex links between the different dimensions and, especially,

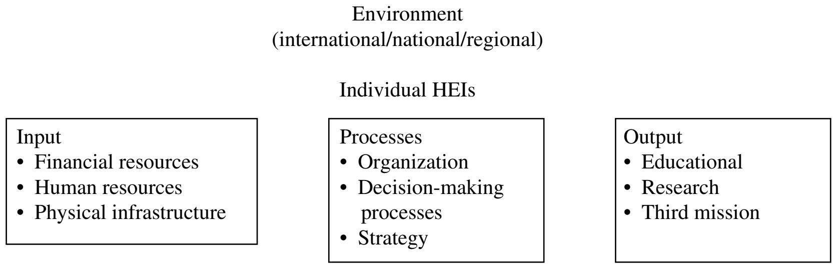

Figure 12.1 A framework for HEI indicators

between input and output variables (endogeneity): for example, at least where performance-related allocation models are in place, financial resources also depend on output. Another example is PhD students, who are, at least in some countries, at the same time an input for HEI research, as well as an educational output.

Note the richness of this approach and the emphasis on a detailed analysis of governance issues. However, it is an open question whether such a complex approach, which would require detailed case studies of each institution, is also applicable for producing broad comparisons of HEIs in different countries. For this reason, the Aquameth project decided to focus from the beginning on a subset of variables that meet the following requirements:

1. They are already available at national level, either from statistical agencies or from other official bodies collecting these data for different purposes (Conference of Rectors, national evaluation agency and so on).
2. They refer to input and output of university activity.
3. They are amenable to some quantitative measures, such as monetary indicators (for funding and expenditures) or counts of persons, publications, licences and so on.

Qualitative information is then introduced at the level of interpretation of results using the expert knowledge of the participating teams. The methodology developed by Aquameth has to a large extent been adopted in the project 'Changes in University Incomes: Their Impact on University-Based Research and Innovation' (CHINC), which was financed by the European Commission from January 2005. The main research question of CHINC is to assess changes in university funding – concerning both its absolute level and its composition – and, based on these results, to investigate the implications for research activities and outputs.

However, CHINC relied on a mix of quantitative data – being basically the same as that used in Aquameth – and qualitative information collected through questionnaires and interviews. The main methodological issue which is being addressed in the project is how to design a standardized questionnaire that also includes some complex issues such as strategies, governance mechanisms and funding allocation mechanisms in universities.

The coverage of the two projects also differs, since Aquameth includes all PhD-awarding institutions in six countries (Italy, Norway, Portugal, Spain, Switzerland and the UK), while CHINC covers a limited subset of HEIs but in a larger number of countries (also including the Netherlands, the Czech Republic, France, Denmark and Hungary) and also some non-PhD-awarding institutions. The total number of institutions in the database is 271 for Aquameth and 108 for CHINC. Since methodological issues concerning quantitative data are largely the same, we shall refer to the results of both projects. Note also that research teams in the Netherlands, France and Hungary have joined the Aquameth project in its phase 2.

Finally, contextual information on national systems is introduced at the level of analysis and interpretation of the data and drawn either from national reports (see national chapters in this book) or from existing literature.

12.2.1 Variables and Definitions

Aquameth/CHINC variables are organized in the following six broad areas:

1. general information on HEIs;
2. revenues;
3. expenditures;
4. personnel;
5. education production; and
6. research and technology production.

The choice of the indicators has been to a large extent a compromise between the theoretical model presented and the kind of questions to be addressed in the project on the one hand, and practical issues of data availability on the other.[1] In our view, the definition of a minimal core of variables which can be collected from existing sources with a reasonable effort – for example, not requiring new surveys – has been a major result of the project.

For a detailed description of the variables collected for each area, see Chapter 13. These are not statistical definitions, but rather instructions to the national correspondents on how to complete the forms starting from the information available at national level. Allowing some leeway to national correspondents has been necessary given the differences in the data availability in each country.

Note that conforming to the OECD and Eurostat practice, all financial data for the euro countries have been converted to euro fixed parity series to preserve the coherency of the series before and after the introduction of the euro. Financial data for other countries are left in the national currency. For cross-country comparisons and time series the CHINC database contains the series of purchasing power parities (PPPs) to US$ and of the national GDP deflators.

12.2.2 Coverage, Data Collection Strategy and Sources

In the collection of data, the national correspondents faced an extreme variety of situations concerning the availability of data and their quality, depending on the different organizations of the national HE systems and of the systems of production of science and technology (S&T) indicators (Esterle and Theves, 2005). While the situation differs slightly according to the considered variables, we encountered basically the following situations (see Table 12.1 for more detailed information):

- Countries where data on individual HEIs are centrally collected by national statistical services, normally for purposes of reporting and accounting. These include Norway, Switzerland and the UK.
- Countries where there is some centralized information, but not from statistical services. This is the case of Italy, where most of the data come from the publications of the Italian Conference of Rectors (CRUI) or the National Committee for the Evaluation of the University System (CNVSU), as well as Spain. Portugal used non-public information from the Ministry of Education.
- Countries where the information had to be collected directly from individual universities. These include France, Germany and Hungary. These are the most problematic cases since harmonization of the data even at national level was quite difficult; moreover, the data available and their quality differ significantly from university to university.

The analysis of the data showed some evident problems of data quality and uniformity across years. This includes evident breaks in time series for the aggregate which we would expect to be quite stable – such as the total expenditures of a university – incomplete time series and inconsistencies between data from different sources. These problems are particularly severe where data had to be collected by non-statistical sources, such as annual reports or rector conference reports. As expected, using secondary sources where statistical information is not available results in a lower quality of data. Methodologies to treat this kind of information should be developed.

Table 12.1 National sources of data per category

Country	Data sources				
	Funding	Expenses	Personnel	Education production	Research production
Czech Republic	Institute for information on education (UIV)	UIV	UIV	UIV	Publications: ISI and national database for 'Other publications' Patents: Industrial propriety office of the Czech Republic
Denmark	Directly from the institutions (no central data available) Grants and contracts (including the subcategories): Danish Centre for Studies in Research and Research Policy	Directly from the institutions (no central data available) R&D expenditures: National R&D statistics	Directly from the institutions (no central data available)	Danish Ministry of Education PhD degrees: Danish Centre for Studies in Research and Research Policy	–
France	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)
Germany	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)

Hungary	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)	Directly from the institutions (no central data available)
Italy	CRUI Publications, ministry of education sources and CNVSU	CRUI Publications, ministry of education sources and CNVSU	CRUI Publications, ministry of education sources and CNVSU	CRUI Publications, ministry of education sources and CNVSU	CRUI Publications, ministry of education sources and CNVSU
Netherlands	Several sources (ministry, statistical office, research institutions)	Several sources (ministry, statistical office, research institutions)	Several sources (ministry, statistical office, research institutions)	Several sources (ministry, statistical office, research institutions)	ISI publications: Centre for Science and Technology Studies (CWTS) Leiden Other publications: directly from the institutions Spin-offs: study for the Ministry of Economic Affairs
Norway	National database for statistics on higher education General university funds (GUF): Ministry of Education and Research	National database for statistics on higher education R&D expenditures: NIFU STEP R&D statistics	National database for statistics on higher education Share of time devoted to R&D: NIFU STEP R&D statistics	National database for Statistics on higher education Number of PhD degrees: Register of PhD degrees, NIFU STEP	Publications: ISI analysed by Studies in Innovation, Research and Education (NIFU STEP) Patents: national patent office and NIFU STEP
Portugal	DGES (Directorate on Higher Education, Ministry of Science and HE)	DGES	OCES (Observatory on Science and HE, Ministry of Science and HE)	OCES	OCES

Table 12.1 (continued)

Country	Data sources				
	Funding	Expenses	Personnel	Education production	Research production
Spain	Vice-Chancellors' Conference of the Spanish Universities (CRUE)	CRUE	National Institute of Statistics, publication Higher Education Statistics	Council of University Coordination: university Statistics	Publications: ISI Web of Science Licensing revenues, spin-offs and patents Technology transfer offices
Switzerland	Swiss Higher Education University System of the Swiss Federal Statistical Office	Swiss Higher Education University System of the Swiss Federal Statistical Office	Swiss Higher Education University System of the Swiss Federal Statistical Office	Swiss Higher Education University System of the Swiss Federal Statistical Office	ISI publications: data are for the Centre d'Études sur la Science et la Technologies (CEST)
United Kingdom	Higher Education Statistical Agency (HESA)	HESA	HESA	HESA	Publication: Research Assessment Exercise

Source: CHINC project report (Slipersæter et al., 2005) with information on Portuguese sources.

Note that the quality and coherency of data varies strongly according to their sources and data collection procedures. Thus, where statistical services are available, they usually ensure some harmonization and check the coherency of the data. Note also that the quality of available data reflects to some extent the national governance structure of higher education: for instance, the well-developed system of HE statistics in Switzerland has been created for the purposes of budgetary allocation; the same holds to a large extent for the UK. In contrast, it is practically impossible to calculate an aggregate value for total expenditures for French universities since most of the permanent staff are paid directly by the ministry; the limited budgetary autonomy of French HEIs impacts on the availability of data. In using this kind of data, we should be aware that in most cases they have been produced to manage funds and decision making and thus they are to a large extent entrenched in institutional structures and power relationships rather than in theoretical reflections on definitions of categories and variables (Godin, 2005).

Time coverage for most institutions was from 1994/1995 to 2003 with some missing years for quite a number of HEIs. Both our experience and the findings of the European Network of Indicator Producers (ENIP) confirm that in many countries a major break in the series in S&T data occurred at the beginning of the 1990s; a longer time coverage would require an expensive work of data gathering, analysis and correction.

12.3 DATA AVAILABILITY AND ISSUES

In the data collection and analysis a number of issues arose concerning the quality and comparability of the available data, as well as in some cases from their availability. We discuss these below for each main category of the Aquameth database. Further, we discuss how from these data it is possible to build indicators to answer some of the main research questions discussed in this book, which, however, are not significantly affected by data problems.

12.3.1 Expenditures and Revenues

For most HEIs it has been possible to reconstruct an aggregate value of total expenditures with the main exception of France, where until now it has been impossible to reconstruct meaningful totals for individual HEI expenditures since a large part of the permanent staff is directly paid by the ministry, which publishes only national aggregates. Clearly, the limited budgetary autonomy of French universities is also reflected in the poor availability of data. (Table 12.2.)

Table 12.2 Availability of consolidated data on revenues

Country	Govt nat.	Govt reg.	GC govt	GC inter-national	GC private	GC private non-profit	GC total	Other	TF	Remarks
Czech Republic	Yes	NR	Yes	No	No	No	Yes	No	NR	There are no data on fees or on international and private contracts except for 2004. Total revenues include all kinds of support for education and research and all other revenues of the institutions, possibly also from private sources, supplementary activities etc. In government appropriations we have stated the basic formula funding of educational activity from the state. However, there are other sources of state income, direct or indirect, and other private sources. The difference between total revenues and government national appropriations comprises all kinds of non-formula funds, but also funds for research, and possibly other revenues which do not come from governmental sources. Current accounting reports do not give us any further clues to support reasonable and comparable information
France	No	No								See comment in the text; some data are available but are not comparable for problems due to wages not taken into account in the data given by each university. The main problem is due to the scarcity of the overall database regarding France. Information is available for some of the categories, but not for all institutions or for the whole period

Germany	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	For data and contracts there is much missing data on breakdown in subcategories
Hungary	Yes	NR	No	No	No	No	No	Yes	Yes	A major problem is that all of the selected universities (except for one) were merged between 1994 and 2000), so the data are generally available only since 2000. Only one university (Semmelweis) was able to provide the data for 1994–2000. The most problematic part relates to the contracts since only two universities of seven in the sample managed to provide all the required data even though time coverage was limited
Netherlands	Yes	NR	No	No	No	No	Yes	Yes	Yes	–
Italy	Yes	NR								Some data on contracts, but coverage is not clear. Fees: undergraduates are included, but we do not know if master's fees are included
Norway	Yes	NR	Yes	Yes	Yes	NR	Yes	Yes	NR	–
Spain	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Spanish data on revenues also include a large part of investment funding, which cannot be separated from the other funds; moreover, it is also impossible to separate grants and contracts from government allocations. Thus these data are not fully comparable with the other countries
Switzerland	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	–
UK	Yes	NR	Yes	Yes	Yes	Yes	Yes	Yes	yes	–

Note: Govt nat. = government appropriations national; Govt regional = government appropriations regional; GC = grants and contracts; TF = tuition and fees; NR = not relevant (because the funding source is not important in that country).

Source: CHINC project.

However, there are a number of issues. First, the perimeter of what is considered as university expenditure might differ between countries, for example, for annex services, social security payments of personnel and the separation between higher education and healthcare expenditures in university hospitals.

Second, we had to exclude data on capital expenditures, since in some countries these are still included in state accounts and even where data are available, accounting conventions for capital expenditure – for example, the distinction between capital investment to be amortized and current expenditure – are largely different across countries, and sometimes also across universities. Also, in some countries some of the larger physical facilities (for example, buildings) may not be located at the university, but at other public or private institutions.

Furthermore, data on physical capital stock – which could be used as an input for the analysis – are almost impossible to find. Limited data on classrooms (number of seats), language and computer laboratories are available in some countries (for example, Italy). No data on experimental equipment and laboratories are available. In some countries (for example, Switzerland) there are data on floor space but cross-country comparison would require a detailed assessment of how these are calculated. For these reasons we decided to exclude capital expenditure from input data in comparative analysis. This amounts to assuming that the use of capital is homogeneous across universities and countries, which is clearly an oversimplification. In the future a dedicated study on technical coefficients should be done, including data on physical facilities.

Solving these issues goes well beyond what can be addressed in an explorative research project like Aquameth, since it would require some form of standardization of accounting and reporting procedures of HEIs at European level. Nevertheless, we devised some strategies for data analysis which are somewhat more robust in the face of these problems.

A major limitation of these data is a lack of disaggregation:[2] in most cases, expenditures can be divided only between personnel and functioning expenditures; according to the national correspondents only in Switzerland and in Norway is it possible to divide expenditures according to some discipline list compatible with the Frascati Field of Sciences classification.

In most public universities total revenues should correspond roughly to the total expenditures, since these institutions have a limited capacity to transfer funds from one year to the other and this generated a simple but useful check for data coherency; where data come from statistical offices, coherency is usually guaranteed, while in the other countries there are still some differences (especially for the Czech Republic and Hungary);

validating these data would require coming back to the original sources and looking in detail at how totals were obtained.

Quality and availability of data differ strongly according to the subcategories. All countries, except France, produced an aggregate value for general funds from the state, even if there were some borderline cases with contracts (as a detailed analysis of the Swiss data suggests). With the exception of the Czech Republic, data are also reasonable concerning tuition fees, although it should be ascertained to what extent fees from continuing education are included. In Spain, it is impossible to exclude grants and contracts from government appropriation and thus these data are not fully comparable with other countries.

The situation is more difficult for contract funding. All countries in CHINC produced totals for the 'grants and contracts' category for most institutions, but the breakdown in subcategories is possible only in Germany, Norway, Switzerland and the UK. Data on private contracts are thus available only in these four countries and, even then, quality might be problematic since, first, many private contracts are managed directly by professors and laboratories and second, accounting conventions may also differ on this issue. In the UK, for example, data include not only industry contracts but also private donations which in other countries are included in other funding. In Switzerland, a break in the series occurred in 1999 when definitions were changed by transferring subsidies from private foundations to the general budget. For public sector contracts we suggest a cross-check with the national aggregates which have been produced (at least for some countries) in the ENIP funding activity (see Lepori et al., 2005).

Finally, note that each comparison, both cross-country and across time, has to comply with the lack of deflators and PPPs specific to the HE sector, since the cost structure for the sector differs significantly from the baskets used for the general GDP deflators and PPPs (especially for a higher share of labour costs). Despite some methodological work at the OECD from the 1970s (see the discussion in the Frascati manual: OECD, 2002: 217ff.) these converters are not routinely produced by statistical offices or used for international comparisons.

12.3.2 Staff

Staff information would seem at first glance easier to get than financial data and, indeed, all universities are able to provide some estimate of the total number of staff. However, in some Aquameth countries (such as Italy and Portugal) only headcounts are available, which can be problematic since not all staff are full time (especially contract researchers). Note that the distinction between permanent and temporary staff is largely bound to the legal structure of the university rather than to their actual position and thus it is highly problematic to get data without a detailed examination of individual cases.

A rather complex issue concerns staff categories which closely depend on the organization of careers at national level: thus, positions like 'researcher' or 'research-only staff' exist only in countries where a specific career for non-teaching staff exists (Italy, Spain, the Netherlands), while in others, personnel with the same function might be labelled differently (for example, a general post-doc category including both research-only and teaching-only staff). Also, in countries with very large numbers of PhDs in natural and technical sciences, some of them could in reality perform functions similar to support and technical staff. All these differences make it difficult to use indicators based on the composition of staff.

However, perhaps the most intriguing issue comes when considering PhD students. Are they inputs of the research process, or just outputs of the educational process? Clearly both, but in unknown proportions. Moreover, the data show dramatic differences between countries concerning the number of PhD degrees and thus we might suspect that the role of PhD students in HEIs is different across countries (Table 12.3).

We propose two solutions to this issue:

1. Ex post *allocation* In some countries (for example, Switzerland, Norway) PhD students receive a separate contract for research, in addition to their grant. If this is the case, they are recorded as academic staff and their research contract is recorded in the funding data. In other countries this is not the case and we are left with the problem of

Table 12.3 PhD degrees awarded per 100 undergraduate students, 2002

Country	PhD/100 undergrad.
Czech Republic	0.54
France	0.89
Germany	1.23
Hungary	1.99
Italy	0.25
Netherlands	0.50
Norway	0.46
Spain	0.46
Switzerland	2.67
UK	0.89

Source: Jongbloed et al. (2005b).

estimating the contribution of PhD students to research. Based on expert discussion we made the decision to add 50 per cent of PhDs to the number of academic staff and hence 50 per cent of the estimated annual cost of PhD students to personnel expenditure.

2. *External variable* Another useful strategy is to let the number of PhD students vary externally to the efficiency model, exploiting the recent developments of robust nonparametric techniques. The ratio between unconditional research efficiency and research efficiency conditioned on the number of PhD students gives a precise measurement of their impact.

12.3.3 Research and Technology Production

The measure of research and technology output is clearly the most problematic domain in the Aquameth database.

The main routinely used indicator is ISI (Institute for Scientific Information) publications. The limitations of ISI data are well known, particularly in human and social sciences (see Hicks, 2004). Several research projects are trying to address this issue, and there have been attempts to use non-ISI sources on a large scale. The importance of this indicator is proportional to the importance of scientific and technical schools and departments in the university, since human and social sciences make less use of the international literature. Obviously, the former problem is larger for non-English-speaking countries. In a comparison between UK and other European countries, the implicit advantage of publishing in the native language should be considered. No formal solution was given to this issue.

We used ISI data from national sources (ministries, evaluation agencies, Conference of Rectors). After extensive discussion we came to the conclusion that data have to be handled with extreme care and have to be validated in a more accurate way. However, they represent the best available estimate of the portion of international published research, which is of the highest quality. Nevertheless, access to data has proved to be very problematic: in the UK, data are owned by a private company and thus cannot be used except at high cost; in other countries where data have been analysed by public services it was practically impossible to get the original data to perform more detailed analysis (such as in Switzerland). Clearly, urgent action is needed to ensure reasonable public access to ISI data on the scientific community.

While PhD numbers and ISI publications can be used to some extent to measure international academic production, the situation is more complex for other outputs. For technological outputs, some indicators have been developed in the literature, such as number of patents and spin-off companies (see Schmoch, 2004 for an overview), but the only standardized data at the level of individual HEIs come from an OECD survey, whose quality, according to information from some universities, is questionable since information was collected from central technology transfer offices and not cross-checked against other sources (OECD, 2003). Service activities towards private companies and the public administration are even harder to measure: a useful proxy could be the number of contracts with these users, but a suitable methodology for collecting these data has yet to be developed.

Note that the issue of measuring research and technology production is particularly severe for non-PhD-awarding institutions, which, at least in Switzerland and in the Nordic countries, are strongly orientated towards service and technology transfer to the regional economy.

12.3.4 Education Production

Although counting the number of students and degrees might seem simple, there are a number of issues to be addressed in this domain. The most important issue concerns the measure of the *quality* of education (see, for example, Salerno, 2003): this is clearly relevant since some strategies to reduce costs, like increasing the number of students in a class, presumably come at the price of a lower quality of education.

For undergraduate students, we assumed the standard ISCED (International Standard Classification of Education) 5A level definition as a starting point, that is, students who attend a course which has a duration of at least three years. Differences in the length of courses across countries in the same discipline are considered unimportant, particularly after the Bologna process, but we should consider that in the Universities of Applied Science (UAS) domain the normal duration is only three years.

Large variability exists across countries, nevertheless, concerning the precise definition of a student. In some countries, such as the UK, for example, there are many part-time students (around 22–25 per cent of the total), who clearly require fewer resources from the university. In other countries, such as Germany, it was reported that many students, in fact, do not attend the university but still apply in order to receive social welfare benefits (healthcare assistance, cheap access to accommodation and food). In Italy this was typical of the situation before the adoption of the 3+2 system. This raises the issue of time series, since changes in enrolment rules might easily modify the number of students over time.

We discussed two possible solutions:

1. *Expert judgement* If reasonably accurate estimates were available, we reduced the figures given by administrative sources by the suggested amount.
2. *Multiple variables* For all countries, we followed the strategy of specifying two separate variables: enrolments and degrees. Systematic differences in the efficiency score when the two variables are used draws attention to possible inefficiencies in the use of university services by students.

Similar problems apply to Master's students. On one hand, the duration of a Master's programme may vary (1–2 years). On the other, in some countries (for example, Portugal) students may stay enrolled in a Master's programme and for several years before graduation, and even leave the programme without a degree.

Even if the number of degrees is normally considered to be more reliable than enrolments, the introduction of the Bologna models leads to some confusion regarding what is considered to be the 'first' university degree: that is, in many continental European countries this is normally considered the Bologna Master's degree (after a 5-year course of study), while the Bachelor's degree would be considered as an intermediate step; however, this in not true for countries such as the UK and for universities of applied sciences. Since these are differences linked to national systems we allowed national correspondents some freedom to decide on what has to be considered as the 'first' university degree, but this could of course affect comparisons.

12.3.5 Construction of Indicators

Except in some specific cases, data in the CHINC/Aquameth database are not used directly as such, but rather to build indicators to answer specific research questions. Even if the two are closely linked – normally indicators are constructed from some mathematical combination of the underlying data – they differ in their meaning: data are simply measures of physical quantities collected according to some definition, for example, counting the number of students in a university or the publications in a database of an author affiliated to a university. In contrast, indicators are constructs which are supposed to measure some abstract property not directly measurable, based in general on a theoretical model, but also normally on a body of empirical research (Van Raan, 2004). For instance, ISI publications are considered to be a measure of *academic productivity* both from sociological considerations on the central role of publications in sciences and from empirical research in bibliometry, while PhD degrees per undergraduate student are usually taken as a measure of *research intensity* of a university (McCormick, 2004). Note that, unlike data, indicators require normalization to have a meaning: for example, a level of 1 PhD degree per 100 undergraduate students is usually considered to be the threshold for research-intensive universities. Of course, this normalization can be quite different for each specific context.

Good indicators should share two main features: first, they should be well founded in theoretical terms, meaning that there should be some underlying explanation for the assumption that they correctly represent a given feature of reality. For instance, the use of publication numbers is well rooted in their role in academic production. In other cases, the foundation has to be sought in some practical evidence: for example, it is known that most of the grants and contracts attributed to a university are used for research purposes and thus their share in total revenues is currently used as a measure of *research intensity*; if we believe that grants are attributed on the basis of the quality of research, this indicator can also be used as a measure of the *research quality* of an institution.

Second, good indicators should be robust against limitations in the underlying data. If we accept that in science and technology available data have quite a number of limitations concerning quality, degree of detail and comparability, we must have some confidence that the results are sufficiently robust against data limitations. An example of this approach is the indicators developed in the chapter on higher education funding (see Chapter 3) where we avoid comparing absolute cost levels per student between institutions since we know that these are affected by differences in accounting systems between countries, and by problems with the use of PPP for international comparisons and, finally, strongly depend on the subject mix of the institution. On the contrary, we expect the evolution of the cost level over two years for a single institution to be more robust since some of these problems should affect both the numerator and the denominator in a similar way.

The list of indicators presented in Table 12.4 is of course not complete, as the dynamics of university functioning is much more complex. This list instead represents the main indicators that we used in our first exploitation of the integrated database to try to characterize the positioning of European universities in their multidimensional space of activity/resources/personnel-based decision making. Following the preceding discussion, we built these indicators according to two main criteria:

- their ability to characterize or at least outline an important aspect of the university activity/profile based on our theoretical understanding of the economics of higher education; and

Table 12.4 Indicators used in the Aquameth project

Domain	Indicator
Expenditures	% capital expenditure over total expenditure % personnel expenditure over total expenditure % current expenditure over total expenditure
Funding	% of government allocations over total revenues % of tuition fees over total revenues % of grants and contracts over total revenues
Staff	Academic staff per undergraduate student % of non-academic staff over total staff
Educational output	PhD degrees per 100 undergraduate students
Research output	ISI publications per academic staff

- their robustness against the major data problems discussed in dection 3.

However, this list and the analysis developed with these indicators in other chapters of the book have to be considered as a contribution of the Aquameth project towards the definition of a basic set of indicators to characterize the structure and the dynamics of European HEIs.

12.4 COMPARABILITY ISSUES AND INTERPRETATION STRATEGIES

In the previous section we illustrated the main data problems and the 'operative' strategy we followed to overcome these limitations. Nevertheless, data problems are only a part of the more general comparability issues, arising from differences among national systems, individual HEIs and so on.

Some comparability problems are embedded in data availability: according to the structure of national systems we get different data. Other problems are located at the level of the individual university and deal with its 'ontology' as an object of analysis. In other cases they rely on the interpretation of data: the same number may tell a different story according to the national context or the type of HEI.

These cases highlight the need to 'contextualize' data and available indicators in their institutional and national context, on the one hand; and to take into account these specificities in the statistic and econometric analysis on the other. Bonaccorsi, Daraio and Simar (Chapter 5) show that by using recently introduced robust nonparametric methods with their related graphical tools (Cazals et al., 2002; Daraio and Simar, 2005a,b, 2006) some national specificities can be caught, without imposing at the beginning a strong formalization of the relationships among the inputs–outputs–external factors as required by traditional (parametric) econometric techniques. They propose these new nonparametric and robust techniques as exploratory tools to detect a first approximation of important phenomena in research and HE policy such as teaching versus research trade-offs; the role of PhDs in research activity; and so on. Obviously, the more we know about the comparability issues, the more we shall be able to understand and interpret the observed inconsistencies in the data. Most of them come down to the general question of the legitimacy of comparing apples with oranges. There is no general answer to this issue. We have to combine careful and detailed knowledge with the appropriate level of abstraction in identifying useful categories and statistical and econometric solutions.

For analytical purposes, it is useful to distinguish between two main sources of comparability problems, which are, however, closely connected.

- First, differences in the *organization and governance structure* of national HE systems which also have an impact on the indicators produced from the Aquameth database. As a consequence of these differences, the same indicators might have a quite different meaning according to the country where the individual HEI is located. For instance, the funding structure of HEIs is to a large extent determined at the national level and, thus, to compare the funding structure of two universities in different countries one has to take into account this effect. Other examples are the existence of a 'second' HE sector (such as the *Fachhochschulen* in Germany) and the relationship between universities and public research organizations (PROs) in the public research sector.
- Second, *heterogeneity* of individual HEIs. While some distinctions – between public and private HEIs and between PhD- and non-PhD-awarding institutions – are rather clear-cut and linked to structural features of the national system, the most intriguing issue is differences in the subject mix between 'traditional' universities. Subject mix is clearly an issue since, according to some results, cost structures are very different according to the domain; the same is probably true for research intensity, even if it is more difficult to define indicators. Also, bibliometric indicators from the ISI clearly tell different stories according to the discipline considered and thus aggregates for a university cannot be compared without taking into account the subject mix.

12.4.1 Institutional Context

An important source of heterogeneity across universities comes from institutional differences across European countries. There is a large body of literature on the system-level governance of HE systems and on its changes during the last 2–3 decades, starting with the scheme proposed by Clark to classify national systems according to the 'coordination mechanism' (Clark, 1983); most analysts have opposed the model prevailing in Continental Europe, where universities are considered as state institutes deserving a public function and subject to strong state control, to the more liberal and market-orientated model in Anglo-Saxon countries (Amaral et al., 2002). Also, it is well known that both models have undergone a profound revision in the last 2–3 decades, characterized as the shift from a state control to a state supervising model (Neave and van Vught, 1994).

Note that, at least in federal countries, the institutional context can be quite heterogeneous even at national level, as the Swiss case clearly demonstrates (see Lepori, Chapter 6). With the increasing role of the regions in the European Research Area (Larédo, 2003), the issue of the impact of regional differences – concerning levels of economic development, industrial structure and support measures – is increasingly important.

12.4.2 Dual Systems

Some European countries have a system in which the HE system includes universities but also a range of non-PhD-awarding institutions, such as *Fachhochschulen* in Germany, *Hogescholen* in the Netherlands and universities of applied sciences in Switzerland (Huisman and Kaiser, 2001). In most cases, these HEIs differ clearly from universities concerning their organization, education and research output.

The appropriate strategy here is to carry out all analyses separately and to specify models differently. In particular, the research output should not be included in models for these institutions. However, the interactions between the two sectors should be considered, especially concerning education, since the existence of these 'second-type' institutions tends to reduce undergraduate student numbers in universities. Moreover, in countries such as Switzerland and Finland, these institutions are relevant players in 'third-mission' activities in the private economy.

12.4.3 Private versus Public Universities

In the European context the public nature of universities is dominant. Nevertheless a number of private institutions still exist, usually recognized by the ministry of education and sometimes partially funded by the government.

The classical solution of a dummy variable is appropriate here. It allows the variability in efficiency across the two categories, as well as the variability internal to the categories to be estimated. Another possibility may be to carry out two separate analyses, one for private and the other for public universities. Note also that data are not fully comparable between the two categories owing to different legal status (for example, commercial accounting in private universities) and different requirements for collecting them.

12.4.4 Public Research Organizations

It is well known that scientific research can be carried out not only by universities, but also by large PROs, such as CNRS (centre national de la recherche scientifique) in France, Max Planck in Germany, CNR (Consiglio Nazionale delle Ricerche) in Italy and CSIC (Consejo Superior de Investigacione Cientificas) in Spain.

In some cases, for example in France, the historical tradition has been one of concentration of most research activity at the level of CNRS, INSERM (for biomedical science) or INRA (for agricultural research), while universities had mainly teaching responsibilities, at least until the large reform that promoted strong institutional collaboration between universities and PROs.

These differences may create estimation problems in all cases where personnel from PROs also work at university level and publish jointly with university researchers, but are not recorded as an input. Similarly, if university researchers obtain access to laboratories at PROs and, due to high-level institutional agreements, do not have to pay for access, then this input is not recorded in terms of funding.

Overestimation of efficiency may result. This issue is particularly relevant for countries such as France and to some extent, Portugal, where mixed units are widespread and, hence, it is difficult to clearly identify the perimeter of higher education.

There is no easy solution to this issue. In principle, one should control for co-authorship patterns in publications and compute a coefficient of joint input from PROs, assuming that the proportion of authors is representative of the proportion of all inputs contributed. More realistically, we might allocate a share of PRO personnel to any individual university, discipline by discipline, according to expert judgement of the actual contribution. For example, if we know that a large INSERM institute is collaborating with a medical school in France, we might want to include in university academic staff a share of INSERM personnel.

More generally, any HEI analysis should take into account the relative role of universities and PROs in the production of research ouput, as well as the interaction between these actors in terms of joint production of research output.

12.4.5 Age and Structure of Universities

In other cases there might be large heterogeneity according to the age of universities, if there has been a discontinuity in the history of higher education. An interesting case is the UK. The definition of a university currently includes the so-called 'new universities'. Originally known as 'polytechnics', they did not carry out research but only technical and professional training. With a national law in 1992, they were transformed into universities, providing research funding.

A solution to this issue is to include the age of the university as a descriptor and to consider it in the estimation. Another solution is to introduce a break in the sample and estimate efficiency separately, once the qualitative analysis has shown large enough differences.

12.4.6 Funding Pattern

One might also consider that large heterogeneity is introduced by institutional differences in the way universities are funded. Thus national funding systems are quite different across Europe concerning the level of funds and their composition; for example, concerning the share of third-party funds versus the share of general government allocations (see Chapter 3).

We have dealt with this source of heterogeneity by computing the share of funding coming from different sources. These composition rates can also be used as external variables in robust nonparametric techniques, in order to understand whether patterns of funding really matter. For example, Bonaccorsi et al. (2006) estimate the relative impact on efficiency of Italian universities of the share of funding coming from private sources.

12.4.7 Heterogeneity of Individual HEIs

Another fundamental issue in the analysis of HEIs is the level of heterogeneity. This is the result of differences in the institutional system, as well as the outcome of historical developments and strategic decisions. Disentangling these effects is rather difficult.

12.4.8 Subject Mix

Universities may have very different profiles in terms of faculties and schools covered, and hence of educational activity and research output.

From the teaching point of view, strong differences in cost per student are likely, due to differences in capital intensity (laboratories), length of courses and type of training (applied, practical experience). This is confirmed by a test performed in the Swiss case, where we could show that resources per undergraduate differ by a factor of about 10 between human and social sciences and medicine (Filippini and Lepori, Chapter 8). Also, from the research point of view, it is well known that disciplines exhibit diverse publication patterns, in terms of average number of co-authors per paper, average number of per capita papers per year, and the like. Ignoring these differences may be misleading.

Tackling the subject mix issues is complicated by the lack of data disaggregated by discipline for most inputs and outputs. In Aquameth it was possible to obtain disaggregated data concerning student numbers at least for the main domains of the OECD field of sciences (FOS) classification (OECD, 2002: 67). Concerning staff, disaggregation would in principle be possible for some countries, but national schemes do not always comply with the FOS, while only in Switzerland can expenditures be divided by field of sciences. Similar problems exist for scientific publications.

In Aquameth, we discussed several possible strategies to address this problem:

1. *Dummy variables* As an example, we have constructed a dummy valued 0 if the university does not have medical schools; 1 if the university has a medical school (but not a hospital), and 2 if the university maintains both a medical school and a hospital. This is particularly relevant since faculties of medicine can account for half of the total expenditure of a universities and the separation of costs between higher education and healthcare is highly problematic (OECD, 2001).
2. *Categorization* We propose to build a concentration index by computing the distribution of students in four broad disciplinary areas (human and social sciences; technical sciences; natural sciences; and medicine). Following a standard notation in economics, C_1 is the concentration index for the largest discipline, C_2 for the first two and so on. We define as 'specialist' a university with $C_1 \geq 0.70$ or $C_2 \geq 0.90$ and 'generalist' otherwise. Of course, other specifications can be explored. Once a categorization is accepted, analyses can be carried out separately. As an example, Bonaccorsi, Daraio and Simar (Chapter 5), have used this categorization to compare the productivity of research across

European countries by analysing only 'generalist' universities according to this definition.

3. *External variables* In some cases it is not well understood whether heterogeneity in the subject mix really matters for the analysis. A suitable technique is made possible by conditional robust nonparametric techniques (Daraio and Simar, 2005a,b, 2006). These allow the estimation of the variability of efficiency score when an external variable is included in the model. Technically speaking, the ratio between conditional and unconditional efficiency is computed. For example, one might be interested in seeing the variation of efficiency in teaching when the number of medical students is included as an external variable. In this way a number of controls can be done. Since the variables should be continuous, they can be transformed appropriately (for example, the discrete variable 'age' can be made continuous without loss of precision).
4. *Test of hypotheses* Once the external variable has demonstrated that efficiency is strongly influenced by heterogeneity, it would be possible to split the sample according to the identified variable and perform the analysis separately.
5. *Multi-layer models* Still another possibility is to apply nested models, that is, start with small samples and include them in larger ones by weighing them virtually (Sarrico and Dyson, 2004).

12.5 TOWARDS 'MESO' DATA ON THE EUROPEAN HIGHER EDUCATION SYSTEM

In this section we should like to draw some general conclusions on the outcome of the approach chosen in Aquameth (and then also developed in CHINC; see Slipersæter et al., 2005) for data collection and analysis, to point out some issues needing urgent improvements and then to propose some ideas for a European platform of indicators on HEIs.

First, we think that the chosen approach has led to real progress: despite all the limitations discussed above, the collected data proved to be *usable* not only for national analysis, but also for comparative analysis across countries concerning efficiency (Bonaccorsi, Daraio and Simar, Chapter 5) and changing funding models (Lepori et al., Chapter 3).

In our view, this positive result is due to a combination of three main elements:

- The definition of a *minimum core of data* which can realistically be collected from existing sources, without posing unsolvable methodological problems. This strategy led, for instance, to the exclusion of data on investments or capital costs.

- The project organization based on *national correspondents* in charge of retrieving and collecting the data based on their knowledge of data sources and of the national HE system. This allowed the use of unconventional sources for countries without a well-developed HE statistics, such as Italy and Spain. Correspondents' expert knowledge was also essential for data handling and interpretation of results.
- A *careful usage strategy* of the collected data which is aware of their limitations and tries to overcome them with suitable techniques.

Second, this positive result should not overshadow the limitations of the collected data and the need for improvements. Some of them, as already discussed, simply reflect the heterogeneity of national HE systems, which is a major difference between Europe and the USA, as well as of individual HEIs. These issues call for a further interpretation effort systematically linking quantitative data with contextual information on individual countries and institutions. However, some limitations lead more directly to issues of data quality and availability. First, despite our efforts, we are aware that to ensure the quality of data more systematic validation procedures should be introduced. Second, in many cases we still lack a clear understanding of the level of comparability of the data. For instance, on some of the issues raised in this chapter – such as differences in accounting systems or in definitions of students and degrees – we do not know whether they simply produce noise in the observed patterns or whether they alter them fundamentally. Finally, there are some domains where little progress can be made without fundamental advances in the methodology and the practice of data collection: this is the case for capital costs and for most output indicators, including scientific publications and third-mission indicators.

Summing up, we claim that in a long-term perspective, the maintenance of a dataset like that developed in Aquameth and CHINC goes well beyond a research project, especially if the intention is to widen its reach to most of the European HEIs (that is, to about 3000 institutions). At the same time, the approach chosen here is a long way from the one adopted by statistical agencies, since it is more centred on the production of indicators from existing data than on the collection of coherent datasets in statistical terms; thus, it largely exploits the work of these organizations at national and international levels (Eurostat, OECD), but it adds a further layer of complexity in three directions: first, it takes explicitly into account the heterogeneity of national systems and of individual institutions rather than trying to harmonize statistics; second, it focuses on individual HEIs rather than on national systems as the main reference level; third, it covers a wider range of indicators concerning issues like third mission and output of HEIs, in the future regional indicators, which are not adequately covered by official statistics, but are clearly needed both for research in HE and for policy analysis.

Finally, the need for a European Science and Technology Indicators Platform (ESTIP) fulfilling these requirements has been convincingly demonstrated by another PRIME project, the European Network of Indicator Producers (ENIP) (Esterle and Theves, 2005) and some concepts in this direction are currently under discussion (Barré, 2005). In this respect, the Aquameth project has brought a number of interesting results and indications.

NOTES

1. During the Aquameth 1 project data on patents and technology indicators for many countries were not available. The collection of these data is under way in Aquameth 2.
2. A related problem is the allocation of funds to teaching and research which is not possible in some countries. In consequence, the allocation of contract income to research is not realistic for all countries. In some countries a large share of base funds is used not only for teaching but also for research.

REFERENCES

Barré, R. (2005), 'Towards a European STI indicators platform (ESTIP)', Position Paper, paper presented at the Second annual PRIME conference, Paris, February 2006.

Bonaccorsi, A., C. Daraio and L. Simar (2006), 'Size, scope and trade-off in the productivity of universities: an application of robust nonparametric methods to Italian data', *Scientometrics*, **66** (2), 389–410.

Cazals, C., J.P. Florens and L. Simar (2002), 'Nonparametric frontier estimation: a robust approach', *Journal of Econometrics*, **106**, 1–25.

Clark, B.R. (1983), *The Higher Education System: Academic Organization in Cross-national Perspective*, Berkeley, CA: University of California Press.

Daraio, C. and L. Simar (2005a), 'Introducing environmental variables in nonparametric frontier models: a probabilistic approach', *Journal of Productivity Analysis*, **24** (1), 93–121.

Daraio, C. and L. Simar (2005b), 'Conditional nonparametric frontier models for convex and non convex technologies: a unifying approach', Discussion Paper no. 0502, Institut de Statistique, Université Catholique de Louvain, Louvain la Neuve (UCL), Belgium, forthcoming in *Journal of Productivity Analysis*.

Daraio, C. and L. Simar (2006), *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*, New York: Springer.

Esterle, L. and J. Theves (2005), 'Analysis of the different European systems for producing indicators', paper presented at the Lisbon Workshop on S&T Indicators Production, Lisbon, 22–23 September.

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

Griliches, Z. (1994), 'Productivity, R&D, and the data constraint', *American Economic Review*, **84** (1), March, 1–23.

Hicks, D. (2004), 'The four literatures of social science', in Moed et al. (eds), pp. 473–96.

Huisman, J. and F. Kaiser (eds) (2001), *Fixed and Fuzzy Boundaries in Higher Education: A Comparative Study of (Binary) Structures in Nine Countries*, The Hague: Adviesraad voor het Wetenschapsen.

Jongbloed, B., B. Lepori, C. Salerno and S. Slipersæter (2005b), 'European Higher Education Institutions: Building a Typology of Research', CHINC report, IPTS, Seville.

Kyvik, S. (2004), 'Structural changes in higher education systems in Western Europe', *Higher Education in Europe*, **29** (3), 393–409.

Larédo, Ph. (2003), 'Six challenges facing the future of European science and technology policy', *Science and Public Policy*, **30** (1), 4–12.

Lepori, B., M. Dinges, B. Poti and E. Reale (2005), 'Public project funding of research activities. A comparative analysis', paper presented at the Lisbon Workshop on S&T Indicators Production, Lisbon 22–23 September.

McCormick, A.C. (2004), 'The 2005 revision of the Carnegie classification system', Presentation to the Washington Higher Education Secretariat, Washington, DC 8 June, www.carnegiefoundation.org/Classification/future.htm.

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

Neave, G. and F.A. van Vught (1994), *Government and Higher Education Relationships Across Three Continents*, Oxford: Pergamon.

Organization for Economic Cooperation and Development (OECD) (2001), *Measuring Expenditures on Health-related R&D*, Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2002), *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development*, Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2003), Turning Science into Business: Patenting and Licensing at Public Research Organisations, Paris: OECD.

Salerno, C. (2003), 'What we know about the efficiency of higher education institutions: The best evidence', Vol. 99, The Hague: Ministry of Education, Culture and Science.

Sarrico, C.S. and R.G. Dyson (2004), 'Restricting virtual weights in data envelopment analysis', *European Journal of Operational Research*, **159**, 17–34.

Schmoch, U. (2004), 'The Technological Output of Scientific Institutions', in Moed et al. (eds), pp. 717–32.

Slipersæter, S., B. Lepori, B. Jongbloed and C. Salerno (2005), 'Collecting Institutional Level Data for European Higher Education Institutions: Evidence from the CHINC Project', CHINC report, IPTS, Seville.

Van Raan, A. (2004), 'Measuring science', in Moed et al. (eds), pp. 19–50.