

# Patterns of Subject Mix in Higher Education Institutions: A First Empirical Analysis Using the AQUAMETH Database

Benedetto Lepori · Lukas Baschung · Carole Probst

Published online: 24 March 2010  
© Springer Science+Business Media B.V. 2010

**Abstract** Teaching and research are organised differently between subject domains: attempts to construct typologies of higher education institutions, however, often do not include quantitative indicators concerning subject mix which would allow systematic comparisons of large numbers of higher education institutions among different countries, as the availability of data for such indicators is limited. In this paper, we present an exploratory approach for the construction of such indicators. The database constructed in the AQUAMETH project, which includes also data disaggregated at the disciplinary level, is explored with the aim of understanding patterns of subject mix. For six European countries, an exploratory and descriptive analysis of staff composition divided in four large domains (medical sciences, engineering and technology, natural sciences and social sciences and humanities) is performed, which leads to a classification distinguishing between specialist and generalist institutions. Among the latter, a further distinction is made based on the presence or absence of a medical department. Preliminary exploration of this classification and its comparison with other indicators show the influence of long term dynamics on the subject mix of individual higher education institutions, but also underline disciplinary differences, for example regarding student to staff ratios, as well as national patterns, for example regarding the number of PhD degrees per 100 undergraduate students. Despite its many limitations, this exploratory approach allows defining a classification of higher education institutions that

---

B. Lepori (✉) · C. Probst  
Centre for Organisational Research, Faculty of Economics, University of Lugano, Lugano,  
Switzerland  
e-mail: benedetto.lepori@usi.ch

C. Probst  
e-mail: carole.probst@usi.ch

L. Baschung  
Observatoire Science, Politique et Société, University of Lausanne, Lausanne, Switzerland

accounts for a large share of differences between the analysed higher education institutions.

**Keywords** Subject mix · Classification of higher education institutions · Generalist vs. specialist · Indicators · Staff composition

## Introduction

The objective of this paper is to provide some empirical evidence concerning differences in subject composition between Higher Education Institutions (HEI), as well as the impact of these differences on the construction of institutional-level indicators, for example for mapping higher education systems or for measuring the research output.

Differences between scientific disciplines and specialities concerning the organisation of research and educational work are a well known issue in higher education studies (Clark 1983; Becher and Trowler 2001), which has been mostly addressed through qualitative approaches comparing disciplines across institutions and national systems. Internal differentiation of scientific domains is also considered a major driving force in the evolution of higher education overall and of increasing diversity between HEIs (Clark 1995; Meek et al. 1996).

It is also well known that individual HEIs strongly differ in the distribution of subject domains and that a number of very specialised institutions can be identified, like technical universities or business schools. However, there is no systematic account of these differences and of the role of specialised vs. generalist institutions in national systems; despite claims that subject differentiation is a widespread phenomenon and a relevant dimension of differentiation in higher education systems in general, almost all current mapping exercises do not include categories taking into account subject composition, like some kind of distinction between “generalist” and “specialist” institutions; for example, recent work on a typology of European HEIs includes as a dimension to be considered the range of subjects offered in education, but no indicator of the importance of subject domains in HEIs’ activities in general is proposed (van Vught et al. 2008).

To some extent, this situation can be explained by the separation between two traditions in higher education studies, one focusing on the organisation of the academic work with a sociological rooting and a focus on scientific disciplines (for example represented by Becher and Trowler 2001), and the other viewing higher education institutions as strategic units, mostly from an economic perspective (for example, represented by Bonaccorsi and Daraio 2007b). To a large extent, this situation is, however, also due to the lack of quantitative indicators allowing systematic comparisons of subject mix across large numbers of institutions in different countries (Bonaccorsi et al. 2007). Yet, these indicators are needed to go beyond individual case studies towards a broader analysis considering the interplay between higher education governance, strategies of individual institutions and the development of subject domains. Moreover, they would help to overcome a major limitation of today’s institutional-level indicators, since one cannot assess to what

extent differences between institutions are driven by their different subject composition (as shown by some national-level studies).

This paper represents an exploratory effort in this direction. We exploit the database developed in the AQUAMETH project (Bonaccorsi and Daraio 2007b) to characterise the subject composition of PhD-awarding universities in six European countries, and to build classes of HEIs with similar subject composition, using staff effort as a proxy of their effort by field. Secondly, we examine the characteristics of these classes with regard to size, age, national distribution, etc. Thirdly, we compare these results with measures of subject composition derived from product mix in education (using student repartition by subjects) and in research (using the repartition by subject of PhD students and of scientific publications). Finally, in the last section we draw methodological conclusions on the construction of HEI-level indicators including some consideration of different subject mixes.

### **From Disciplinary Diversity to Subject Mix**

The organisation in scientific disciplines and specialities with different cognitive and social structures is one of the most distinctive characteristics of higher education; while social studies of science have extensively dealt with the characteristics and dynamics of scientific communities, higher education studies have also analysed the complex relationships between the organisational level of higher education institutions and the socio-cognitive level of disciplines (Becher and Trowler 2001); HEIs have thus been characterised as loosely coupled organisations with extensive work division between disciplines and limited interdependencies and economies of scope between fields (Musselin 2007), while double affiliation to disciplinary communities and universities is considered as a key feature of the academic profession (Clark 1983).

In organisational terms, scientific disciplines constitute the basic building blocks of most higher education institutions; comparative studies reveal that most of them display two basic organisational levels (Clark 1983), one corresponding to broad groupings of knowledge areas preparing to specific occupations—*faculties, schools, colleges*—and one corresponding to narrower groupings around a professional specialty or a subject discipline—*chairs, institutes, departments*.

Starting from the medieval quadrivium, European universities developed in the nineteenth century a differentiated organisation in faculties, adopting, however, specific national groupings (Rothblatt and Wittrock 1993; Rüeegg 2004). While this basic organisation has been largely maintained until today, an increasing variety of subdisciplines and specialities has been documented as a response to internal differentiation of scientific work, as well as to differentiation of social demands for higher education (Becher and Trowler 2001); this has led to the multiplication of organisational structures at the level of subunits (research centres and subjects of curricula), but also to the creation of new faculties like educational sciences or environmental sciences, beyond the historical disciplinary structure. In this perspective, disciplinary differentiation is considered as a driving force for the

evolution of the higher education system in general and of the organisation of individual institutions (Clark 1995).

However, while the differentiation process of scientific disciplines and internal differentiation in individual HEIs have been widely studied, differences between institutions concerning the domains and specialities, which are present, and their relative importance have been much less investigated. Most work concerned with the differentiation in the HEI population has focused on vertical differentiation (concerning, for example, quality of research) and differentiation concerning missions (Bleiklie 2003; Huisman et al. 2007), for example, between education and research, but largely neglected the question to what extent differentiation takes place alongside the subject dimension, with individual HEIs selectively focusing on specific subject domains depending also on the opportunities and resources available in each field (see, however, Rossi 2009 for the Italian case).

Yet, there are good reasons to consider that this question is highly relevant if we want to understand the construction of individual profiles of HEIs, their diversity and differentiation processes, and their impact on system's performance.

Firstly, there is evidence that available resources and opportunities for education and research are largely subject-specific; for example, third-party funding for research is concentrated in natural and technical sciences, private funds are highly focused on specific domains (business studies, technology), while students' growth in most countries is concentrated in social sciences. Accordingly, HEIs with different subject compositions will be faced with different environmental conditions and opportunities, and thus the interaction between institutions and environment is likely to be largely subject-specific.

Second, sociological studies, but also economic analyses of higher education reveal that differences between subjects concerning internal organisation and the production process of research and education are very large and display systematic patterns between countries and institutions. Thus, differences in costs per student between subject domains are systematic, with medicine at the top followed by natural sciences and technology, while social sciences and humanities have lower average costs (Jongbloed et al. 2003); in the case of UK universities, 70% of differences in unit costs between HEIs can be explained by subject mix (Johnes 1990). Using disaggregated data at the field level in the Swiss case, Filippini and Lepori (2007) show that differences between domains are systematically larger than differences between individual HEIs, and this pattern is consistent across a wide range of indicators, including students per professor, educational and total costs, number of PhD students and degrees. The few available econometric studies also display that disciplinary composition is a major explaining factor of productivity of individual HEIs and that institutional comparisons need to control for it (Sarrico et al. 2009). When benchmarking individual institutions, due consideration of different underlying production structures is thus required.

Third, existing studies show that prioritising and allocation of resources between subject domains are a central issue for higher education strategies and internal allocation of resources (Fumasoli and Lepori 2010), and these processes are largely driven by the perception of HEI management of the strengths and weaknesses of specific domains, of environmental opportunities, as well as by the power of

different departments inside the organisation (Salancik and Pfeffer 1974). The scarce empirical evidence of the interplay between institutional-level strategies and disciplinary units displays complex patterns of interaction which are likely to be highly dependent on the specific institution and national context (Meek et al. 2000; Morphew 2000).

These remarks led some scholars in higher education to consider that disciplines are the relevant unit of analysis, which should be compared across institutions and countries, and that it is their dynamics that drive the differentiation of the system and the construction of individual profiles at the institutional level through the power of departments and faculties (influenced also by differentials in the availability of resources; Meek et al. 1996). Other scholars consider instead that, both, public policies granting more autonomy to individual HEIs and the strengthening of central management increasingly transform HEIs into strategic units which are able to manage their portfolio of inputs and outputs and to actively define their profile (Bonaccorsi and Daraio 2007a). Moreover, in this perspective, institutional-level processes like isomorphism and competition profoundly shape HEIs' profiles and system's configuration beyond subject differences (van Vught 2007). Thus, analysing the subject composition of higher education institutions and the forces shaping it comes back to a long-standing debate on the drivers of the differentiation process of higher education systems in general (Meek et al. 1996). Theoretical reflections and existing empirical studies provide support for both positions, but there is a distinct lack of studies that consider, at the same time, subject diversity and institutional diversity and look at their interplay.

It goes beyond the scope of this paper to provide answers to these questions, which would require extensive comparative work that takes into account changes of subject profiles over time and looks at their drivers. Our more modest objective is to provide some preliminary evidence on subject composition of HEIs in European countries, to analyse national differences in this respect, and to identify specific classes of institutions with regard to subject composition—for example, concerning the role of specialised institutions in higher education systems. Moreover, on the methodological side, we aim to provide some insight on the extent to which different measures of subject composition might yield different results in terms of specialisation patterns and of comparisons between institutions.

## **Constructing Classification Schemes**

In general terms, subject mix can be defined as the repartition of activities of a higher education institution between its subject domains. Comparing HEIs for subject composition raises the issue of which activities to use as benchmarks and how to construct and operationalise classifications; this requires also to consider variations across space and time and to choose the right level of granularity in the adopted classification. As a matter of fact, the available options are strongly limited by the availability of data. However, an in-depth discussion of methodological issues is relevant to understand the limitations of the results presented in this paper and to look for potential improvements.

As shown by sociology of science, subject domains are complex social, cognitive and organisational constructions which should be analysed through suitable network techniques like social networks and cognitive structures. While this approach is well suited for studies of individual disciplines, it is impractical for system-level analyses, also because it does not provide any basis for institutional comparisons; moreover, this approach is suitable for investigating scientific work, but can be hardly applied to examine other activities like education and transfer.

Resorting to organisational units—for example faculties—is a simpler solution when these units are relatively homogenous concerning their domain of activity and there is some expectation that a unit is active in the same domain for most of its activities. This assumption can be reasonably maintained for broad domains, but it is questionable for fine-grained analyses since specialisations in research and in education are likely to be very different and matrix organisational schemes—with faculties responsible for education and departments for research—are becoming more widespread.

The approach of mapping organisational units to a general classification scheme of scientific disciplines is the one recommended by the Frascati manual for R&D statistics (OECD 2002), where a broad classification in six Fields Of Science (FOS), as well as a finer-level classification by more specific subfields, is introduced. The manual states that the statistical unit for attribution should be the smallest homogeneous unit for its activities and thus requires, for example, splitting up interdisciplinary faculties. Concretely, this can be done most easily for staff and, with more problems concerning data quality, for expenditures (Bonaccorsi et al. 2007), while attribution to organisational units of outputs is a cumbersome process for research, since it requires matching publications data with institutional affiliations. In AQUAMETH, a simplified version of FOS considering four domains, namely medical sciences, engineering and technology (including agriculture), natural sciences and social sciences and humanities has been adopted.

An alternative option is to consider HEIs as institutions which transform inputs—staff, infrastructure, new students—in different types of outputs, including degrees, research products, technology and knowledge transfer and to classify outputs by subject, disregarding the organisational units from which they are produced. This approach can be adopted for educational products—students and degrees—based on the classification of educational programs by fields of education provided in the handbook of educational statistics (OECD 2004), as well as for scientific publications using journal-based subject classification as the subject domains adopted by Web of Science. However, while for students and degrees subject profiles can rather easily be constructed using data from educational statistics, there are well known issues in constructing research output profiles by subject since the most widely used database for this purposes displays very uneven coverage by subject (Nederhof 2006).

The first approach therefore characterises subject mix in terms of the repartition of internal activities—as, for example, measured by repartition of staff—and provides information on how the considered institution allocates resources to subject domains, while the second one provides information on the realised product mix by subject and is therefore more relevant when analysing competition between institutions. In this paper we use both of them; thus, in section “[Constructing](#)

**Classification Schemes**” we build a classification based on the repartition of staff between domains, while in the following sections **“Data Sources, Limitations and Comparability Problems”** and **“Identifying Institution Classes through Staff”** we compare this classification with the repartition of students from one side and of research products from the other, and inquire to what extent these display systematic differences.

A number of limitations of the measures adopted in this paper are related to the problem of applying general classification schemes to different and contingent realities. Firstly, using a classification of scientific fields to four or six domains blends the increasing level of internal differentiation between disciplines documented by higher education studies (Clark 1995; Becher and Trowler 2001); this is relevant since much of the diversity in subject composition between higher education institutions takes place at the specialty level and thus, for instance, two institutions specialised in arts and in theology are very different with regard to their subject specialisation, even if both were considered as specialised institutions in social sciences and humanities. Second, given the dynamics of scientific disciplines, it is somewhat questionable to what extent classification schemes developed in the 1960s like FOS still adequately represent current patterns of subject specialisation and, especially, the emergence of new specialities across the traditional disciplinary borders (attempts by the OECD to revise FOS, however, have merely led to minor adaptations). Thirdly, the application of uniform classification schemes to different national situations and types of institutions can lead to comparability problems, especially in domains like social sciences and humanities where organisation of science is largely national; furthermore, activity fields of vocationally-oriented HEIs largely based on professional categories (e.g. health-related professions) are not easily mapped into classifications built for academic institutions.

To a large extent, these problems cannot be avoided when developing large-scale comparisons since these can be done only by using existing statistical data; they point, first, to the need of great care in interpreting the results of this analysis and, second, to the fact that the approach based on classification schemes can be safely used only for broad subject domains, but not pushed to further disaggregation; finally, they point to the need of complementing the analysis with more in-depth investigations of subject mix using detailed data from the institutions themselves, but forcefully limited to a much smaller set of institutions.

A final issue concerns the choice of the benchmark against which to compare subject mix. A first option is to use absolute concentration indexes, thus considering all institutions where the share of a single domain exceeds a threshold (75% of students in Bonaccorsi and Daraio 2007c for specialist HEIs). The main difficulty is that repartition by domain in the whole sample might be quite skewed: thus, in the Swiss context where 2/3 of the students are enrolled in social sciences and humanities, an institution with 75% of the students in this domain would not be considered as strongly specialised. However, looking at which domains are present and at their respective share will be useful to identify specialised institutions.

An alternative option is to build specialisation indexes against the whole population or against all HEIs in the same country. Thus, we use the relative competitive advantage index (RCA) introduced by Balassa (1965).

$$\text{SPE}_{ik} = 100 \times \tanh \ln \left\{ \frac{\text{STAFF}_{ik} / \sum_k \text{STAFF}_{ik}}{\sum_i \text{STAFF}_{ik} / \sum_{ik} \text{STAFF}_{ik}} \right\}$$

$$\text{SPE}_i = \sum_k \text{ABS}(\text{SPE}_{ik}/4).$$

$\text{SPE}_{ik}$  is the relative specialisation of an institution  $i$  in the subject field  $k$  ( $k = 1, 2, 3, 4$ ) and  $\text{SPE}_i$  is the average of the specialisation for each of the four fields we consider here. The sum runs over the whole sample (sample specialisation) or over the institutions in the same country (national specialisation). Thus, this index measures the distance of an individual HEI from the “average” institution in the sample or at the national level; the normalisation implies that  $\text{SPE}_{ik} = 0$  if the share of a field is the same as in the average,  $\text{SPE}_{ik} = 100$  if the field is the only one in the institution,  $\text{SPE}_{ik} = -100$  if the field is not present in the institution.

In the following analysis, we will use these different benchmarks—absolute share of domains, national specialisation, sample specialisation—to characterise the subject mix of HEIs, depending on the type of analysis; details are given in methodological comments to figures and tables.

## Data Sources, Limitations and Comparability Problems

The analysis is based on the database developed in the AQUAMETH project, which contains a large set of data for 488 higher education institutions in 11 European countries (Bonaccorsi and Daraio 2007b), covering funding and expenditures, students, personnel, degrees and research output. While most data are aggregated, for some of them an attempt was made to get data divided by scientific field. This is possible for a sizeable number of countries for enrolled students, undergraduates and publications from the Web of Science, for fewer of them for academic staff, PhD students and PhD graduates. Time series are available as follows in Table 1:

**Table 1** Availability of data by subject domain in the AQUAMETH database

	Enrolled students	Graduates	Academic staff	PhD students	PhD graduates	Publications
Finland	1994–2006	1994–2006	1994–2006	–	–	1994–2004
Norway	1995–2003	1995–2003	1995–2003	–	1995–2003	1995–2003
Netherlands	–	–	1995–2001	–	–	–
Switzerland	1994–2002	1994–2002	1994–2002	1994–2003	1994–2003	1994–2001
Italy	1997–2004	1997–2005	1997–2004	–	–	1995–2001
Portugal	1997–2001	1997–2001	–	–	–	1994–2001
Spain	1994–2004	1994–2004	–	1994–2004	–	1994–2004
UK	1996–2003	1996–2003	1995–2005	1996–2003	1996–2003	1994–2005
France	1999–2005	1999–2003	1995–2006	1999–2005	1999–2003	–



In this paper, data for 242 HEIs (respectively 236 concerning students, see below) from Finland (20), Norway (4), the Netherlands (13), Switzerland (12), Italy (77/74) and the United Kingdom (116/113) are analysed. Data for France are disaggregated per HEI, however, without providing the HEIs' names. As it would be necessary to identify their names for a deepened analysis, French data are not integrated in the analysis.

While the availability of data for such a large number of institutions and countries allows new analyses, one should not disregard the limitations of the dataset (see Bonaccorsi et al. 2007); since data have been collected from different national sources, full comparability cannot be ensured, despite attempts made to check for major differences. In some cases, the dataset showed inconsistencies that made it necessary to correct the data through direct contacts with researchers in the respective countries. Finally, the AQUAMETH database covers only PhD-awarding institutions, while in most of the considered countries non-university HEIs enrol a large share of students (Kyvik 2004); the absence of these institutions is likely to bias the comparisons, especially with regard to students.

Other comparability problems stem from differences in the classification of staff (especially the repartition between technical and academic staff) and from different treatment of PhD students (included in staff in some countries, but not in others; Bonaccorsi et al. 2007) and by the fact that FTE data were available only for Switzerland and the UK, while for the other countries we had to use headcounts. For Norway, disaggregated data are available only for the four largest universities and coverage is therefore not complete, while for the Netherlands, a specific problem emerges from the transfer of medical staff to University Medical Centres, which is not shown in the statistics for some universities (the Radboud University Nijmegen, the University of Leiden, the University of Utrecht, the University of Amsterdam and the VU University Amsterdam); we thus had to interpolate data with information on changes in student numbers. In most countries, staff numbers also included a subject category “multidisciplinary”. For the analysis, the—often small—numbers in this category were proportionally attributed to the other categories at the level of each institution.

## Identifying Institution Classes Through Staff

In this section, we use data on repartition of staff by domain to characterise the institutions in the sample and to develop criteria to identify institutional classes (see Table 2). We choose the year 2001 for which we have data for the largest number of institutions as reference.

While the share of social sciences and humanities is rather similar in the six countries (from 29% in Switzerland to 39% in Finland and UK), the share of the other fields varies to a larger extent: for example, technical sciences account for 10% of staff in Norway, but 28% in Finland. It remains open to what extent these differences are due to different national orientations, to the composition of the sample excluding the non-university sector or to different classification schemes.

**Table 2** Institutions in the sample and their staff

Country	N Inst	Tot	EngTech (%)	Med (%)	Nat (%)	HumSoc (%)
Finland	20	14,238	28	14	19	39
Italy	77	86,587	17	24	21	38
Norway	4	8,762	10	36	21	33
Switzerland	12	18,105	20	20	32	29
Netherlands	13	24,528	25	24	16	36
United Kingdom	116	105,593	14	23	24	39

Data for the year 2001. Data are in FTE for Switzerland and UK and in headcounts for Finland, Italy, Norway and the Netherlands

The SPE index allows us to understand how these shares are distributed on individual HEIs or, in other words, to what extent a HEI is specialised in (a) certain subject field(s) compared to all HEIs of a same country. The SPE index computation for these institutions (average of absolute values for each domain) provides wide variations between a minimum of 11 for two HEIs in Italy (Sassari, Roma La Sapienza) and three in the United Kingdom (Paisley, Plymouth and North-East Wales Institute of Higher Education) and a maximum of 99 for two specialised universities (Harper Adams University College and The Royal Veterinary College in the UK). About 51 institutions have SPE indexes of 90 and more (we shall see that these are the institutions specialised in a single field), further 72 above 50, while the remaining 119 are below this value.

If we average these indexes, we get a rough measure of the overall specialisation per sector in each country (referred to the *national* average; see Table 3). Some interesting patterns emerge: thus, there is a much stronger specialisation of individual HEIs in Finland and the Netherlands than in Norway, the United Kingdom and Italy (for Norway this can be explained by the sample used, which excludes specialised institutions). Concerning sectors, we get higher values of specialisation for technology and medicine than for natural sciences and social sciences and humanities.

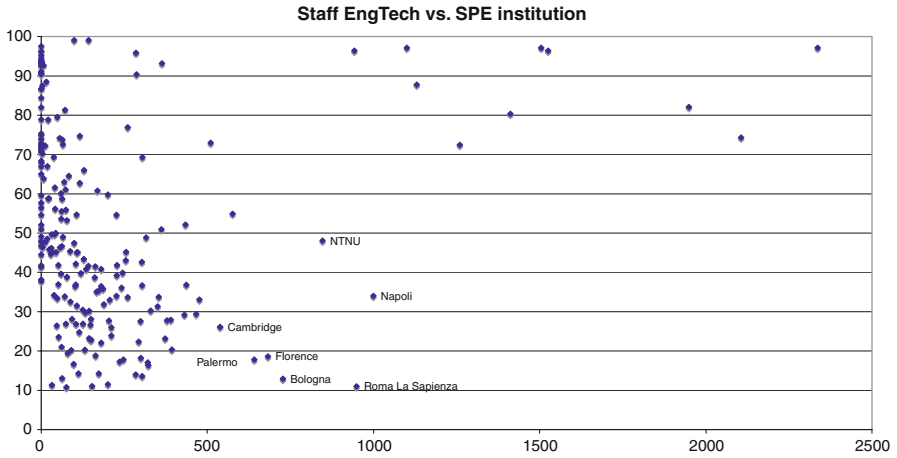
### Patterns of Specialisation by Domain

A more precise analysis displays different patterns of specialisation according to the field considered. As the following figure displays (Fig. 1), almost all institutions

**Table 3** Average specialisation index per country and subject domain, 2001

	EngTech	Med	Nat	HumSoc	Average
Finland	85	88	79	63	79
Italy	57	69	49	49	56
Norway	86	31	23	11	38
Switzerland	89	74	45	53	65
The Netherlands	97	61	73	47	69
United Kingdom	54	63	42	39	49

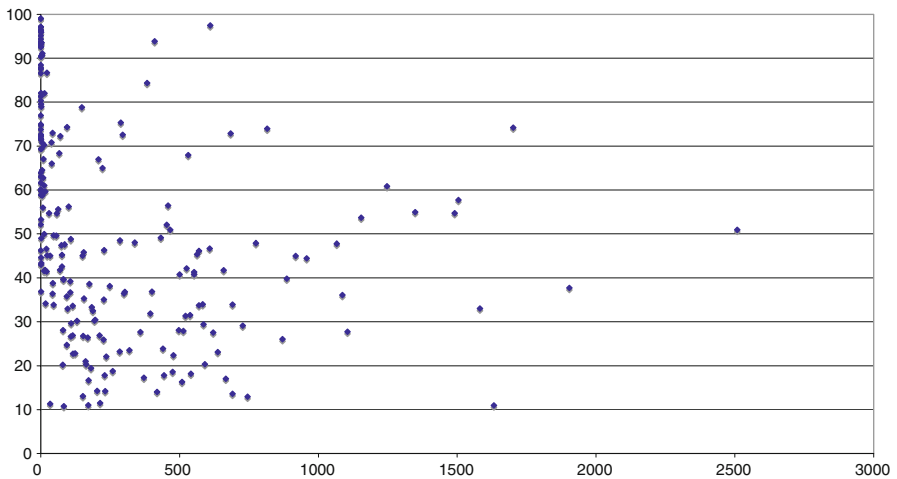
Unweighted averages of absolute values



**Fig. 1** Specialisation index in technical sciences (2001). X axis: Staff in technical sciences. Y axis: average SPE of the institution at a national level

with a large technical department (staff > 500) are highly specialised, with just seven exceptions, namely five generalist Italian universities (Palermo, Florence, Bologna, Napoli and Roma La Sapienza), one generalist university in the UK (Cambridge) and the Norwegian University of Science and Technology NTNU.

Medicine (Fig. 2) presents also a strong concentration, but with a different pattern. Only 35% of the institutions in our sample has a sizeable medical department (larger than 200 employees), while the 25 institutions with the largest



**Fig. 2** Specialisation index in medicine (2001). X axis: Staff in medicine. Y axis: average SPE of the institution

departments concentrate about half of the total staff. However, these represent, in fact, the largest generalist universities in their country: the largest medical departments can be found at the University College for the UK, in Rome for Italy, in Zurich for Switzerland and in Oslo for Norway. Among the sample there is a smaller group of 13 universities where medicine accounts for more than half of the total staff, including Pavia, Catanzaro, Siena and Insubria in Italy, Rotterdam and Maastricht in the Netherlands, the University College and King's College in London, UK, and the University of Kuopio in Finland, as well as four specialised institutions in this field (at least 80% of staff in medicine), two in Italy (Milano San Raffaele and Roma Campus Biomedico) and two in the UK (University of Wales College of Medicine and London School of Hygiene & Tropical Medicine).

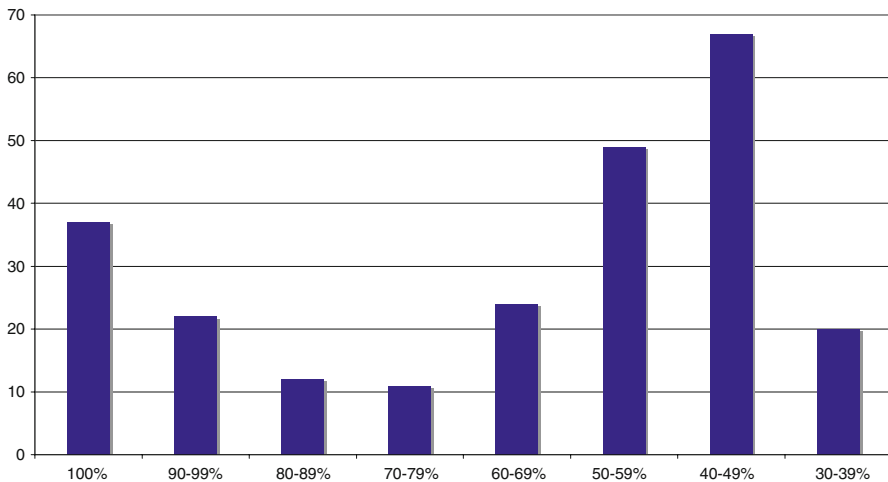
On the contrary, natural sciences and human and social sciences are much less concentrated and are present in most institutions in the sample. Natural sciences account for at least 10% of staff in 70% of all HEIs in the sample, human and social sciences in 90%.

### Identifying Classes of Institutions

This discussion suggests that not all possible combinations of the four domains are equally present in our sample and therefore puts forward two criteria for distinguishing classes, namely specialisation in a single domain and the weight of the medical department.

From an ideal-type point of view, *specialist institutions* should be institutions having staff in just one of the four domains. However, we need to take into account a few specific cases:

- A first case are technological schools with a large natural science department, being probably a sign of the integration process of the two domains; these are a number of large institutions, like ETH Zurich or the Polytechnic of Milan, which one would consider as typical examples of technical schools. Thus, we include them among technical schools. We notice that this could be considered a case where the division between natural and technical sciences in the FOS classification does not match the organisation of science anymore, as some evident border cases indicate (e.g. informatics).
- A second case consists of institutions with a small additional domain with a few units of personnel. We therefore decide that we include them among specialised institutions if the share of the main domain is higher than 85% (for medicine 80%). The choice of the cut-off point is justified by the fact that only 33 institutions in the sample have their main domain between 70% and 95% of total staff (see Fig. 3). Thus, generalist and specialist institutions represent two distinct populations with little overlap which can be readily separated.



**Fig. 3** Institutions by the share of their main domain. Number of institutions where the most important domain accounts for a given share of total staff, considering natural + technical sciences, medicine, human and social sciences

Secondly, we adopt as a main criterion to disregard whether the remaining *generalist institutions* have a *medical department*, since the previous analysis indicates that large medical departments are concentrated in a rather small share of institutions. We distinguish between two classes:

- Generalist institutions without a medical department, in practice where medicine accounts for less than 10% of the total staff (excluding multidisciplinary). In fact, the cut-off is quite clear, since most often the share of medicine is less than 2%.<sup>1</sup> Most of these institutions have all three remaining subject fields, but there are cases with only natural sciences and human and social sciences (especially in countries where technical sciences are concentrated in specialised institutions), as well as a few cases of institutions with technology and human and social sciences.
- Generalist institutions with a medical department: most of them cover all four subject fields, with exceptions where technology is concentrated in specialised institutions, like in Switzerland. Among them, we can single out a small group where medicine covers more than half of the total staff (institutions with strong medicine).

By applying the previous criteria we can identify two main classes of specialised institutions—in technical sciences (including technical-natural sciences) and in social sciences and humanities—as well as two main classes of generalist institutions, those with and those without medicine (Table 4).

<sup>1</sup> However, in some, especially British HEIs, the proportion of staff in medicine is around 8–9%.

**Table 4** Classes of HEIs by subject mix. Classification criteria

Category	<i>N</i> Institutions	Criteria
<b>Specialist HEI</b>		
Engineering & Technology	10	The main domain has more than 85% of the total staff (80% for medicine), considering natural and technical sciences together. Subdivisions according to the main specialisation domain
Engineering & Technology and Natural Sciences	5	
Medicine	4	
Natural Sciences	1	
Humanities & Social Sciences	45	
<b>Generalist HEI</b>		
With strong medicine	9	Medicine accounts for more than 50% of the staff
With medicine	109	Medicine accounts for 10–50% of the total staff
Without medicine	59	Remaining institutions which are not specialised and where medicine accounts for less than 10% of the staff

### Characterising Institutional Classes

It is now interesting to look at systematic differences between classes of institutions concerning size, age, research intensity, etc. to understand to what extent these features are specific of classes instead of individual institutions or of national systems. A first basic characterisation of the HEIs in the sample is provided in Table 5.

Thus, roughly 1/4 of the sample is composed by specialist institutions which, however, account for less than 13% of the total staff. Technical-natural institutions are mostly large institutions, which account for a significant share of the staff in their domain, while human and social sciences institutions are mostly smaller niche institutions.

**Table 5** Classes of HEIs by subject mix

Category	<i>N</i>	Staff	Average Staff
<b>Specialist HEI</b>			
EngTech	10	8,693	869
EngTechNat	5	10,517	2,103
Med	4	1,604	401
Nat	1	58	58
HumSoc	45	10,816	240
Total	65	31,688	3,672
<b>Generalist HEI</b>			
With strong medicine	9	16,836	1,871
With medicine	109	165,731	1,520
Without medicine	59	43,559	738
Total	177	226,125	4,129
Grand Total	242	257,813	7,801

Basic characterisation

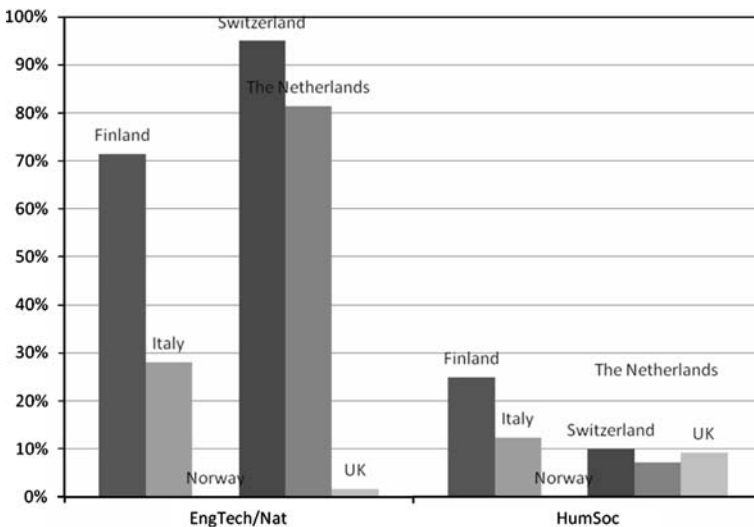
The core of the higher education system is composed by the 118 generalist institutions with medicine accounting for more than 2/3 of the total staff; most of them are large institutions, including the oldest and largest universities in each country.

Finally, the smaller group of generalist institutions without medicine (59 institutions) is composed by medium-size institutions, comprising about 17% of the total staff; some of them are, however, rather large, like Fribourg in Switzerland, Salerno, Lecce and Trento in Italy and the Open University in the UK.

### Looking at National Patterns

When one looks at individual countries, differences emerge. Thus, in all considered countries, generalist institutions with medicine constitute the core of the higher education system; however, the role of specialised institutions differs rather strongly. These include about 1/3 of the staff in Finland—where 12 universities among 20 are specialised—and in Switzerland—owing to the strength of the two federal institutes of technology—about 20% in the Netherlands, but much less in Italy and the UK, whereas they are inexistent in the observed Norwegian sample, given its composition.

As shown in Fig. 4, differences are particularly relevant for technical sciences, where in three countries most of the staff is concentrated in specialised institutions, while there are very few technical universities in the UK and not one in Norway (at least in the sample we included). Italy represents an intermediary case with some large technical universities alongside technical departments in generalist institutions. On the contrary, specialisation in human and social sciences is more a niche phenomenon in all of the countries considered.



**Fig. 4** Share of domains in specialised institutions. Share of staff in technical sciences and human and social sciences in specialised institutions at a national level

These preliminary results suggest that we use these classifications by classes to systematically compare specialisation patterns between different countries and then look for specific national explanations; however, for this aim a critical improvement would be to have a broader and more comparable perimeter of HEIs across all countries, for example covering also the entire non-university sector.

## Understanding Dynamics and Historical Development

A further relevant issue is to look at the dynamics over time of subject mix and the repartition in classes. For doing so, we look at the long term dynamics inquiring whether there is a link between the type of HEIs and their foundation year and then we observe changes in subject mix between 1997 and 2003. In order to address the first issue, we distinguish seven time periods (see Table 6).

Even though we need to be careful since we are comparing age with current specialisation, some patterns emerge. Namely, among the oldest universities, generalist institutions with medicine dominate: 33 out of the 41 institutions founded before 1800 belong to this class and this can be readily explained by a model where new domains are integrated in existing institutions, taking into account that medicine constituted one of the cores of the medieval university. After 1800 a new pattern emerges where specialised institutions, especially in humanities and engineering and technical sciences, are established, as well as many generalist institutions without medical department (especially after 1950). In a historical perspective, differentiation by subject domains seems therefore to be a rather recent phenomenon, probably related to expansion and differentiation of activities of HEIs in the last century.

However, one could inquire if, at least among generalist institutions, the share of different subject domains is being modified, which would also mean that individual

**Table 6** HEI types and foundation periods

	– 1400	1400– 1599	1600– 1799	1800– 1899	1900– 1949	1950– 1979	1980– 2001	Total
<b>Specialist</b>								
EngTech			1	1	4	3	1	10
EngTechNat				2	1	1	1	5
HumSoc	1	1	1	13	11	11	7	45
Med				1	1		2	4
Nat						1		1
<b>General</b>								
With strong medicine	2			2		3	2	9
With medicine	12	13	6	38	19	14	7	109
Without medicine	1	2	1	24	4	16	11	59
<b>Total</b>	<b>16</b>	<b>16</b>	<b>9</b>	<b>81</b>	<b>40</b>	<b>49</b>	<b>31</b>	<b>242</b>



**Table 7** Change in staff per country and domain, 1997-2003

	EngTech (%)	Med (%)	Nat (%)	HumSoc (%)	Total (%)
Finland	36	23	35	16	26
Italy	39	61	34	91	61
Norway	26	28	-3	18	17
Switzerland	12	9	14	15	13
The Netherlands	13	13	28	14	15
United Kingdom	-5	39	20	15	17
Total	16	40	23	36	30

For the Netherlands, data for 2003 are not available, therefore change refers to 1997-2001

HEIs are able to reallocate resources between sectors according to their strategic priorities. For this aim, we compare the share of the four domains across four countries for the years 1997 and 2003 (Finland, Italy, Norway, Switzerland and UK; 1997 and 2001 for the Netherlands).

A look at the increase of staff (Table 7) shows that even within this short period of time there would have been room for reallocating resources since the staff increased by 30% over these 6 years. Differences between domains emerge, but are essentially limited to distributing differently the increase in staff numbers.

If we compute for each institution the change of the share of each domain in total staff, the *average* of the absolute values between 1997 and 2003 is just 3.9%, which shows that the shares have been very stable not only at the national level, but also regarding individual institutions and, thus, there is no evidence of systematic specialisation of institutions in specific domains at least at the level of broad subject domains (the picture is likely to be very different for specialities); these results confirm recent work on educational specialisation of Italian universities (Rossi 2009).

If one looks at the institutions with at least 100 staff members in 2003, for which the share of at least one subject domain has changed by at least 10 percentage points, we get 48 institutions over the 217<sup>2</sup> in the sample (see Table 9). Two of them are Swiss HEIs, 20 are Italian, and 26 from the United Kingdom. 13 of them are fast-growing institutions, with an increase in total staff of at least 80% over 6 years, 12 of them being located in Italy.

Among the remaining cases, we find some rather large institutions where the share of a domain changed significantly, like medicine at the Imperial College (from 28% to 44%) and in Milan (from 38% to 57%). We also find cases where the absolute size of a domain has been reduced, even if the institution itself has grown. A reduction of at least 100 units of staff in one field occurred in London South Bank University, The University of Greenwich, De Montfort University and Southampton Institute (engineering and technical sciences), at Milano, Lausanne, De Montfort University and Camerino (natural sciences), London South Bank University (social sciences and humanities) and Milano Cattolica (medicine).

Due to the limitations of the data and to possible differences between years in coverage, these results should be taken with caution. However, the example of the

<sup>2</sup> Excluding the 12 institutions for which no information for 1997 was available.

University of Lausanne, which chose to transfer most of the natural sciences (except biology) to the Federal Institute of Technology in Lausanne to focus on human and social sciences and life sciences, points to the fact that these changes might be the outcome of explicit strategic choices. Noteworthy, for Italy and the UK, the number of universities which significantly modified their allocation of resources in the considered period is not negligible, and this phenomenon is not limited to small institutions, but also includes some very large ones. For the other countries in the sample, however, stability between these large subject fields seems to prevail.

### **The Educational and Research Profile: How Indicators Are Related to Classes**

In this section, we examine subject composition of HEIs from the perspective of the distribution by subject domains of educational and research products; as discussed in section “[From Disciplinary Diversity to Subject Mix](#)”, not only are we looking for different dimensions of specialisation (resources vs. products), but we are also using different types of measures and, thus, differences to the previous discussion are likely to emerge.

We perform the analysis for the following indicators: number of students, being a proxy for educational production, number of PhD degrees, the basic indicators used in the US Carnegie classification of universities for research intensity, and the number of publications in the WoS, a widely used measure of international research production. Of course, this choice is also partially limited by the fact that we use large measures of the volume of educational and research production, which do not consider quality (for example, citation impact indicators would be preferable in this respect).

There are some potential sources of differences between these indicators of specialisation. One are differences in the production structure and productivity by classes of institutions, related, for example, to economies of scope between institutions covering different fields or to economies of scale of concentrating effort in a single large domain. Another explanation are differences in the organisation of scientific production by field, for example that classes in technical sciences need to be smaller than in social sciences because of the subject taught, and this, of course, would influence the number of students and degrees produced by level of input (as measured by the number of staff).

A third possibility is that the same indicator is measuring different things across scientific fields, like in case of measures of research output through the number of publications in the Web of Science. The reader will notice that this is also a normative issue related to how we define the notion of scientific production—which kind of results we include, how we value different results, etc.—and, thus, there is no general answer to these questions.

To disentangle these aspects would go well beyond the scope of this paper and would require recourse to some modelling techniques, as well as the integration of more qualitative information. In the following, we limit ourselves to a descriptive analysis looking at systematic differences between classes of institutions (by subject

mix), subject domains and countries and highlighting some regularities which might hint to underlying structural effects.

### Comparing Staff and Students Subject Mixes at National and Institutional Levels

To understand differences between student and staff composition, we have performed the attribution to the identified classes also based on the student numbers. Two relevant results emerge. Firstly, practically all institutions belong to the same class for both indicators and, secondly, there are systematic differences in staff endowment per student between domains independent of the country considered. This is a highly relevant result for the study of higher education systems since it shows that in broad terms specialisation by effort and by educational products coincide; to some extent this can be considered an outcome of using the broad subject domains for classification according to which higher education is organised.

Thus, when looking at the distinction between generalist and specialist institutions, only 17 HEIs out of 236 change their profiles (for 6 HEIs no information on student numbers was available). This reflects to some extent the fact that, in the whole sample, the share of students is systematically larger than the one of staff for social sciences and humanities and lower for medicine: thus, at the University of Rotterdam, for example, the staff is equally divided among medicine and human and social sciences, but 85% of the students are in the latter domain. Most of the institutions changing from generalist to specialist with this additional classification already have above 70% of staff in human and social sciences. A similar case are 26 institutions which are generalist with medicine in terms of staff (10.1%–49.9%, 22 of them below 30%), but without medicine in terms of the students' share (3.4%–9.7%).

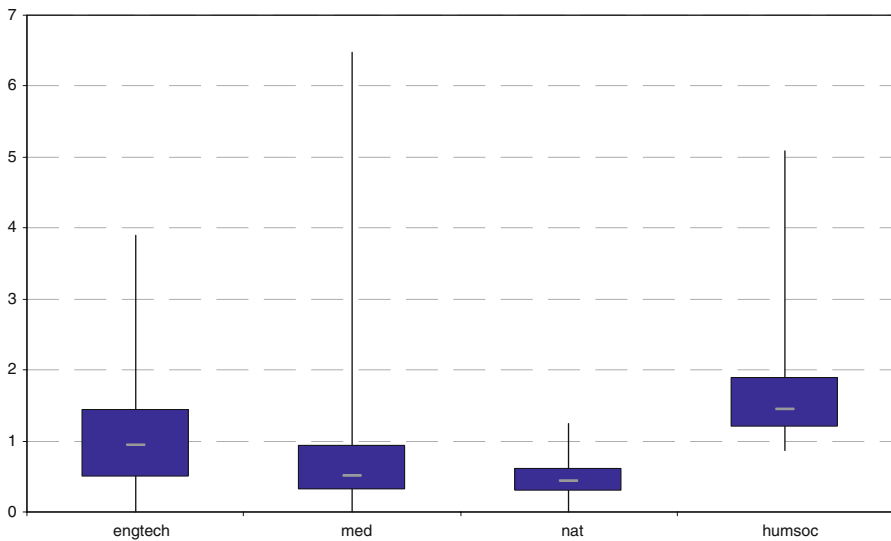
However, as Table 8 displays, differences between staff and students' proportions are systematic between disciplinary sectors. Medicine and natural sciences have staff percentages which are well above the students' proportion, percentages of staff and students in engineering and technical sciences are close to one another, and the students' proportion in social sciences and humanities is well below the one of the staff.

It is interesting to examine to what extent these patterns also appear at the institutional level. For this aim, we calculate for each institution the ratio of students to staff in each subject domain, and we normalise it against the institutional average (thus, a ratio of 1.2 for a domain means that it has 20% more students per unit of staff than the average for the whole institution). We limit this analysis to the generalist institutions as shown in Fig. 5.<sup>3</sup>

<sup>3</sup> However, we had to exclude 9 HEIs from this analysis (1 from Switzerland, 2 from the Netherlands and 6 from the United Kingdom): for these HEIs, student numbers were reported for fields for which no staff was reported.

**Table 8** Staff and students' distribution per subject domain and country (2001)

	EngTech		Med		Nat		HumSoc	
	Staff (%)	Student (%)	Staff (%)	Student (%)	Staff (%)	Student (%)	Staff (%)	Student (%)
Finland	28	23	14	5	19	15	39	56
Italy	17	20	24	9	21	7	38	64
Norway	10	7	36	9	21	17	33	66
Switzerland	20	10	20	9	32	15	29	66
The Netherlands	25	17	24	12	16	7	36	64
United Kingdom	14	14	23	13	24	11	39	62
Total	17	17	23	11	22	9	38	63



**Fig. 5** Student to staff ratio for all institutions in the sample, 2003. 1 = the average for each individual institution. Median of the values for each institution; boxplots of 1st and 3rd quartiles

In human and social sciences, 50% of the institutions in the sample have shares between 1.19 (19% more than the institutional average) and 1.89, with the median being 1.44. At the other extreme, in natural sciences, 50% of the institutions have shares between 0.29 and 0.61, with a median value of 0.43 (i.e. 57% less students per staff than the institutional average). Medicine has also much less students per staff than the institutional average, but the spread between institutions is larger, while technical sciences are around the average, but with an even larger spread. Differences by field are larger than differences by institution: thus, just 6% of institutions get a score under 1 in human and social sciences (meaning they have

less students per staff than the institutional average), while just 3% of the institutions get a score higher than 1 in natural sciences and in medicine.

These results are highly significant if we consider the diversity of the sample concerning countries, share of different domains and average students to staff ratios: in the whole sample, the median value of the student to staff ratio is 14, but the first quartile is 8 and the third quartile 19; moreover, differences in the average student to staff ratios vary widely between countries from 4.4 (Switzerland) to 19.4 (Italy). Of course, this raises the question whether these differences are related to the disciplines' general modes of research organisation, the balance of power between disciplines, historical constraints or any other factors.

### Does Research Intensity Depend on Subject Mix?

A further relevant question is to what extent are indicators concerning research intensity systematically different between classes. Despite its limitations, we use the main indicator adopted in the US Carnegie classification, namely the number of PhD degrees per 100 undergraduate students, knowing that in the American classification 1 is taken as the threshold to identify research-intensive institutions.

Of course, there would be good reasons to expect systematic differences between classes; for example, one could expect that technical schools—most of them having a strong research reputation—display systematically higher values of this indicator; the same would be expected for generalist institutions with medicine, being the largest, oldest and most reputed institutions in their country.

However, data show that national patterns are much more evident than sectoral patterns for this indicator (Table 9). This should cast some doubt on the usability of this indicator to classify higher education institutions at the European level.

**Table 9** Research intensity by subject mix class

	Finland	Italy	Norway	Switzer-land	The Netherlands	UK	Total
<b>Generalist</b>							
With medicine	1.02	0.24	0.79	3.62	1.50	1.14	0.71
With strong medicine	1.43	0.24			0.98	3.45	1.15
Without medicine	0.91	0.15		1.31	1.80	0.57	0.46
Total	1.01	0.23	0.79	3.16	1.41	0.97	0.67
<b>Specialist</b>							
EngTech	0.69	0.30			2.18	1.35	1.16
EngTechNat		0.26		5.23			1.04
HumSoc	0.42	0.21		2.93	0.60	0.59	0.47
Med		0.00				0.93	0.62
Total	0.57	0.24		4.65	1.72	0.64	0.78
Total	0.87	0.23	0.79	3.49	1.47	0.95	0.68

PhD degrees for 100 undergraduate students, 2001

Thus, the average number of PhD degrees per undergraduate students varies from 3.49 in Switzerland to 0.23 in Italy, and this seems to reflect completely different functions of the doctorate related to the work market, as well as roles of PhDs in the workforce of the universities (countries with high ratios hiring PhD students systematically as workforce in research projects or as teaching assistants; (Filippini and Lepori 2007)). Differences by class are present in some cases, like the higher value for technical universities in Switzerland, but these seem to be less systematic (taking into account also the number of institutions in some classes is very low).

If we look at the 71 institutions exceeding the threshold of 1 PhD degree per 100 undergraduate students, no regularity emerges: 41 are generalist with medicine, 17 without medicine, 6 are specialised in technical sciences or technical and natural sciences, and 7 are specialised in social sciences and humanities. 11 are Dutch, 10 Swiss, 3 Finnish and 2 Italian. The large share is formed by 45 HEIs in the United Kingdom, a country where, on the other hand, we find 18 HEIs that did not attribute any PhD degree in 2001, and 14 more with less than 0.1 PhD recipients per 100 students.

Unfortunately, the AQUAMETH database includes data on PhD degrees disaggregated by subject mix only for few countries, thus it is impossible to compute specialisation indexes similar to those for staff and students; this is in our view a further priority for future improvements, which would allow a development towards a research profile, as well as first comparisons with educational profiles.

### **To What Extent Does Publications' Output in the Web of Science Depend on Subject Mix?**

A final issue concerns the measurement of research output using publications from the Web of Science database (WoS). It is well known that its coverage is strongly different between subject domains, being much more complete in natural sciences and medicine than in social sciences and humanities and, to a lesser extent, in technical sciences (Nederhof 2006). Hence, there is the possibility that publication counts and measurements of productivity at the level of all HEIs are strongly influenced by subject mix.

To address these issues, we use disaggregated data by domain for staff and scientific publications in our sample for the year 2001. These data have to be handled with care since there is no one-to-one correspondence between publication fields (based on journals) and staff fields (based on organisational units). However, at the level of broad domains, we assume a reasonable correspondence, even if some problematic cases can be expected (for example, at the border between technical and natural sciences). Thus, in AQUAMETH, a correspondent table has been built between the 29 fields of Current Contents and the four fields we are considering here.

Table 10 shows that in our sample social sciences and humanities count for 38% of the total staff, but only for 9% of WoS publications. There is also a large difference between technical sciences on one side, and natural sciences and

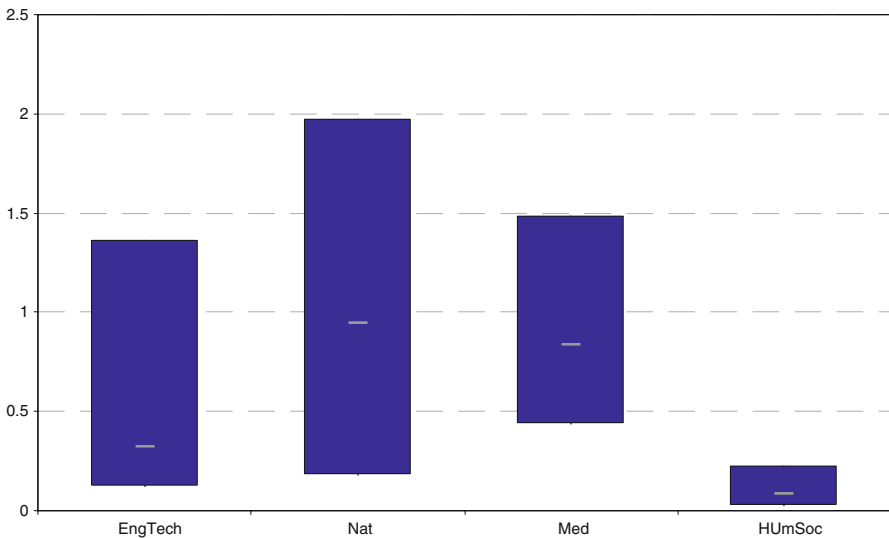
**Table 10** Number of Web of Science publications by domain (2001)

	Staff	EngTech	Med	Nat	HumSoc	Tot
<b>Specialist</b>						
EngTech	8,693	1,888	490	3,211	181	5,770
EngTechNat	10,517	1,088	291	2,452	134	3,965
Med	1,604	171	1,944	863	403	3,381
Nat	58	8	36	256	12	313
HumSoc	10,816	55	191	234	722	1,201
<b>Generalist</b>						
With strong medicine	16,836	660	9,542	3,608	2,100	15,910
With medicine	165,731	12,322	48,970	47,627	9,392	118,312
Without medicine	43,559	3,954	4,459	12,663	2,886	23,962
<b>Total</b>	<b>257,813</b>	<b>20,146</b>	<b>65,923</b>	<b>70,915</b>	<b>15,830</b>	<b>172,814</b>

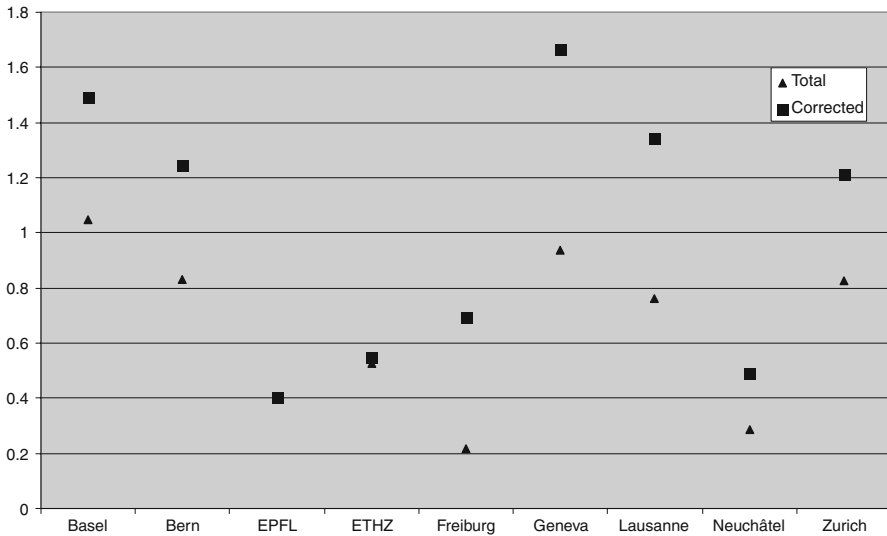
medicine on the other. This corresponds by large to the current state of the art concerning Web of Science coverage.

If we compute the number of publications per unit of staff by subject domain at the level of individual institutions, human and social sciences display systematically much lower values than the other domains for most institutions in the sample; on the contrary, differences between individual institutions in the other fields are larger than average inter-field differences (see Fig. 6).

A more refined analysis looking to the correlations between different factors—field, class and country—would be at place here to investigate in-depth these differences. However, in terms of analyses at the institutional level, the main



**Fig. 6** Scientific publications per unit of staff, 2003. Median of the values for individual institutions; boxplots of 1st and 3rd quartiles. Extremes values not plotted since some are out of range



**Fig. 7** Publications per staff of Swiss universities. Total refers to all WoS publications by total staff, corrected to all WoS publications by staff excluding human and social sciences staff

methodological implication of the previous discussion is that research profiles using WoS data *de facto* exclude social sciences and humanities; this implies that, first, specialised institutions in that field should be excluded from institutional comparisons based on WoS data and, second, when calculating numbers of publications per staff, staff in social sciences and humanities should be excluded.

This correction does not require having publications data by field, but would at least correct the most important difference related to subject mix; as the above figure (Fig. 7)—based on Swiss HEIs—displays, the effect is not only to raise the average productivity level, but also ranking by productivity of individual institutions is affected. We therefore strongly advise to adopt this approach for institutional comparisons.

## Conclusions

This analysis should be considered as preliminary, not only because of the limitations of the data, but also since we restrained ourselves to a descriptive analysis based on simple statistics, without trying to look for correlations between different factors. We consider this approach adequate for a first exploration, since it allows investigating some basic patterns and to take into account the limitations of the data sources, as well as some individual specificity. Its strength has been to identify some broad patterns in a sample of 242 institutions in six European countries which possess very different higher education systems; thus, there is some trust that these results might be valid also for a broader set of countries.



The main result is the identification of a few classes of HEIs, which account for most of the differences between European universities concerning subject mix; thus, it has been possible to single out specialist and generalist institutions, with a rather clear-cut separation, and to show that there are basically two types of specialist institutions, the technical schools on one side, the institutions specialised in human and social sciences on the other. Beyond this, the core of the higher education system in all considered countries is constituted by the old large generalist institutions, which concentrate most of the staff and activities in medical sciences (being thus a distinct feature of this group); especially in the twentieth century, a number of generalist institutions without medicine have been created (for example in Italy) to accommodate the increase in student numbers.

Patterns of subject mix in HEIs—at least at the broad level we considered—are thus largely related to long-term historical developments and are likely to be very stable. On the other side, however, differences between countries exist regarding the proportions of generalist and specialist HEIs, which might be explained by varying national policies, which include the establishment of such specialist HEIs, the increased student demand in the field of social sciences and humanities, as well as varying institutional strategies to concentrate on certain subject fields. However, an exploratory analysis comparing institutional profiles between 1997 and 2003 identified relatively few cases with radical changes between the four subject fields. If this observation could be confirmed over the long term, this could mean that changes in institutional profiles take place rather within than between broad subject fields.

As a further step in this analysis, we strongly suggest to compare these results with more qualitative reconstructions of higher education systems in the considered countries; not only would this provide a better check of robustness of results, but it could also provide directions to understand the forces beyond the observed patterns, as well as explaining the national differences we observed.

A comparison between staff and student numbers regarding subject fields at the institutional and national levels brought further insight: on the one side, considering students' presence in the different subject fields, very few HEIs would be differently classified. From this point of view, there is a broad correspondence between profiles derived from effort (staff) and educational production. On the other side, the relationship between staff and students' percentage varies systematically across countries according to disciplinary groups. As a consequence, one may raise the question whether these patterns are related to the disciplines' general modes of research organisation, the balance of power between different disciplines, historical constraints or any other factors. Further insight could emerge in this respect by refining the analysis using the production of degrees and differentiating between levels of education.

The analysis of research production constitutes the most experimental part of this work due to limitations in the data availability, but also because one should question the adequacy of the broad classification in four subject fields to analyse subject mix in research; at the same time, the only data allowing a finer disaggregation, namely those on publications, suffer of substantial limitations in coverage especially for social sciences and humanities. Some progress in this area might be achieved by

using disaggregated data on PhD degrees and third-party research funding, but we suggest to first perform this work on a smaller sample of institutions and countries in order to look in-depth for classification issues.

These results are finally relevant for the on-going attempts to develop a European classification of HEIs, as well as for institutional benchmarking. While in a long-term perspective one would hope to integrate detailed information on subject mix in these classifications, the identified classes provide a much simpler approach, which nevertheless allows to correct some of the strongest effects of subject mix when comparing institutions without recourse to detailed data. We therefore strongly suggest including the distinction between specialist and generalist institutions—and its operationalisation suggested in this paper—in all typologies of HEIs and to compare only institutions in the same class, since differences between all kinds of indicators are indeed very large. Moreover, we showed that in some cases, simple strategies allow to correct at least the largest distortions related to subject mix. The most evident case is scientific productivity based on counts of publication in the Web of Science: since the number of publications in social sciences and humanities is very small anyway, computing this indicator excluding staff in social sciences and humanities would certainly improve the quality of comparisons between HEIs. This approach would be certainly worth of further development for other indicators.

## References

- Balassa, Bela. 1965. Trade liberalization and 'revealed' comparative advantage. *The Manchester School of Economic and Social Studies* 33: 99–123.
- Becher, Tony, and Paul R. Trowler. 2001. *Academic tribes and territories. Intellectual enquiry and the culture of disciplines*. Ballmoor: The Society for Research into Higher Education and Open University Press.
- Bleiklie, Ivar. 2003. Hierarchy and specialisation: On the institutional integration of higher education systems. *European Journal of Education* 38(4): 341–355.
- Bonaccorsi, Andrea, and Cinzia Daraio. 2007a. Theoretical perspectives on university strategy. In *Universities and strategic knowledge creation. Specialization and performance in Europe*, eds. Andrea Bonaccorsi, and Cinzia Daraio, 3–30. Cheltenham: Edwar Elgar.
- Bonaccorsi, Andrea, and Cinzia Daraio. 2007b. *Universities and strategic knowledge creation. Specialization and performance in Europe*. Cheltenham: Edwar Elgar.
- Bonaccorsi, Andrea, and Cinzia Daraio. 2007c. Universities as strategic knowledge creators: Some preliminary evidence. In *Universities and strategic knowledge creation. Specialization and performance in Europe*, eds. Andrea Bonaccorsi, and Cinzia Daraio, 31–81. Cheltenham: Edwar Elgar.
- Bonaccorsi, Andrea, Cinzia Daraio, Benedetto Lepori, and Stig Slipersaeter. 2007. Indicators on individual higher education institutions: Addressing data problems and comparability issues. *Research Evaluation* 16: 66–78.
- Clark, Burton R. 1983. *The higher education system. Academic organization in cross-national perspective*. Berkeley: University of California Press.
- Clark, Burton R. 1995. *Places of inquiry: Research and advanced education in modern universities*. Berkeley: University of California Press.
- Filippini, Massimo, and Benedetto Lepori. 2007. Cost structure, economies of capacity utilization and scope in Swiss higher education institutions. In *Universities and strategic knowledge creation. Specialization and performance in Europe*, eds. Andrea Bonaccorsi, and Cinzia Daraio, 272–305. Cheltenham: Edwar Elgar.
- Fumasoli, Tatiana and Benedetto Lepori. 2010. Patterns of strategies in Swiss higher education institutions. Submitted paper.

- Huisman, Jeroen, V. Lynn Meek, and Fiona Wood. 2007. Institutional diversity in higher education: A cross-national and longitudinal analysis. *Higher Education Quarterly* 61(4): 563–577.
- Johnes, Jill. 1990. Unit costs: Some explanations of the differences between UK universities. *Applied Economics* 22(7): 853–862.
- Jongbloed, Ben, Carlo Salerno, and Frans Kaiser. 2003. *Kosten per student. Methodologie, schattingen en een internationale vergelijking*. Enschede: CHEPs Center for Higher Education Policy Studies.
- Kyvik, Svein. 2004. Structural changes in higher education systems in Western Europe. *Higher Education in Europe* 29(3): 393–409.
- Meek, V. Lynn, Leo Goedegebuure, Osmo Kivinen, and Risto Rinne. 1996. *The mockers and mocked: Comparative perspectives on differentiation, convergence and diversity in higher education*. Oxford: Pergamon.
- Meek, V. Lynn, Leo Goedegebuure, and Jeroen Huisman. 2000. Editorial: Understanding diversity and differentiation in higher education: an overview. *Higher Education Policy* 13: 1–6.
- Morphew, Christopher C. 2000. Institutional diversity, program acquisition and faculty members: Examining academic drift at a new level. *Higher Education Policy* 13: 55–77.
- Musselin, Christine. 2007. Are Universities specific organisations? In *Towards a multiversity? Universities between global trends and national traditions*, eds. Georg Krücken, Anna Kosmützky, and Marc Torca, 63–84. Bielefeld: Transcript.
- Nederhof, Anton J. 2006. Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics* 66(1): 81–100.
- OECD. 2002. *Frascati manual. Proposed standard practice for surveys on research and experimental development*. Paris: Organisation for Economic Co-operation and Development.
- OECD. 2004. *OECD handbook for internationally comparative education statistics. Concepts, standards, definitions and classifications*. Paris: Organisation for Economic Co-operation and Development.
- Rossi, Federica. 2009. Increased competition and diversity in higher education: An empirical analysis of the Italian University system. *Higher Education Policy* 22: 389–413.
- Rothblatt, Sheldon, and Björn Wittrock. 1993. *The European and American University since 1800*. Cambridge: Cambridge University Press.
- Rüegg, Walter. 2004. *A history of the University in Europe, Vol. III: Universities in the nineteenth and early twentieth centuries*. Cambridge: Cambridge University Press.
- Salancik, Gerald R., and Jeffrey Pfeffer. 1974. The bases and use of power in organizational decision making: The case of a University. The entity from which ERIC acquires the content, including journal, organization, and conference names, or by means of online submission from the author. *Administrative Science Quarterly* 19(4): 453–473.
- Sarrico, Claudia S., Pedro Teixeira, Maria João Rosa, and Margarida F. Cardoso. 2009. Subject mix and productivity in Portuguese Universities. *European Journal of Operational Research* 197(1): 287–295.
- van Vught, Frans. 2007. *Diversity and differentiation in higher education systems*. Oslo: The Norwegian Association of Higher Education Institutions, Jubileumsseminar i anledning Universitets- og høgskolerådets 50-årsjubileum.
- van Vught, Frans, Jeroen Bartelse, David Bohmert, Jon File, and Christiane Gaethgens. 2008. *Mapping diversity. Developing a European classification of higher education institutions*. Enschede: CHEPS Center for Higher Education Policy Studies.