Measuring institutional diversity across higher education systems

Jeroen Huisman[1,*], Benedetto Lepori[2], Marco Seeber[1],
Nicoline Frølich[3] and Lisa Scordato[3]

[1]*Department of Sociology, Ghent University, Korte Meer 5, 9000 Ghent, Belgium,* [2]*Faculty of Economics, Università della Svizzera Italiana, Via G. Buffi 13, CH-6904 Lugano, Switzerland and* [3]*Nordic Institute for Studies in Innovation, Research and Education (NIFU), Wergelandsveien 7, 0167 Oslo, Norway*
*Corresponding author. E-mail: jeroen.huisman@ugent.be

Institutional diversity—the variety of higher education institutions within a higher education system—has been a popular theme in higher education (policy) research. Despite the large amount of work, key challenges to a better understanding of institutional diversity have hardly been addressed. We argue that particularly the problems of conceptualizing and measuring diversity across higher education systems are underestimated. This article therefore focuses on selecting salient dimensions of difference (i.e. important dimensions of organizational action), setting the perimeter of the populations studied, dealing with outliers and choosing appropriate diversity measures. We highlight the careful steps that need to be taken to make comparisons across higher education systems meaningful and use European data to illustrate the consequences of particular choices.

1. Introduction

Scholarly attention to institutional diversity in higher education can be traced back to sociologically oriented studies addressing the shape and size of the US higher education system. Acknowledging the growing variety of higher education institutions (HEIs) over time, analysts expressed concerns about processes of academic drift, in which 'lower-tier' institutions tried to emulate elite institutions. The expectation was that this could lead to potential problems in relation to serving the increasing numbers and diversity of students and the varying demands of other stakeholders (local, regional, or national) on higher education (e.g. Jencks and Riesman 1968). The general picture emerging from the US analyses is not different from the situation in Europe (e.g. Neave 1979; Huisman and Van Vught 2009), in Australia (Meek 1991), Canada (Skolnik 1986), and China (Zha 2009). Almost all of these studies stress diversity being under pressure. It is therefore not a surprise that maintaining a certain level of diversity has been high on the policy agendas of many governments (Reichert 2009).

Surprisingly, higher education studies have paid limited attention to the conceptualization and measurement of diversity. Governments often vaguely allude to institutional variety being wished for—in light of the (often assumed, not always substantiated) increasingly diverse student body—but are not always explicit about the particular way in which institutions should differ. This vagueness can also be traced to the many diversity studies in higher education that largely provide narrative accounts and impressions of diversity levels. For sure these studies serve a purpose (see e.g. Neave 1979 for a historical account), but are unsatisfactory for the conceptualization and measurement of diversity.

A handful of studies have, however, addressed conceptual and measurement issues. Birnbaum (1983) investigated diversity in US higher education in 1960 and 1980. He argued that institutional control (four values), size (three values), sex of students (two values), programmes offered (four values), degree levels (four values), and minority enrolment (two values) are salient characteristics of HEIs and developed a typology of institutions on the basis of the different configurations of values of the characteristics. He found 141 types in 1960

(of the theoretical maximum of 768 types) and 138 types two decades later. Using Lorenz curves, he showed that the diversity—technically speaking the level of inequality—was slightly larger in 1980 than in 1960. Tight (2007) investigated the diversity of the UK's higher education system using readily available descriptive data on the institutions, particularly the student body (gender, degree-level enrolments, full-time/part-time, etc.). Stanley and Reynolds (1994) used cluster analysis and similarity groupings as a technique, on the one hand they used governmental data on university characteristics (e.g. number of staff by discipline, percentages of funding by source, percentage of female students, percentage Aboriginal enrolments) and performances (e.g. amount of research grants, degree completions) and on the other hand data from the Good Universities Guide (GUG) (e.g. admissions flexibility, gender balance, employment prospects) to show that quantitative governmental data leads to different similarity groupings than evaluative GUG data. A study by Lepori et al. (2014) also made use of the notion of 'distances' between (Swiss) HEIs to exemplify diversity within and between the two types of HEIs in that system.

Comparative and longitudinal research is rare, the study of Huisman et al. (2007) being an exception. Their study ranked higher education systems in terms of their overall levels of diversity—looking at institutional control, disciplines, degree level, and size, using diversity measures comparable to those used by Birnbaum (1983). The brief summary of these studies shows that researchers use different characteristics to conceptualize diversity, although most pertain to education, research, and/or outreach as a point of departure to reflect on similarities and differences between HEIs within a system and use different methods and techniques to measure or exemplify these differences.

To better understand whether and how higher education studies contributed to our understanding of diversity, it is useful to benchmark these studies against a general framework for analysing diversity offered by Stirling (2007). He defines diversity as '... an attribute of any system whose elements may be apportioned into categories'. Three properties, and no more or no less, he argues (Stirling 2007: 709–10) are key to understanding diversity: the variety (number of categories into which the elements can be categorized), the balance (the dispersion of elements across categories), and the disparity (the manner and degree of difference between elements). While this conceptual framework provides a suitable basis to systematically analyse diversity, there are however two preliminary questions one needs to address, which are at the core of the methodological development proposed in this article. First, to analyse diversity one needs to identify categories and to address the robustness of diversity measures in respect to the definition of categories. Yet, for HEIs 'natural' categories (e.g. universities versus polytechnics in the pre-1992 UK system) are often not sufficiently meaningful, because

these nominal distinctions hide considerable diversity within these two categories (e.g. Tight 1988). Second, related to the problem of meaningful categories is the problem that it is not always clear which HEIs belong to the total population. To take the UK example again, Tight (2011) argues that depending on different criteria, the UK higher education system contains 108, 138, or 162 institutions. Arguably, the way one sets the perimeter for the system affects the level of diversity. While it may be possible to offer specific solutions for a within-country analysis, problems loom large when comparing higher education systems, for each system (including its governments) has its own definitions and regulations to set the perimeters of the system. We note that the problem of meaningful categories and setting the boundaries of the system is not a problem solely pertaining to the study of diversity of HEIs, but a general problem in organizational studies. As McKelvey and Aldrich (1983: 116) argue: 'Two principal difficulties remaining for organizational scientists are the selection of taxonomic characteristics and the initial selection of organizations for study'.

In summary, concerning comparative studies on diversity in higher education, there are very challenging questions. First, we cannot or should not *a priori* rely on existing classifications of HEIs, for they likely hide significant internal variety and these classifications are unlikely to be comparable across countries. Second, even if there is agreement on the perimeter of the system(s) under investigation, there are questions regarding how to deal with outliers. Third, the researcher is left with questions regarding the dimensions he or she should use to make distinctions between the HEIs in the system. Which characteristics of the HEIs or activities are most important to take into account? Are they equally important and can they be aggregated? The struggle to define relevant dimensions is visible in higher education studies (see above). There appears to be some consensus that e.g. size and type of control matter and that dimensions should relate in one way or another to the core functions of HEIs, but otherwise studies diverge considerably in their choice of salient dimensions. Fourth, what are the meaningful and relevant (statistical) measures of diversity, once dimensions have been chosen and operationalized?

This contribution picks up these questions and aims to form a better understanding of the conceptual and methodological choices a comparative researcher must make. The structure of the contribution is as follows. We first discuss the issue of meaningful dimensions, then address the challenges of setting the perimeter (sampling) and then propose robust statistical measures for diversity. Subsequently, we exploit the European Microdata Collection (EUMIDA) data set of European HEIs to exemplify the conceptual and methodological choices made and analyse the impact of these choices on diversity measures. We finish with some critical reflections on our approach and set out directions for further research.

2. The dimensions and measurement problems

We organize our discussion around three main issues, namely the relevant dimensions of diversity and their operationalization, the sampling of the population, and the construction of measures of diversity.

2.1 The relevant dimensions of diversity and their operationalization

The first issue, the identification of relevant dimensions, requires a conceptual understanding of the important characteristics of organizations, which might be related to theoretical/epistemological or functional considerations. Regarding the former, it matters whether one looks at diversity from a rational perspective (dimensions can be objectively measured) or from a constructivist's perspective (diversity is in the eye of the beholder). With respect to functions, it is important to consider the kinds of services higher education is (expected) to provide (education, research, third mission: to be generally understood as activities geared towards the application and exploitation of knowledge). Different choices lead to different definitions (and operationalizations, see next section) of diversity.

Potentially all kinds of dimensions can be considered, but it depends on what one wants to explain: generic organizational diversity or specific aspects of diversity, the latter making it easier to test hypotheses concerning the underlying mechanisms driving that specific dimension of diversity and its evolution over time. Indeed, some studies focused on the provision of specific educational services (see e.g. Huisman and Morphew 1998; Rossi 2009, 2010; Teixeira et al. 2012 on programme diversity) and have offered useful insights in the specific dynamics regarding the programme supply. Other studies have taken a more encompassing perspective, stressing a multi-functional approach and including generic organizational dimensions like size, control and student characteristics, and to some extent educational dimensions (degree level) (Birnbaum 1983; Morphew 2009) or more specifically education, research, and the third mission (Van Vught 2009; Bonaccorsi et al. 2007; Daraio et al. 2011). The latter attempts must be seen in the context of the desire to create a flexible European classification of HEIs, thus allowing stakeholders to compare (and benchmark) institutions.

Whether one takes a narrow or more generic take on the dimensions to be considered, suitable operationalizations need to be chosen. Obviously, operationalization means *approximating* the dimensions. In this respect it is reassuring that studies that have narrowed in on specific elements of the higher education fabric (e.g. programme diversity), and been able to reach considerable consensus on which variables to consider (Rossi 2009, 2010; Teixeira et al. 2012). And, as explained, authors like Birnbaum, Morphew, and Huisman et al. use similar operationalizations. These studies do not deviate much—in terms of the operationalizations—from other attempts to classify HEIs in the US (Carnegie Classification) and in Europe (e.g. Van Vught and Westerheijden 2010).

The studies cited above are valuable for their attempts to make the operationalization of (dimensions of) diversity transparent. Importantly, the broad set of dimensions used in these studies has become widely accepted (Van Vught 2009; Bonaccorsi and Daraio 2007) and been adopted in large-scale mapping projects like EUMIDA and MULTIRANK (see Van Vught and Westerheijden 2010). For sure, practical considerations matter in diversity studies, and some of the conceptual dimensions identified are not well-covered in the data we are using; at the same time, it is a great progress in diversity studies to be able to use a data set which has been explicitly constructed from an understanding of the relevant dimensions to characterize HEIs.

2.2 Sampling

If populations are 'natural' and boundaries clear, sampling is not a key issue, but—as said—problems loom for organizational studies. For this challenge, two alternatives can be distinguished, one based on functional criteria considering all organizations that belong to an organizational field, independent of their legal status and labelling (see e.g. Ruef 2000 for such an approach), the other considering only the organizations which belong to a particular organizational form, for example, analysing diversity only among universities in a country (see Huisman and Morphew 1998). When comparing countries with different organizations of higher education, these two approaches are likely to provide different results. In some countries, governments grant all HEIs university status (unitary systems), whereas in others they allocate different functions to different types of HEIs (universities, colleges, polytechnics), e.g. in binary systems. The level of diversity in university sectors may therefore differ significantly between unitary and binary systems (Lepori et al. 2014). Most scholars of higher education would argue that a broad organizational field approach is more adequate to capture the current dynamics of higher education, with its fuzzy and changing boundaries between forms (Huisman and Kaiser 2001; Kyvik 2004). However, one caveat must be made. Perimeters of higher education are often related to national policies and are likely to change over time, thus affecting cross-country and longitudinal comparisons. We notice that, with all issues related to cross-national differences, the EUMIDA data set made an attempt to provide a general definition of an HEI perimeter (see also Section 3 and footnote 1), which is broadly comparable across European countries (Lepori and Bonaccorsi 2013). Further, the introduction of the Bologna system increased

the comparability of tertiary education in Europe, as higher education can be largely equated to HEIs graduating at least at the bachelor level.

While adhering to the general principle of considering the whole higher education field, we suggest that careful decisions on inclusions or exclusions (including decisions on outliers) might be required to achieve comparability and that the rationale for the choice of a perimeter should be made explicit in each study. Further, the sensitivity of indicators against the definition of the perimeter needs to be analysed.

2.3 Measuring diversity

The literature shows various ways of measuring diversity (e.g. Birnbaum 1983; Huisman 2000). Although it may be a good point of departure to try to bring on board the three properties of diversity as Stirling (2007) proposes, there are often good reasons to focus on only one (or two) of the properties. From the perspective of parsimony, it is tempting to work with a single diversity measure (indicator) across all variables. But it is important to note that the nature of the data does not always allow this. Some variables are nominal (e.g. public/private), others are ordinal (age), and some are measured on a ratio scale (number of staff). Another crucial issue is whether indexes have a normal distribution or not. Both issues have important consequences for the choice of reliable diversity indicators.

In the case of discrete events (nominal and categorical variables), entropy measures are most appropriate, since they are suitable to treat variables that describe a different set of events (Jost 2006). The Simpson diversity index (S) is commonly used to measure ecological diversity (e.g. Pianka, 1994). It considers the number of existing types and their relative abundance. In this case, the index points out the probability that two randomly selected HEIs will not belong to the same type or category. The index is given by the following formula:

$$\lambda = \Sigma_{i-->R} p_i^2$$

where R is the number of types and p is their relative abundance. The Gini-Simpson index is often employed, which represents the probability that the two entities belong to different types and results from $1 - \lambda$, so that higher values correspond to higher diversity (Jost 2006).

For indicators built on continuous variables, it is most appropriate to preserve the information richness of the data (Bonaccorsi et al. 2010), by using measures like the standard deviation and the mean Euclidean distance (from here on: Euclidean distance). The standard deviation measures dispersion of data by considering the square root of the average squared differences of the values from the mean. The Euclidean distance is given by the average difference (distance) between each unit in the sample; normalization occurs by dividing the Euclidean distance by the mean value of the sample.

For these indicators, some properties analysed by the inequality literature need to be discussed: scale independency, population independency, and the assumptions about the underlying distribution of the data (Litchfield 1999). Scale independency relates to the fact that ranges of values should not impact the diversity estimate. For instance, two systems that differ in HEI size along an identical distribution function should score a similar diversity value, even if in one system size ranges between 1,000 and 2,000 students and in the other between 5,000 and 10,000 students. The Euclidean distance normalized is scale independent, similar to other measures such as the Gini-Simpson coefficient—while this is not the case for the standard deviation (Litchfield 1999). In turn, the selection of the indicator should consider this property (particularly important for cross-national comparison) when the values have very different ranges of variation on an absolute scale. Arguably, scale independency may preferably be avoided when the indicator ranges on a given interval. For instance, consider an analysis of diversity of internationalization between HEIs in two countries where the share of foreign staff ranges from 1% to 6% and from 10% to 60%, respectively. The Euclidean distance normalized for the two systems would be the same, while it seems reasonable to claim that diversity is much higher in the second case.

Another property identified by the literature on inequality is that the metric should not be affected by the number of units in the population (population independence) (Litchfield 1999). This property is satisfied by the measures considered.

Choosing measures also implies adhering to different assumptions regarding the distribution of data. For instance, the standard deviation assumes a normal distribution, whereas it will be greatly inflated when applied to skewed distributions.

In sum, the selection of the measure should take into consideration their desired properties and the data at hand. In particular, when scale independency is a desirable property, the Euclidean distance normalized and the Gini-Simpson coefficient are preferable; otherwise, the Euclidean distance and the standard deviation are suitable. Standard deviation is preferable when the data follow a normal distribution, whereas the Euclidean distance and the Euclidean distance normalized are preferable when data are skewed.

3. Methodology

In order to illustrate the implications of the choices and challenges set out above, we provide a case study of diversity of European higher education based on data gathered in the EUMIDA project, a study 'to test the feasibility of a regular data collection of micro-data on HEIs' in Europe

(Bonaccorsi et al. 2010: 9). The feasibility study was partly a technical exercise, which should be seen against the background of a perceived need to gain better insight into the diversity of HEIs across Europe and their performance. The project has focused on data availability, confidentiality issues, and the resources needed to carry out a full-scale exercise. The EUMIDA data set offers a unique opportunity to compare the level of diversity across higher education systems. The data set comprises key variables relating to the core functions of HEIs (teaching, research, third mission). Eventually, data was gathered on 28 European countries (EU-27 member excluding France, plus Norway and Switzerland) and 2,457 HEIs, covering about 90% of the students enrolled in tertiary education in these countries (Lepori and Bonaccorsi 2013). We excluded two countries for which there are too few cases, and hence, measures of diversity would be unreliable, namely Luxembourg (nine HEIs) and Malta (four HEIs), leading to a sample of 26 countries.

For most of our analyses, we rely on the EUMIDA restricted data set of so-called 'research active HEIs'. The EUMIDA project adopted a multi-criteria approach, according to which an institution is considered research active if it satisfies at least three criteria out of a list of six.[1] The list of criteria was designed with the explicit goal that any combination of three or more of them would describe an institution that might sensibly be considered as systematically active in research. The sample of research active institutions comprises 58% of the HEIs, 91% of students, and 94% of academic staff in the whole EUMIDA data set, including almost all doctorate-awarding HEIs.

3.1 Operationalizing and measuring diversity

Regarding the relevant dimensions and their operationalization and measurement, we follow a generic approach (see Section 2.1) to allow for a broad understanding of the phenomenon of diversity in higher education. We focus—in line with most other research on diversity—on key dimensions of the higher education fabric, related to the central missions of HEIs: teaching, research, and third mission.

These were supplemented with the dimension of internationalization, which in the past decades has received increasing attention from scholars in higher education (Ackers and Gill 2008; Horta 2009). In fact, it has been argued that in the era of globalization, nation-states become increasingly interdependent, and internationalization missions are currently part and parcel of universities' strategies and organization. For sure, this dimension is intertwined with the existing missions of teaching, research, and the third mission (Scott, 2006). We also included structural features that are deemed particularly relevant for organizational functioning, according to traditional contingency theory as well as more recent configurational approaches (Short et al. 2008). Size is a key

characteristic of organizations, which impacts most of their activities, while legal characteristics associated with public regulation have a deep impact on organizational behaviour.

For the purpose of this article we have selected indicators that represent five dimensions in a robust way (see also Table 1): educational profile, research involvement, third mission, international orientation, and structural features. We are not arguing that this is the best possible set of dimensions and indicators, but sufficient to address the main contribution of this work, as to the selection of the variable as well as sampling decisions and its consequences for diversity measures.

Educational profile characterizes the level of the qualifications awarded and the offer of educational programmes through two indicators on:

- *Degree structure:* For each institution, the sum of the number of diplomas (short degrees) and bachelor degrees is divided by the number of total degrees (diploma, bachelor, and master). This indicator has some limitations, for there is a minority of countries that have not yet introduced an undergraduate-graduate degree model. This indicator ranges between 0 and 1 and the values follow either a right skewed distribution, with most HEIs providing mainly masters degrees (like in Germany), or left skewed distribution, with most HEIs providing bachelor and master degrees (like in Finland or the Netherlands), while no country displays a clear normal distribution.[2] As a consequence, the Euclidean distance appears to be the most suitable measure of diversity.

- *Subject areas:* We employ the Herfindahl index to measure how evenly spread the distribution of students by discipline[3] is. This indicator ranges between 1, when all students are enrolled in the same discipline, and $1/n$ when the students are homogeneously distributed across the n disciplines (Hirschmann, 1964). The values of this index follow a left-skewed distribution. This suggests that the Euclidean distance would be a good measure of diversity.

As a proxy of *research involvement* we consider the ratio between the number of PhD students (ISCED 6) and the total number of students (PhDs plus bachelor and master students), which is a widely available and used indicator of research orientation (Bonaccorsi et al. 2007). This indicator ranges between 0 and 1 and values follow a right-skewed distribution, with a high number of low or no research involvement and few highly research intensive HEIs. The Euclidean distance is the most suitable measure.

Due to data availability and methodological issues, EUMIDA does not include many indicators for *third mission*. In light of our definition focusing on the application and exploitation of knowledge, academic patents would be a suitable measure, but cross-country analyses

Table 1. Diversity indicators and measures

Dimensions	Variables	Definition	Indicator for each HEI	Diversity measure	Indicator code
Educational profile	Degree structure	Share of Diploma and Bachelor degrees of all degrees (PhD not included)	(no. of bachelor and diploma degrees)/(no. of bachelor, diploma and master degrees)	ED	BACHELOR
	Subject diversity	Distribution of students by field	Herfindahl index on the number of students by each of nine fields of educational statistics	ED	SUBJECT
Research involvement	Doctorate students	Share of ISCED 6 students (PhDs) as share of total student number	ISCED 6 / (ISCED 5 + ISCED 6)	ED	DOCTORATE
International orientation	International students	Number of foreign ISCED 5 students as share of total students	Foreign ISCED 5 / ISCED 5	ED	INT_5
	International PhD students	Number of foreign ISCED 6 students as share of total students.	Foreign ISCED 6 / ISCED 6	ED	INT_6
Knowledge exchange	Third party funding	Third party funding as share of total income	Third party / total income	ED	THIRD
Structural features	Governance	Public or private (yes/no)	0 = private (including government-dependent private). 1 = public	Gini-Simpson index (1-Simpson)	GOVERNANCE
	Size	Total number of staff	Number of staff	ED normalized	SIZE_STAFF
		Total number of students (ISCED 5 and 6)	Number of students	ED normalized	SIZE_STUDENT

are plagued with comparability issues related to differences in national regulations (Lissoni et al. 2008), while data on spin-off companies are not readily available (Mustar et al. 2006). We adopt the share of third party funding from total income as the sole indicator for this dimension. This indicator ranges between 0 and 1, and data follow a right-skewed distribution, so that the Euclidean distance is preferable to the standard deviation.

Two indicators have been chosen to cover different dimensions of the *international orientation*. The share of foreign students enrolled at ISCED 5 level focuses on attractiveness to international students, while the share of international PhD students is related to international reputation and openness, especially concerning research activities. Data on international academic staff are unfortunately available for only a few countries. These indicators range between 0 and 1, with data following a right skewed distribution in all countries, which leads us to consider the Euclidean distance.

Finally, we include three indicators concerning structural features of HEIs. *Governance* (in particular the private-public dimension) is deemed one of the key elements of this dimension. Following what has been set out in Section 2.3, we choose the Simpson index for this categorical variable. For *size*, we select two measures, one on staff numbers and the second considering student numbers. Unlike the previous measures, these indicators span on an absolute scale rather than a given range.

Following the discussion in Section 2.3, we employ the Euclidean distance normalized. Table 1 summarizes our operationalization and measures.

4. Analysis

Based on these measures, we perform three types of analyses. First, we discuss the impact of choosing perimeters on the measurement of diversity concerning organizational size. Second, using the dimension of PhD intensity, we analyse the impact of outliers on diversity measures. Third, we compare measures of diversity across different variables and we inquire about the extent to which they provide different results, in order to shed light on the question of aggregating measures.

4.1 HEI perimeters and their implications for diversity

As highlighted, there is no 'natural' definition of the perimeter for the analysis of higher education diversity. While HEIs can be expected to share some common features and functions (like offering education at ISCED 5 and/or 6 level), the precise delineation of which kind of educational institutions are considered as part of a higher education system is a complex socio-cultural and political process. The EUMIDA data set provides an opportunity to assess robustness of diversity measures against the definition of the perimeter. To analyse the impact of the choice

Table 2. Diversity of size of HEIs (number of students)

Country	Number of HEIs			ED normalized		Diversity rank	
	All	Research active (RA)	Coverage (%)	All	RA	All	RA
AT	68	68	100	1,5	1,5	3	1
CH	36	36	100	1,2	1,2	15	6
NL	59	59	100	1,1	1,1	18	10
SK	33	33	100	1,1	1,1	19	11
BE	85	85	100	0,9	0,9	20	19
FI	49	49	100	0,9	0,9	21	20
IE	21	21	100	0,8	0,8	23	22
ES	47	47	100	0,7	0,7	24	24
UK	149	147	99	0,8	0,8	22	21
SE	49	41	84	1,2	1	14	16
HU	71	58	82	1,4	1,3	9	2
DE	378	306	81	1,3	1,2	13	5
RO	82	57	70	1,3	1,2	12	4
NO	68	46	68	1,3	1,1	11	9
GR	60	40	67	1,3	1	10	15
BG	59	39	66	1,1	1	17	17
CZ	73	26	36	1,5	1	2	12
LT	46	16	35	1,1	0,9	16	18
IT	243	81	33	1,6	1,1	1	7
LV	61	20	33	1,4	1	8	14
EE	34	7	21	1,4	1,1	7	8
PL	457	91	20	1,4	1	6	13
CY	37	7	19	1,4	1,2	5	3
PT	138	18	13	1,4	0,7	4	23

Denmark is not included in the Research Active database; also Slovenia was excluded (only four research active HEIs).

Table 3. Diversity in research involvement: ratio of PhD to PhD plus undergraduate students

	Research active HEIs			Research active HEIs without outliers			
	N	ED	Rank	N	Coverage (%)	ED	Rank
CZ	26	0,19	1	26	100	0,19	1
CH	36	0,08	4	36	100	0,08	2
BG	39	0,07	5	39	100	0,07	3
FI	49	0,06	6	49	100	0,06	4
GR	40	0,05	8	40	100	0,05	5
DK	14	0,04	11	14	100	0,04	10
NO	44	0,04	12	44	100	0,04	11
RO	57	0,04	13	57	100	0,04	12
SE	41	0,04	14	41	100	0,04	13
SK	33	0,04	15	33	100	0,04	14
ES	47	0,03	16	47	100	0,03	15
HU	61	0,03	17	61	100	0,03	16
IE	21	0,03	18	21	100	0,03	17
PT	109	0,03	19	109	100	0,03	18
EE	7	0,02	20	7	100	0,02	20
LV	20	0,02	21	20	100	0,02	21
NL	59	0,02	22	59	100	0,02	22
PL	75	0,02	23	75	100	0,02	23
LT	16	0,01	24	16	100	0,01	24
DE	306	0,04	10	303	99	0,04	9
UK	148	0,05	9	146	99	0,05	6
AT	68	0,05	7	67	99	0,04	8
IT	81	0,11	3	76	94	0,02	19
BE	43	0,11	2	40	93	0,04	7

Cyprus and Slovenia excluded (only three doctorate-awarding HEIs).

of perimeter, we compare diversity indexes and rank for the full and restricted perimeter (Table 2).

Overall, sensitivity to the perimeter is rather moderate (with a correlation of 0.60 for the values and 0.55 for the rank positions), taking into account that the full sample includes 70% more HEIs than the restricted sample. However, when the perimeter is inflated by a large number of very small HEIs, like in Poland or Portugal, then the impact is significant. The findings confirm that perimeter issues are indeed central to measuring diversity, meaning differences across countries need to be considered when interpreting results.

4.2 The impact of outliers

Analysing diversity scores in doctoral education allows one to consider the impact of outliers on diversity measures. The indicator we use is the ratio between the number of PhDs and the total number of undergraduate and PhD students. As this is a continuous variable based on a ratio from 0 to 1, we use non-normalized Euclidean distance. Table 3 provides scores considering all research-active HEIs in each country and scores excluding outliers from the sample.

Descriptive statistics show that the ratio of PhD to total students for the entire sample is highly skewed (mean = 0.04, standard deviation = 0.08) and country means are widely different. There are 17 HEIs in the sample for which PhD students are more than half of the total students: these include five Italian graduate schools, three Belgian academies (university consortia focused on graduate education), one UK cancer research institute, five Czech universities, the Bulgarian Academy of Science, and two small HEIs in Austria and Germany.

For most countries, no outliers were identified, so the overall correlation between country-level diversity for the two samples is high (0.81 for the values and 0.86 for the ranking positions). On the other hand, for the countries where outliers exist, they have a strong impact on diversity scores and rank. For instance Belgium and Italy, respectively the second and third most diverse system when outliers are included, drop to the 7th and 19th position in the ranks when outliers are excluded.

The measure without outliers is more consistent with findings than systems with a rather strong binary divide (Finland and Switzerland), and competitive unitary systems (UK) are more diverse (Huisman et al. 2007). But from a theoretical standpoint there is no principle

Table 5. Correlation between diversity measures

	Bachelor	Subject	Doctorate	Third	Int_5	Int_6	Governance	Size_student	Size_staff
Subject	0,50*								
Doctorate	0,25	0,20							
Third	−0,09	0,22	−0,04						
Int_5	0,15	0,24	0,47*	0,20					
Int_6	0,13	−0,35	0,46*	−0,06	0,65**				
Governance	0,00	0,46*	−0,28	0,06	0,34	−0,09			
Size_ student	0,48*	0,70**	0,04	0,10	0,43	−0,11	0,63**		
Size_ staff	0,25	0,80**	0,24	0,16	0,64**	−0,02	0,56*	0,77**	

*Correlation is significant at the 0.05 level (two-tailed).
**Correlation is significant at the 0.01 level (two-tailed).

Denmark, Belgium and Hungary, and a group of countries displaying low diversity, including Lithuania, Spain, Greece, Ireland, the Czech Republic and Portugal. This largely corresponds to the divide between Northern and Western European countries on the one hand, and Mediterranean and Eastern European countries on the other hand. Interestingly, there are various systems (Finland, the Netherlands, Latvia, Romania and the UK) that are both diverse (top-5 positions for one or more indicators) and fairly homogeneous (bottom-5 positions for one or more indicators) at the same time. This might point to the fact that governmental policies for differentiation were aimed at specific activities, like selective research findings in the UK.

Second, we compute pairwise correlations between the country-level diversity measures presented in Table 5. Expectedly, the two size indicators are highly correlated; they are also rather strongly correlated with diversity concerning governance, internationalization of undergraduate students and level of education. Most of these correlations can be readily interpreted. For example, private HEIs are typically smaller and more specialized in specific subjects than public ones and, therefore, systems with high levels of governance diversity will also display higher levels of diversity in size and subject profiles. A second group of associated variables are the two internationalization variables and the research orientation variable (doctorate), which shows that research orientation and internationalization are closely associated. On the contrary, there are no significant associations between the variables measuring diversity in education (bachelor and subject), research (doctorate), and third mission (share of third party funding). This suggests that they are independent dimensions of HEI activities.

We stress that the results above are not the definitive results of an encompassing research project on institutional diversity in these countries. The findings were presented to show patterns based on fairly robust data and to explore correlations. We included an overall diversity rank to actually highlight the potential dangers of aggregating the data: many relevant details—particularly the fact that many countries somewhere in between high and low overall diversity display levels of extreme low and high diversity at the level of specific indicators.

The analysis shows that contrasting different measures is helpful in order to identify independent dimensions across which HE diversity might unfold. Identifying main dimensions would allow for constructing composite indicators, for example, combining measures concerned with internationalization, thus adding to the robustness and parsimony of the analysis.

5. Conclusion

This article discussed the challenges of operationalizing and measuring diversity in higher education. We offer an arguably robust and valid approach, focusing on key dimensions and meaningful variables that largely coincide with those employed in earlier research. We used data from the majority of European HEIs and systems to present a sensitivity analysis that showed that diversity studies should clarify in advance the criteria used, and explicate decisions regarding the perimeter and the inclusion/exclusion of outliers. Measurements should also be accompanied by thorough data inspection. Our analysis yielded the following. First, perimeter issues are central to measuring diversity, both from a conceptual and statistical perspective, but in practice, diversity measures are rather robust when compared to perimeter changes. Second, distribution properties matter, and we argued that the measurement of diversity is vulnerable to the choice of (statistical) diversity measures. Third, outliers do significantly affect measurement outcomes, especially when using continuous measures. Including outliers (or not) is however not only a technical decision, but largely depends on the nature of and framework for analysis. Fourth, we found differences and similarities for different dimensions of diversity, which shows the relevance of analysing the interdependencies between the dimensions. Overall, our conclusion is that there is no textbook recipe for how to measure diversity, but the selection of dimensions, variables and analytical methods must be seen

in the context of the specific goal of the study and its analytical framework. Good knowledge of the characteristics of measures and variables and their contextualization in the specific empirical setting considered are important issues in this respect.

Given the many anecdotic and simplified approaches in past research on diversity, we argue that our study is an important methodological step forward in analysing diversity in higher education, and towards gaining significant empirical insight in the levels of diversity of higher education systems, acknowledging that other methods are available as well (e.g. Schubert 2014; Huisman 2000).

Our work has some limitations, which point at future areas for research and sophistication. The preference given to measurement and measurable indicators led to a focus on institutional activities and outcomes. 'Softer' dimensions of diversity could not (yet) be considered (e.g. organizational climate and culture). The data did not include sophisticated measures of outputs, such as data on publications or research grants, nor longitudinal information that could be used to detect patterns and causal connections over time. Existing data sets can be expanded by including bibliometric data from sources like the Leiden ranking (Waltmann et al. 2012) or the SCIMAGO institutional ranking and participation in European programmes, which would provide an important addition to the dimensions of research and internationalization.

These limitations are useful to identify the next research steps. First, by complementing the existing data with information on the systems' configuration, it is possible to investigate what factors explain lower and higher levels of diversity, for instance by considering differences between system's binary-unitary arrangement and the level of competition (Lepori and Seeber 2011) or the role of specific governmental policies (Lepori et al. 2014). Moreover, diversity is not solely determined by environmental and system factors, as universities increasingly act as strategists and their choices can have an important direct or indirect impact on the level of system diversity (Huisman 2000). Further research should, therefore, also be oriented towards a fine-grained understanding of these micro-level dynamics and the interaction with external pressures (e.g. Fumasoli and Lepori 2010; Frølich et al. 2012; Fumasoli and Huisman 2013).

Funding

Acknowledgement

Notes

1. Criteria for inclusion were: (i) the existence of an official research mandate; (ii) the existence of research units institutionally recognized (for example, on the institutional Web site); (iii) the inclusion in the R&D statistics (availability of R&D expenditure data), as sign of institutionalised research activity; (iv) awarding doctorates or other ISCED 6 degrees; (v) consideration of research in institutions strategic objectives and plans; (vi) regular funding for research projects either from public agencies or from private companies.
2. Some countries, instead, display non-skewed distributions, such as Spain with a rather flat distribution, and Italy following a quasi-normal distribution.
3. Education; humanities and arts; social science, business, and law; science; engineering, manufacturing, and construction; agriculture; health and welfare; and services.

References

Ackers, L. and Gill, B. (2008) *Moving People and Knowledge— Scientific Mobility in an Enlarging European Union*. Cheltenham: Edward Elgar

Birnbaum, R. (1983) *Maintaining Diversity in Higher Education*. San Francisco, CA: Jossey-Bass.

Bonaccorsi, A. *et al.* (2010) *Feasibility Study for Creating a European University Data Collection*. Final study report. Brussels: European Commission.

——. (2007) 'Indicators on Individual Higher Education Institutions: Addressing Data Problems and Comparability Issues', *Research Evaluation*, 16/2: 66–78.

Bonaccorsi, A. and Daraio, C. (eds) (2007) *Universities and Strategic Knowledge Creation. Specialization and Performance in Europe*. Cheltenham: Edward Elgar.

Daraio, C. *et al.* (2011) 'The European University Landscape: A Micro Characterization Based on Evidence from the Aquameth Project', *Research Policy*, 40/1: 148–64.

Frølich, N. *et al.* (2012) 'A Reinterpretation of Institutional Transformations in European higher Education: Strategising Pluralistic Organisations in Multiplex Environments', *Higher Education*, 65/1: 79–93.

Fumasoli, T. and Huisman, J. (2013) 'Strategic Agency and System Diversity: Conceptualizing Institutional Positioning in Higher Education', *Minerva* 51/2: 155–69.

Fumasoli, T. and Lepori, B. (2010) 'Patterns of Strategies in Swiss Higher Education', *Higher Education*, 61/2: 157–78.

Hirschman, A. O. (1964) 'The Paternity of an Index', *The American Economic Review*, 54/5: 761.

Horta, H. (2009) 'Global and national prominent universities: internationalization, competitiveness and the role of the State', *Higher Education* 58(3): 387–405.

Huisman, J. (2000) 'Higher Education Institutions: as Different as Chalk and Cheese?', *Higher Education Policy* 13: 41–53.

Huisman, J. and Kaiser, F. (eds) (2001) *Fixed and Fuzzy Boundaries in Higher Education. A Comparative Study of (binary) Systems in Nine Countries*. The Hague: AWT.

Huisman, J. *et al.* (2007) 'Institutional Diversity in Higher Education: A Cross-National and Longitudinal Analysis', *Higher Education Quarterly* 61/4: 563–77.

Huisman, J. and Morphew, C. C. (1998) 'Centralization and Diversity: Evaluating the Effects of Governmental Policies

in U.S.A. and Dutch Higher Education', *Higher Education Policy* 11/1: 3–13.

Huisman, J. and van Vught, F. (2009) 'Diversity in European Higher Education: Historical Trends and Current Policies', In: van Vught, F. (ed.) *Mapping the Higher Education Landscape. Towards a European Classification of Higher Education*, pp. 17–37. Dordrecht: Kluwer.

Jencks, C. and Riesman, D. (1968) *The Academic Revolution*. Garden City: Doubleday.

Jost, L. (2006) 'Entropy and Diversity', *Oikos* 113/2: 363–75.

Kyvik, S. (2004) 'Structural Changes in Higher Education Systems in Western Europe', *European Journal of Education* 29/3: 393–409.

Lepori, B. and Bonaccorsi, A. (2013) 'The Socio-political Construction of a European Census of Higher Education Institutions: Design, Methodological and Comparability Issues', *Minerva* 51: 271–93.

Lepori, B. *et al*. (2014) 'Convergence and differentiation Processes in Swiss Higher Education: An Empirical Analysis', *Studies in Higher Education* 39/2: 197–218.

Lepori, B. and Seeber, M. (2011) *'Forces Shaping the Structure and Diversity of Populations of Higher Education Institutions in Europe'*. Paper presented at the 22nd Annual CHER conference. Reykyavik, Iceland, 23-25 June.

Lissoni, F. *et al*. (2008) 'Academic Patenting in Europe. New Evidence from the KEINS Database', *Research Evaluation* 17/2: 87–102.

Litchfield, J. A. (1999). *Inequality: Methods and Tools*. New York: Worldbank. <http://icm.clsbe.lisboa.ucp.pt/docentes/url/analeco2/Papers%5cLitchfield_Inequality_Methods_Tools.pdf> accessed 3 August 2015.

McKelvey, B. and Aldrich, H. (1983) 'Populations, natural selection, and applied organizational science', *Administrative Science Quarterly*, 28, 101–28.

Meek, V. L. (1991) 'The Transformation of Australian Higher Education from Binary to Unitary System', *Higher Education* 21: 461–94.

Morphew, C. C. (2009) 'Conceptualizing Change in the Institutional Diversity of U.S. Colleges and Universities', *Journal of Higher Education* 80/3: 243–69.

Mustar, P. *et al*. (2006) 'Conceptualising the Heterogeneity of Research-Based Spin-Offs: A Multidimensional Taxonomy', *Research Policy* 35/2: 289–308.

Neave, G. (1979) 'Academic Drift: Some views from Europe', *Studies in Higher Education* 4/2: 143–59.

Pianka, E. R. (1994) *Evolutionary Ecology*. New York: Harper & Row.

Reichert, S. (2009) *Institutional Diversity in European Higher Education. Tensions and Challenges for Policy Makers and Institutional Leaders*. Brussels: European University Association.

Rossi, F. (2009) 'Increased Competition and Diversity in Higher Education: An Empirical Analysis of the Italian University System', *Higher Education Policy* 22/4: 389–413.

——. (2010) 'Massification, Competition and Organizational Diversity in Higher Education: Evidence from Italy', *Studies in Higher Education* 35/3: 277–300.

Ruef, M. (2000) 'The Emergence of Organizational Forms: A Community Ecology Approach', *American Journal of Sociology* 106/3: 658–714.

Short, J. C. *et al*. (2008) 'Research on Organizational Configurations: Past Accomplishments and Future Challenges', *Journal of Management* 34/6: 1053–1079.

Schubert, T. *et al*. (2014) 'Is there a European University Model? New Evidence on National Path Dependence and Structural Convergence'. In: Bonaccorsi A. (ed.) *Knowledge, Diversity and Performance in European Higher Education: A Changing Landscape*, pp. 47–83. Cheltenham: Edward Elgar.

Scott, J. C. (2006) 'The Mission of the University: Medieval io Postmodern Transformations', *Journal of Higher Education* 77/1: 1–39.

Skolnik, M. L. (1986) 'Diversity in Higher Education: The Canadian Case', *Higher Education in Europe* 11: 19–32.

Stanley, G. and Reynolds, P. (1994) 'Similarity Grouping of Australian Universities', *Higher Education* 27: 359–66.

Stirling, A. (2007) 'A General Framework for Analysing Diveristy in Science, Technology and Society', *Journal of the Royal Society Interface*, 4: 707–719.

Teixeira, P. N. *et al*. (2012) 'Competition and Diversity in Higher Education: An Empirical Approach to Specialization Patterns of Portuguese Institutions', *Higher Education* 63/3: 337–52.

Tight, M. (1988) 'Institutional Typologies', *Higher Education Review* 20: 27–51.

——. (2007) 'Institutional Diversity in English Higher Education', *Higher Education Review* 39/2: 3–24.

——. (2011) 'How many Universities are there in the United Kingdom? How many Should There Be?', *Higher Education* 62: 649–63.

Van Vught, F. (ed.) (2009) *Mapping the Higher Education Landscape: Towards a European Classification of Higher Education*. Dordrecht: Kluwer.

Van Vught, F. and Westerheijden, D. F. (2010) 'Multidimensional Ranking: A New Transparency Tool for Higher Education and Research', *Higher Education Management and Policy* 22/3: 1–26.

Waltman, L. *et al*. (2012) 'The Leiden Ranking 2011/2012: Data Collection, Indicators, and Interpretation', *Journal of the American Society for Information Science and Technology* 63/12: 2419–32.

Zha, Q. (2009) 'Diversification or Homogenization: How Governments and Markets have Combined to (re)shape Chinese Higher Education in its Recent Massification Process', *Higher Education* 58: 41–58.