Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Dipartimento Economia e Statistica Cognetti de Martiis

Working Paper Series

18/18

SCIENTIFIC OUTPUT OF US AND EUROPEAN UNIVERSITIES SCALES SUPER-LINEARLY WITH RESOURCES

BENEDETTO LEPORIA, ALDO GEUNA and ANTONIETTA MIRA

brick Bureau of Research on Innovation, Complexity and Knowledge

UNIVERSITÀ DEGLI STUDI DI TORINO

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Scientific output of US and European universities scales super-linearly with resources

Authors: Benedetto Lepori[a*], Aldo Geuna[b,c], Antonietta Mira[d]

Affiliations:

a: Faculty of Communication Sciences, Università della Svizzera italiana, Via Lambertenghi, 10, Lugano, 6900 (Switzerland), blepori@usi.ch (corresponding author).

b: Department of Economics and Statistics Cognetti De Martiis, University of Turin, Lungo Dora Siena 100 A, 10153 Turin, Italy, Tel: +39 0116703891, Fax: +39 011 6703895; email: aldo.geuna@unito.it

c: BRICK, Collegio Carlo Alberto, Via Real Collegio, 30, Moncalieri Torino, 10024 (Italy)

d: Institute of Computational Sciences, Faculty of Economics, Università della Svizzera italiana, Via Buffi, 13, Lugano, 6900 (Switzerland), antonietta.mira@usi.ch.

Classification: Social Sciences/Economic Sciences

Abstract:

By using a comprehensive dataset of US and European universities, we demonstrate super-linear scaling between university revenues and their volume of publications and citations. We show that this relationship holds both in the US and in Europe. In terms of resources, our data show that three characteristics differentiate the US system: (1) a significantly higher level of resources for the entire system, (2) a clearer distinction between education-oriented institutions and doctoral universities and (3) a higher concentration of resources among doctoral universities. Accordingly, a group of US universities receive a much larger amount of resources and have a far higher number of publications and especially citations when compared to their European counterparts. These results demonstrate empirically the pervasiveness of a social order where financial resources are tightly coupled with a measure of 'excellence' associated with international rankings and, additionally, where the widely accepted measures of 'excellence' in reality 'prime' resources. They therefore raise important questions for policy-making and for the management of higher education institutions.

Significance statement:

Thanks to a unique dataset, the paper provides novel empirical evidence on a) the relationships between university budgets and their scientific output and visibility and b) on differences in the structure of higher education and in the distribution of resources between US and Europe. For the first time, these results show how pervasive is the tight coupling between financial resources and international 'excellence' as measured by international rankings. Second, they show how widely used measures of ‘excellence’ are closely associated with the amount of resources, raising therefore concerns on their value in measuring research ‘quality’. At the policy and HEI level, these findings also suggest that fighting for top-ranking positions should not be the main concerns for policymakers and for most university managers.

1 Introduction

An extensive amount of literature has documented scaling properties that exist in the science system, including the distribution of scientific citations (Peterson, Presse and Dill 2010), the structure of scientific networks (Barabási and Albert 1999), the relationships between publications and citations at the country level (Katz 1999) and in cities (Nomaler, Frenken and Heimeriks 2014). These relationships have been frequently approximated with power-law distributions (Newman, Mark EJ 2005; Leitao, Miotto, Gerlach and Altmann 2016). A super-linear power-law relationship with an exponent above one between the volume of publications and citations has been observed for the 500 largest universities worldwide (van Raan 2013). Super-linear scaling implies that the average number of citations per paper at the university level, increases more than linearly with the volume of scientific production and, therefore, universities with a larger volume of output will also appear at the top of international research rankings, which are heavily correlated with bibliometric measures.

Unlike parallel literature on cities, where scaling is measured against populations, or the volume of economic production (Bettencourt 2013), the literature on science scaling does not rely on a consistent measure of resources. Yet, the economics literature argues that scientific production is affected by the volume and distribution of resources (Partha and David 1994) and, specifically, suggests that the dominance of US universities in international rankings is due to a better-funded and more competitive funding system (Aghion, Dewatripont, Hoxby, Mas-Colell and Sapir 2010).

We contribute to this debate by first demonstrating that there is a striking statistical regularity in the number of university publications and citations that scale super-linearly in respect to the volume of resources, and that these relationships are similar in the US and in Europe.

Second, we show fundamental differences in the distribution of revenues within the two systems. Accordingly, the US system includes a number of universities with a far larger budget than their European counterparts, which are also at the top of international rankings (Bonaccorsi, Cicero, Haddawy and Hassan 2017; King 2004).

Third, we speculate on some interpretations of our findings in light of the literature on cumulative effects in science (Stephan, Paula E. 1996) and on the impact of rankings (Hazelkorn, E. 2009; Deem, Mok and Lucas 2008a). We interpret these results as an indication of how pervasive have become 'universal' measures of research 'excellence', such as those conveyed by international rankings (Sauder and Espeland 2009) and we suggest that these underlying mechanisms lead to the observed tight coupling between resources, research output and visibility (irrespectively from size and region).

Fourth, our findings on super-linear scaling add a further worrisome dimension to this debate as they show that widely accepted measure of 'excellence' in reality prime the richest universities, therefore further contributing to cumulative effects (Abramo and D'Angelo 2016) and proving misleading signals to university managers (Paradeise and Thoenig 2013).

The reminder of the paper is structured as follows. In Section 2, we provide a detailed description of the new database that includes comparable information on 5,551 Higher Education Institutions (HEIs) in the US and in Europe. Section 3 presents the methods and results for the analysis of supra-linear scaling. The results are contextualized in Section 4 where we analyze comparatively the differences in the level and distribution of resources between the US and Europe. In the last section we speculate on why such strong correlation might exists and we offer a few conclusions on the implication for policy making and management of HEIs.

2 The data

We have created a matched database composed of 3,287 HEIs in the US and 2,264 HEIs in Europe, which represents the population of HEIs delivering degrees at least at the bachelor level.

Our database is derived from the Integrated Postsecondary Education Data System for the US (IPEDS; http://nces.ed.gov/ipeds/) and the European Tertiary Education Register dataset (ETER; www.eter-project.com). Both datasets provide information on HEIs, defined as institutions delivering degrees at the tertiary level, corresponding to levels 5 to 8 of the International Standard Classification of Educational Degrees (ISCED; http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx). Since ETER does not adequately cover HEIs delivering only short tertiary degrees (ISCED 5), we limit our sample to the HEIs delivering at least a bachelor (ISCED 6), therefore excluding associate colleges in the US. When compared with international students statistics from EUROSTAT and OECD, the coverage of student enrolments at levels 6 (bachelor), 7 (master) and 8 (PhD) in the US is 100% and 96% in Europe.

2.1 Variables

Table 1 provides an overview of all variables that are used for this paper. The methodological and comparability issues for financial, staff and bibliometric variables are discussed in more detail below. Information on how the other variables have been mapped (for example, subject fields) is available on request.

Data availability is fairly good for all variables, except for European budget data. However, availability is much better for doctoral universities and for larger HEIs and, therefore, missing data have a limited impact on the regression analysis (see section 2.5 below). Only specific analyses on the composition of funds for European HEIs have to be taken some care.

Table 1 Variables used in the paper and number of available cases

Variable Name	**Definition**	Valid cases US	Valid cases Europe
Highest Degree Delivered*	0= ISCED 5 diplomas with duration of less of three years; 1=ISCED 6 bachelor (3 or 4 years); 2=ISCED 7 master or equivalent diploma in the pre-Bologna system (for example 4/5 years license); 8=qualification equals doctorate.	3,287	2,191
PhD Awarding	1= HEI has the right to deliver the PhD, 0 otherwise	3,287	2,191
Legal Status	0 public institutions (IPEDS = public, ETER= public or private government-dependent); 1 private institutions (IPEDS = private for profit or private non-profit, ETER= private)	3,287	2,262
Staff	ETER = academic staff. IPEDS = instructional, research and public service staff, both in Full Time Equivalents	3,195	1,719
Enrolments	Total number of students enrolled at levels ISCED5-8	3,150	2,264
Publications	Publications count (Web of Science)	3,287	2,264
Citations	Field normalized citations count	3,287	2,264
Subject mix*	Herfindahl index of the distribution of students by field using the fields of education and training classification	3,056	2,003
Total current revenues	Total Revenues euro PPP (excluding hospital revenues and subsidiaries)	3,062	1,316
Basic state instalment	State allocation to the general university budget in EURO PPPs	2,570	606
Private donations and endowment	Private donations to the general university budget and revenues from the endowment in EURO PPPs	2,570	606
Third Party Funding	Public and private contracts, including those from public agencies (NSF etc.) in EURO PPPs	3,062	1,115
Student fees funding	Fees paid by students, including also indirect state support (for example loans) in euros PPPs	3,062	1,154
*these variables are used to classify HEIs using the Carnegie classification criteria			

2.2 Revenue and staff variables

Our main variable is *Total current revenues*. For a more fine-grained analysis, revenues were divided into four streams:

- *Basic state instalment*, i.e. the funds provided by the state for the general functioning of the HEI.
- *Private donations and payouts from the endowments*. These funds are usually managed at the university level (even if a share might be earmarked to specific activities).

- *Third-party funds* mostly for research, like research grants from public funding agencies and contracts from companies.
- *Funding from student fees* paid by students and families.

The construction of these data has to deal with differences in how the perimeter of revenues is constructed and in accounting systems between types of HEIs (specifically public vs. private) and between US vs. Europe (as well as national differences in Europe).

To limit comparability problems, we devised the following strategy:

- first, we considered only *current revenues*, therefore excluding capital income, like state contributions for facilities and buildings. Besides data availability issues, this avoids a major comparability problem due to different treatments of capital costs and revenues depending on the university accounting system (cash vs accrual-based) and depending on whether university facilities are owned by the university or by the state. Investment income (for example interests on assets) are however included.
- Second, we excluded *healthcare revenues* and the *revenues from ancillary enterprises*. Following EUROSTAT definitions, these are generally excluded in Europe (the main exception was Germany, where healthcare revenues were excluded manually), while in the US data these are singled out in two specific subcategories and were be excluded; this is particularly important since in some US universities sales and services from auxiliary enterprises, like intercollegiate athletics, and medical care in hospitals constitute a large share of revenues.
- Third, we devised a detailed mapping scheme based on the subcategories of revenues provided by IPEDS and ETER (Table 2). Such a disaggregated approach allows a more precise control of the revenue perimeter and of comparability problems. While there might be residual comparability problems within categories, this mapping suggests that on the whole comparability between countries and accounting systems is good (as compared with the size of the observed effect).

Table 2. Mapping scheme for HEI revenues

<table>
<tr><th>ETER Variable</th><th>IPEDS public HEIs</th><th>IPEDS private non profit HEIs</th><th>IPEDS private HEIS</th><th>ETER</th></tr>
<tr><td rowspan="3">Core Budget (public)</td><td colspan="2">Federal Appropriations</td><td></td><td rowspan="3">Basic government allocation (central or regional)</td></tr>
<tr><td colspan="2">State Appropriations</td><td></td></tr>
<tr><td>Local Appropriation, Education District taxes, and Similar Support</td><td>Local Appropriations</td><td></td></tr>
</table>

	Federal Non Operating Grants			
	State Non Operating Grants			
	Local Non Operating Grants			
Core budget (private)	Gift (Including Contributions from affiliate Organizations)	Private gifts	Other Revenues	Gifts and donations
	Other Revenues and Additions			
	Other Non Operating Revenues			
	Sales and Services for Education Activities			
	Other sources - operating			Interests
	Investment Income	Investment Return	Investment Income and Investment gain (losses) included in net income	Investment income
Tuition Fees	Tuition and Fees, after deducting discounts and allowances	Tuition and Fees	Tuition and Fees	Tuition and Fees
Third Party	Federal Operating Grants and Contracts	Federal grants and contracts	Federal Appropriations, Grants and Contracts	Public grants and contracts (central, regional ,local)
	State Operating Grants and Contracts	State grants and contracts	State and Local Appropriations, Grants and Contracts	
	Local Operating Grants and Contracts	Local grants and contracts		Grants and contracts from abroad
	Private Operating Grants and Contracts	Private gifts, grants, and contracts	Private Grants and Contracts	Private grants and contracts
Unclassifed Revenues	Other sources operating revenues	Other revenues	Other revenues	Other revenues
Excluded	Sales and Services of Hospitals	Hospital revenues		Healthcare revenues
	Independent Operations Revenue			
	Sales and Services of Auxiliary Entreprises	Contributions from affiliated entities	Sales and Services of Auxiliary Enterprises	Sales and Services of Auxiliary Enterprises

We use Purchasing Power Parities in euros from Eurostat, as they are likely to provide a more accurate comparison of general costs in each country. Since PPPs for the US are below one (1 US $ = 0.734 euros), this choice somewhat reduces funding level differences between the US and Europe.

Academic staff in Full Time Equivalents are based on working contracts; in ETER, we include the personnel involved in teaching and research, while in IPEDS, we use the number of instructional, research and public service staff as the nearest equivalent. In both cases, it excludes management, technical and support staff, as well as healthcare staff in the hospitals annexed to universities. Coverage of PhD students and postgraduate staff may not be fully complete. However, when using FTEs, this is less of a concern if part-time staff is not fully covered.

2.3 Scientific output

Indicators on publications and citations were derived from the Leiden Ranking (LR) database (Waltman, Calero-Medina, Kosten, et al 2012), which is based on extensive cleaning of the data from the Web of Science. First, we searched in the LR for the HEIs in ETER and IPEDS; second, we looked in both datasets to match additional candidates focusing on HEIs with a sizeable number of PhD degrees for which there was no match. Publication data were retrieved for 851 HEIs in ETER, which included 97.2% of the PhD degrees in the dataset, and for 421 HEIs in IPEDS, corresponding to 89.8% of the PhD degrees. The lowest coverage for the US is mostly due to two private distance universities, which enrol a large number of PhD students, but few publications (Capella University and Warden University).

The Leiden ranking includes substantial effort to delineate the perimeter of universities and to handle special cases like assigning publications correctly to members of confederate universities (e.g. University of London). A concern has been a correct attribution of publications in medicine, as the delineation between universities and hospitals might affect comparability. To this aim, the publication of hospitals has also been assigned to the respective university if at least one author displays a strong collaboration link to the university (e.g. a large share of publications, which are also affiliated to the university; Waltman, Calero-Medina, Kosten, et al 2012).

For the purposes of our study, we use two variables:

- The total number of publications in the Web of Science for the period between 2012-2015 using fractional counting. While fractional counting at the level of individual HEIs might be affected by the share of co-publications, the correlation between full and fractional counting indicators is above 0.9 and therefore fractional counting is unlikely to affect comparability on a large sample (Waltman and van Eck 2015).

- The total field normalized citation score, i.e. the sum of the citation score of each publication normalized to the average mean number of citations of all publications in that field. This measure takes into account, to some extent, disciplinary differences in the number of citations.
- Data follow the same methodology used in the Leiden ranking system. Only core publications in the Web of Science are included, specifically the count does not include conference proceedings publications and book publications. This is an important limitation in certain research fields, especially in computer science, engineering, and the social sciences and humanities. This is however unlikely to cause bias in system-wide comparisons, particularly since we exclude specialised universities from our analysis.

2.4 Applying the Carnegie classification

The Carnegie classification provides a well-recognized classification of US HEIs (http://carnegieclassifications.iu.edu/). We use the Carnegie classification for two purposes: (a) to provide a comparative analysis of institutional diversity and (b) to identify a subset of research universities on which to focus our analysis. To this aim, we replicate the criteria of the Carnegie classification on our dataset as follows:

- Research universities. HEIs with at least 20 ISCED 8 degrees in the year.
- Masters' colleges and universities. HEIs with less than 20 ISCED8 degrees and at least 50 ISCED 8 degrees.
- Baccalaureate colleges. HEIs with more than 50% of the degrees at level ISCED 6 and with less than 50 ISCED 7 degrees or less than 20 ISCED 8 degrees.
- Baccalaureate/associate colleges. HEIs which can award ISCED 6 degrees, but where more than 50% of the degrees are at level ISCED 5.
- Special focus institutions. HEIs, for which the Herfindal index of the distribution of degrees by educational field (using the OECD-UNESCO fields of education classification) is larger than 0.7, implying that at least 80% of the degrees are in a single educational field.

For the US, these criteria classify 75% of the HEIs in the same category as the 2010 Carnegie classification, with most of the difference being accounted for by a slightly different delimitation of subject fields, by the fact that the Carnegie classification allows for a number of exceptions from the above criteria, and by the fact that our data refer to 2013.

2.5 The doctoral universities sample

For the analysis in this paper, from the US-Europe-HEI database, we have extracted a subpopulation of 936 doctoral universities corresponding to the definition in the US Carnegie classification, i.e. more than 20 PhD degrees in the reference year 2013 (http://carnegieclassifications.iu.edu/).

The subsample of research universities is composed of 570 universities in Europe and 366 universities in the US, and it includes 307 out 340 HEIs in the top-500 of the Shanghai ranking (2015 edition), with most of remaining cases being focused HEIs. US universities include 286 of the 297 research universities in the 2010 Carnegie classification, in addition to 58 colleges and 16 focused institutions. While the latter can be explained by a different delimitation of subject fields, the former inclusion can be explained by a number of masters' colleges and universities attaining the threshold for PhD degrees –the 2015 edition of the Carnegie classification counts 329 research universities.

543 of 570 research universities in Europe and 344 out of 366 in the US could be identified in the Leiden Ranking. Therefore, the coverage of research universities is almost complete.

The sample for regressions drops from 887 to 762 universities due to missing revenue data for a number of European universities; these observations however still enrol 83% of the students in all research universities and include 282 out of 307 research universities covered by the Shanghai ranking.

3 Analyzing superlinear scaling

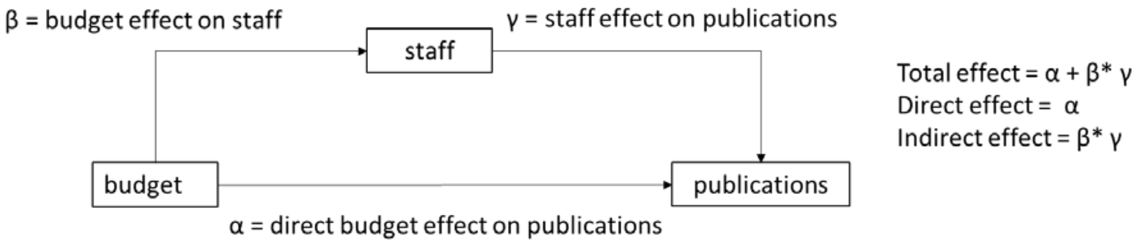
3.1 Methods

The standard approach for fitting power-law relationships is to use an OLS regression on the log-transformed variables and to provide an analysis of residuals to check whether there are potential robustness issues. In general, in our data, the conditions for an OLS regression providing unbiased coefficients are not met (Leitao, Miotto, Gerlach and Altmann 2016).

Our main regression approach relies on a method based on the assumption that the error term of the regression follows a mixture of normal distributions (Bartolucci and Scaccia 2005). This approach has been shown to provide more reliable estimates than OLS when errors are not normally distributed.

Expectedly, descriptive statistics shows strong correlations between university revenues and the number of academic staff; in order to disentangle the association between these variables and with research output, we run a mediation model that allows estimating the two paths that associate budget and publications (respectively citations) directly or through the number of academic staff (Figure 1).

Figure 1. Mediation model



Finally, we perform quantile regressions for two purposes (Koenker and Hallock 2001). First, they are more robust against non-normality than OLS and, therefore, provide a further robustness test of our results. Second, being conditional to the median of the dependent variable, they allow investigating whether the observed effects differ by the levels of the dependent variable. This might provide useful information for the development of institutional policies adapted to the specific position of individual universities.

3.2 Descriptive statistics

Table 3 provides descriptive statistics and the correlation table for the main variables used in the regression and for the doctoral universities sample.

Table 3. Descriptive statistics and correlation table for the doctoral universities sample

	N	Mean	Std. Dev.	Min.	Percentile 25	Median	Percentile 75	Max
Staff	817	1'396	1'259	-	26	994	1'625	6'980
Enrolments	936	18'700	15'258	-	239	15'262	25'279	161'491
Publications	936	1'959	2'834	-	-	804	2'567	32'254
Citations	887	2'410	4'083	-	-	910	2'962	57'380
Budget (x1000)	797	99'800	257'000	1'247	4'354	24'400	78'900	4'940'000

	Public	Private			US	Europe
PhD Awarding	748	148		Region	366	570

Correlation table	Staff	Enrolments	Publications	Citations	Budget
Staff	1				
Enrolments	0.620	1			
Publications	0.832	0.461	1		
Citations	0.764	0.364	0.979	1	
Budget	0.777	0.411	0.887	0.896	1

The distribution of all variables is highly skewed. The Kolmogorov-Smirnov test shows that the hypothesis that our main dependent variable, i.e. ln_budget, is normally distributed cannot be rejected (p-value = 0.8215), while the normality hypothesis is rejected for all other variables even if logged. Descriptive analysis shows that this is mostly due to a shorter right tail than expected from the lognormal distribution, but most of our data points lie in the central part of the distribution, thereby reducing the risk that the coefficients are strongly influenced by the tails.

As expected, the number of publications and of citations are highly correlated between them and with the budget; budget and staff are also highly correlated (0.777), while the correlation is much lower with enrolments (0.411).

3.3 Scaling properties of scientific output

We analyse the relationship between university revenues and two commonly used scientific output measures: publications and the number of citations in the Web of Science. Since data refer to the whole budget, including educational expenditures, we also control for the volume of education.[1]

As reported in Table 1, a linear relationship is observed on the log-log scale, with slope 1.31 for publications (p-value < 0.001) and 1.47 for citations (p-value < 0.001, robust regression using a mixtures model for errors), corresponding to the degree of the power law distribution for publications and citations over revenues (measured in Purchasing Power Parities). A mixtures model with three Gaussian components provides a significantly better fit than the OLS model (AIC= 1,239 against AIC=1,650 for OLS for publications). The coefficients are very similar in the robust and OLS versions, and they remain significantly above one for both publications and citations. Results are also robust with the inclusion of the number of students as a control and of a dummy for US vs. Europe.

Table 1. Regression results for publications and citations

Models with three mixtures. Heteroskedasticity-corrected standard errors (White 1980).

N = 762 (341 in the US and 421 in Europe). 174 cases dropped due to missing data

	Ln_publications				**Ln_citations**			
	b	se	t	pvalue	b	se	t	pvalue
Ln_budget	1.310	0.073	17.927	0.000	1.473	0.072	20.534	0.000
_cons	-18.550	1.416	-13.098	0.000	-21.727	1.397	-15.548	0.000
Number of obs	762				762			
Pseudo Rsquare	0.681				0.664			
	b	se	t	pvalue	b	se	t	pvalue
Ln_budget	1.612	0.059	27.465	0.000	1.865	0.061	30.546	0.000
Ln_enrolments	-0.176	0.054	-3.235	0.001	-0.312	0.060	-5.168	0.000

[1] For European HEIs, it was not possible to build a robust and comparable measure of research funding only. Additionally, it is questionable whether a clear separation between research and education in universities is possible, therefore, we consider a multivariate framework controlling for the volume of education to be more suitable.

Region=US	-1.218	0.077	-15.925	0.000	-1.395	0.088	-15.914	0.000
_cons	-20.904	0.901	-23.194	0.000	-24.240	0.917	-26.445	0.000
Number of obs	762				762			
Pseudo Rsquare	0.780				0.766			

To ensure robustness, we also perform separate regressions for the two regions. In both cases, we observe super-linear scaling, with coefficients of 1.67 in the US and 1.30 in Europe for publications, and 1.84 in the US and 1.44 in Europe for citations, with both differences being statistically significant. When controlling for enrolments, the difference becomes smaller and not significant (1.67 in the US and 1.47 in Europe for publications, p=0.02, and 1.92 in the US and 1.77 in Europe for citations, p=0.13).

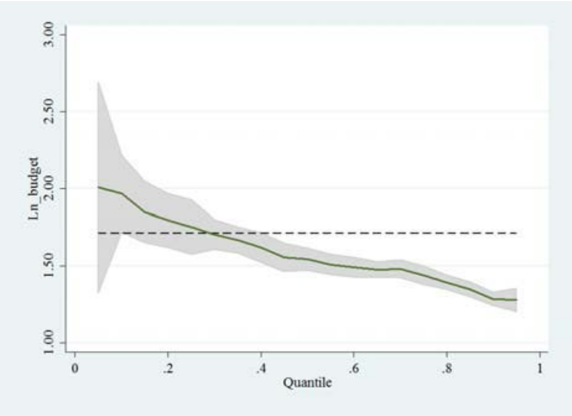
A plot of the standardized residuals against the dependent variable (ln_publications) displays a fairly high correlation, which is concentrated in the left tail of the distribution: the correlation is 0.543** for the whole distribution (N=762), but drops to -.127** when considering HEIs with at least 500 publications per year (N=499). In substantive terms, the analysis of residuals suggests that the model fits fairly well for HEIs with medium and high levels of research output, which is the main focus of our paper, while the relationship is less strong for low-publishing HEIs.

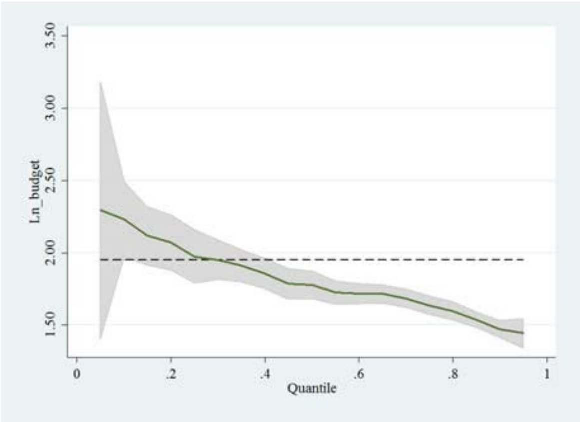
An-in-depth analysis of the cases with large standardized residuals in the regression for ln_publications shows that these cases are explained by particular HEI characteristics. These include graduate schools and large HEIs in terms of enrolments, but with limited research activities as witnessed by low numbers of PhD degrees. Some are on-line universities (Liberty University, more than 70,000 students), others are ex-colleges that award only a few PhD degrees (American International College in the US, Birmingham City University in the UK), others are institutions oriented towards social sciences and humanities, for which the output in the Web of Science is expectedly low (University for Oriental Studies in Naples). We notice that our doctoral universities sample excludes focused universities and therefore this effect is less significant (van Raan 2013).

Finally, a quantile regression displays, first, a decrease in the coefficients with budget size, which remain however well above one for the whole range of quantiles; second, a decreasing standard error for higher levels of the dependent, since less variance is expected for large HEIs because of aggregation effects (Figure 4). On the one hand, this implies that scale effects are stronger especially for the first quartile (as observed also in cities; Bettencourt 2013). On the other hand, the coupling between budget and research output is tighter at the top of the pile, while for smaller universities we observe stronger variation.

Figure 4. Quantile regressions

Coefficient of ln_budget for dependent ln_publications (left) and ln_citations (right). The dashed line corresponds to the OLS estimate, the grey are the coefficient's SE.





All tests performed provide convergent evidence that our results are robust. The regression approaches we use (robust regression and quantile regressions) are more robust than OLS, while the analysis of distributions and outliers provides confidence that results are robust against outliers and that most outliers are in the tail of low-publishing HEIs. This is less of a concern in substantive terms, as our focus is on investigating these relationships for the universities with a sizeable publication output.

3.4 Mediation model for staff

Information on academic staff allows for the disentangling of the two paths through which the budget is associated with scientific production, i.e. by hiring additional staff or by providing more resources (in the form of salaries or research resources) to staff.

As reported in Table 2, the direct coefficient of budget to publications is 1.189, while the indirect coefficient through staff is 0.716*0.696=0.498. The total coefficient is therefore 1.687 for publications and 1.929 for citations, which is consistent with the main regression. Both coefficients are statistically significant, but the former accounts for about two-thirds of the total. Therefore, the main association between the budget on the one side, and publication and citation output on the other side is not through the number of academic staff, but through the amount of resources independent from the number of staff.

As expected, student enrolments have a positive association with the number of staff, implying that with increasing enrolments the budget is used to a larger extent to hire staff. For what concerns publications, the indirect coefficient through an increase in the number of staff is positive (0.191*0.716=0.136), but the aggregate coefficient is negative (0.136-0.324=-0.188), i.e. universities with more students have less publications and citations with the same budget.

Table 2. Mediation models for citations and publications

OLS with robust standard errors.

	Coef	Std. Err.	P>|z|		Coef	Std. Err.	P>|z|
Dependent variable	ln_publications				ln_citations		
ln_staff (γ)	0.716	0.199	0.000		0.769	0.245	0.002
ln_budget (α)	1.189	0.164	0.000		1.393	0.203	0.000
ln_enrolments	-0.324	0.076	0.000		-0.467	0.090	0.000
region= US	-0.981	0.140	0.000		-1.122	0.169	0.000
cons	-17.618	2.038	0.000		-20.544	2.551	0.000
Dependent variable	ln_staff				ln_staff		
ln_budget (β)	0.696	0.024	0.000		0.696	0.024	0.000
ln_enrolments	0.191	0.026	0.000		0.191	0.026	0.000
region=US	-0.471	0.025	0.000		-0.471	0.025	0.000
cons	-8.076	0.342	0.000		-8.076	0.342	0.000
Indirect coefficient (β*γ)	0.498	0.132	0.000		0.536	0.163	0.001
direct coefficient (α)	1.189	0.164	0.000		1.393	0.203	0.000
total coefficient (α+β*γ)	1.687	0.058	0.000		1.929	0.069	0.000
N	700				700		
AIC	2253.404				2453.899		

4 Scaling and differences in resource distribution

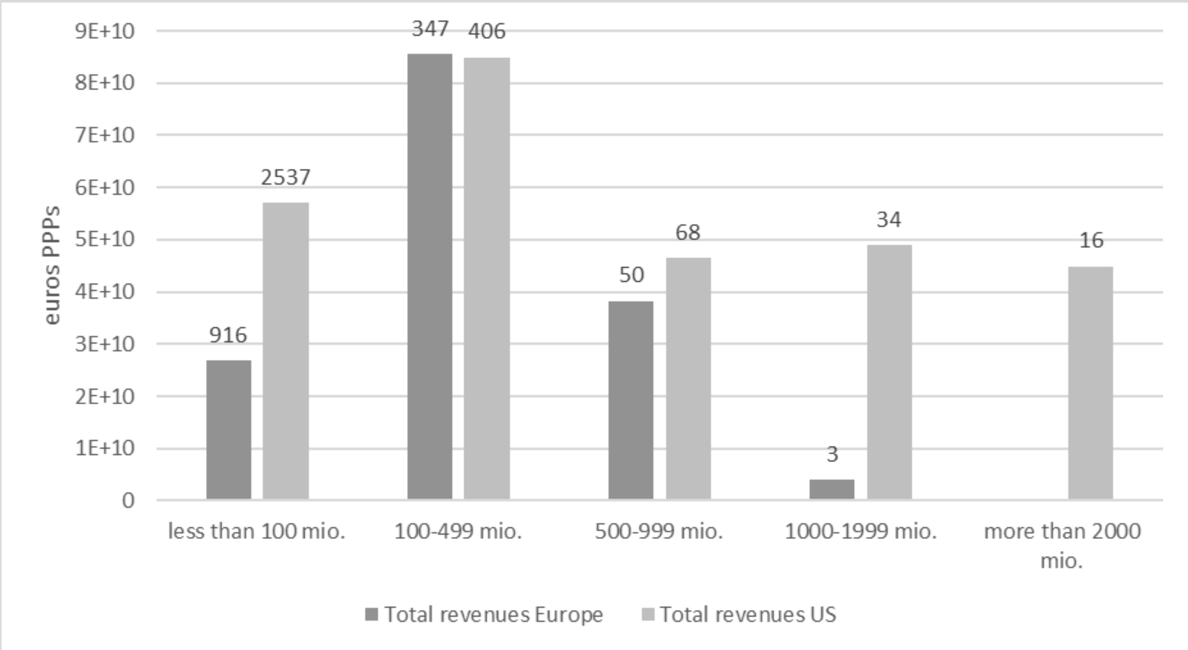
While scaling relationships are similar across systems, the two systems are characterized by major differences in the distribution and composition of HEI revenues. The interaction between stronger concentration of revenues in the US and super-linear scaling results in a much higher number of publications and citations for the top-US universities than the top-European ones (both absolute and normalized by size).

Figure 1 shows that the US system has a larger number of small HEIs (in terms of budget) and a group of HEIs with extremely large budgets, while in Europe the largest portion of resources are directed to middle-size HEIs. On the top of the pile, the US system includes 16 HEIs with total revenues above 2 billion euros in PPPs, which account for 16% of revenues, while the 50 HEIs with a budget above 1 billion constitutes one-third of all system-level resources. On the contrary, in Europe there are only 3 HEIs with a budget above 1 billion euros, while half of the revenues are accounted for by middle-sized HEIs below 500 million Euros.

All top-25 HEIs by budget in the database are in the US, with the list being topped by Harvard, Stanford and Yale, the first European universities are Cambridge (place 26) and Oxford (place 41), i.e. the highest ranked European HEIs in the Shanghai ranking.

Figure 1. HEI budgetary classes

Number of HEIs by region and class. Left axis: sum of revenue by class. Right axis: Median by category of resources per student. Amount in euros PPPs.

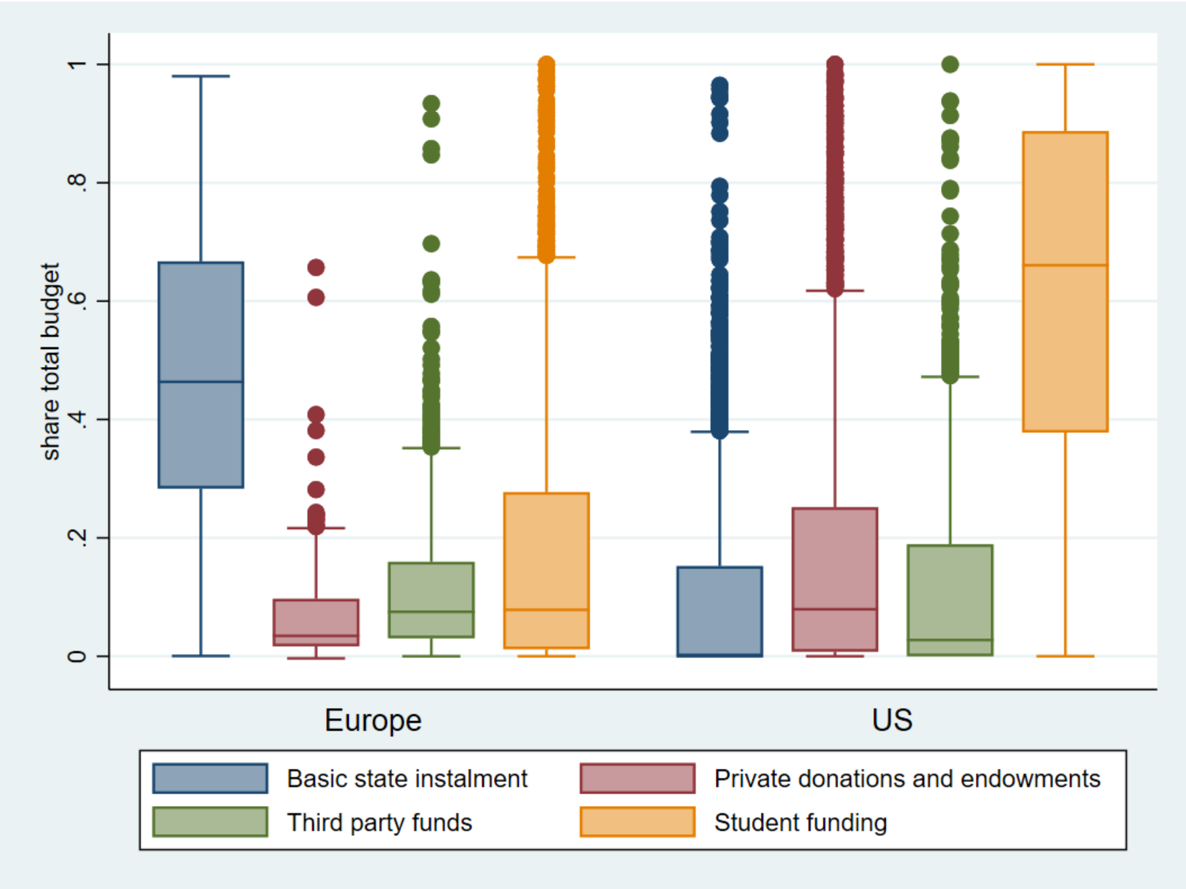


When combined with super-linear scaling of publications and citations over the budget, the distribution of revenues translates into a dominance of the US universities in the ranking by number of publications and citations, respective to the international rankings, which are closely correlated to bibliometric indicators. As a matter of fact, 15 out of the top-25 universities in the Shanghai ranking are among the top-25 HEIs in our database for budgets, and Harvard and Stanford top both lists.

Second, the US system is characterized by a stronger differentiation of revenue sources in the aggregate and between HEIs. To this aim, we divide university budgets into four categories of funds, i.e. the basic state instalment to HEIs, private donations and pay-outs from the endowment at the university level, third-party funds (mostly research contracts) and funding from student fees. As demonstrated in Figure 2, most European HEIs have a funding model where the basic state instalment represents the largest share of funds, while other sources are complementary – the only exceptions are private for-profit HEIs and public UK universities that are mostly funded through student fees.

On the contrary, US universities have a differentiated funding model, where private resources and student funding play a central role – the latter being largely indirect state support through student loans and subsidies. Differences within the system are large. The public (state) universities have a composite funding structure, in which state funds represent a sizeable (even if diminishing) share of the budget (Weerts and Ronca 2012), while private for-profit HEIs are mostly funded by student fees, similar to Europe. Finally, the large private non-profit sector, that comprises most of the top-universities in terms of research output, are funded by a combination of private donations and endowments, and through student fees.

Figure 2. Boxplots of the share of different types of funds over total revenues



Third, we observe differences in how the largest universities in terms of budget are funded (Table 3). In the US, tuition fees are the main funding source for smaller and more education-oriented HEIs, while private donations and endowments are concentrated at the top of the pile. The 16 universities with a budget above 2 billion euros account for 53% of private donations that constitute 49% of their resources. On the contrary, in Europe, the universities with the highest revenues are funded by a combination of state allocation and third-party funds, but the extent of concentration is far lower.

Table 3. Composition of HEI revenues

Median by budgetary class, percentages over total revenues. Given the different availability of data and type of statistics, subcategories do not necessarily add to 100%.

Budgetary class		less than 100 mio.	100-499 mio.	500-999 mio.	1000-1999 mio.	more than 2000 mio.
Europe	Basic state instalment	52 %	35 %	38 %	26 %	
	Private donations and endowments	4 %	3 %	3 %	1 %	
	tuition fees	6 %	11 %	9 %	21 %	
	third party funds	6 %	10 %	21 %	40 %	
US	Basic state instalment	0 %	21 %	18 %	13 %	0 %
	Private donations and endowments	7 %	8 %	14 %	22 %	43 %
	tuition fees	72 %	41 %	30 %	23 %	16 %
	third party funds	2 %	17 %	28 %	29 %	23%

4.1 Decomposing differences in resource distribution

We finally explore some institutional factors that may explain the observed differences in resourcing between the US and Europe.

First, the US higher education system is endowed with more resources on the aggregate. The total amount of resources received by the HEIs in our database is 282 billion euros PPS in the US and 154 billion euros in Europe (Table 4). Data may be affected by missing data, especially in Europe: however, when comparing only HEIs that have financial data, the numbers of staff, enrolments and publications are similar in both systems. Consistent with previous literature, US university citations are significantly higher than European institution citations (Bonaccorsi, Cicero, Haddawy and Hassan 2017). However, US HEIs have about two times the resources of European HEIs.

Table 4. Aggregated data for US and Europe

The second subtable reports aggregates for other variables only for the HEIs for which we have financial data, due to large numbers of missings in Europe.

		N. of HEIs	Total revenues (million euros PPPs)	Academic staff FTE	Student enrolment ISCED 5-7	Number of publications	Number of citations
All HEIS	Europe	2,264	154,722	975,270	17,075,833	1,090,597	6,467,962
	US	3,287	282,401	842,730	13,669,196	866,077	6,803,918
	Europe	1,316	154,722	819,882	13,079,969	907,417	5,557,805

Only HEIs with financial data	US	3,061	282,401	832,116	13,448,559	865,078	6,799,080

The difference in resources from our data is compatible with international statistics where tertiary education spending was 2.7% of GDP in 2014 in the US and ranged between 1% and 2% in European countries (source: OECD, Education at a Glance). This difference is essentially due to resources from the private sector and from students (also including state subsidies to students) that accounted for two-thirds of tertiary education spending in the US, but to less than 40% in most European countries (with the exception of the UK).

Second, we observe a difference in the extent of institutional differentiation between the two systems, as revealed by applying the US Carnegie classification criteria to the dataset (Table 5). Although the European system comprises a large number of colleges and specialized HEIs, doctoral universities account for nearly 70% of academic staff and enrolments at the bachelor and master level, when compared to only 55% of staff and 45% of the enrolments for US doctoral universities. The difference would have been even larger when considering all tertiary education institutions, since HEIs delivering short degrees (associate colleges) are far more important in the US than in Europe.

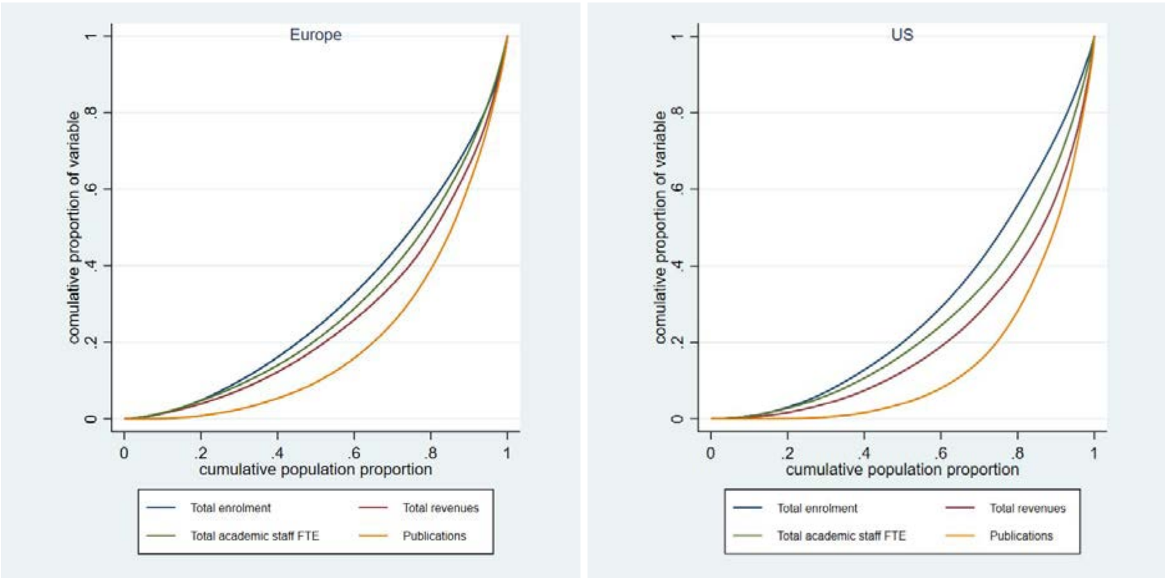
Table 5. Carnegie classification applied to the dataset

	Europe			US		
	N. HEIs	Staff	Enrolments	N. HEIs	Staff	Enrolments
Unclassified	88	190	70,210	231	11,991	307,332
Doctoral Universities	570	671,505	11,212,100	366	469,233	6,291,367
Masters' colleges and universities	549	184,660	3,785,405	815	212,263	4,550,288
Baccalaureate colleges	348	39,255	789,930	637	61,304	959,374
Baccalaureate/associate colleges	20	2,999	58,664	487	33,700	1,049,342
Focused institutions	689	76,661	1,196,661	751	54,239	511,493

Since colleges receive fewer resources per student in both systems, a higher share of students attending colleges translates into more resources for doctoral universities in the US. This difference has lasting historical roots: the US system was grown from different institutional models, including the appearance of the research university as a distinctive type of institution during the 20th century (Geiger 1993; Cohen 2007). Europe was historically dominated by the "Humboldtian" public university model, with attempts to differentiate a second sector of higher education only beginning in the 1970s (Daraio, Bonaccorsi, Geuna, et al 2011).

Third, when focusing on the subpopulation of 936 doctoral universities, we highlight differences in how the variables are distributed (Figure 3). The level of concentration is similar in both systems for enrolments (the Gini coefficient is 0.419 in the US against 0.387 in Europe) and academic staff (0.491 against 0.421). However, in the US, revenues (0.572 against 0.472) and publications (0.768 against 0.571) are more concentrated than in Europe when compared to enrolments. In other words, European HEIs tend to "scale up" with student enrolment, with the distribution of revenues closely following the distribution of students and with research outputs only moderately more concentrated. On the contrary, revenues are more concentrated than students (and staff) in the US when also taking into account super-linear scaling, publications are also far more concentrated.

Figure 3. Lorenz curves of the distribution of variables for doctoral universities. N=936



5 Discussion and conclusions

Our findings move beyond previous results on preferential attachment in scientific networks (Peterson, Presse and Dill 2010), as we demonstrate that the number of publications and citations scale super-linearly over resources. Therefore, bibliometric measures, including those normalized by the volume of scientific production, are all size-dependent and need to be complemented with a measure of resources in order to compare HEIs in a sensible way and to draw policy conclusions (Abramo and D'Angelo 2016).

In a comparative perspective, our results show that the US system includes a group of universities with a far higher level of resources than their best-funded European counterparts. We also showed that this pattern can be associated with three differences in resourcing between US and Europe: a higher level of resources, possibly enabled by the differentiation of funding streams; a more diverse set of HEI types, which dates back to the 19th and 20th century; a higher concentration of resources (absolute and relative to students) also among doctoral universities. When combined with super-linear scaling of publications and citations over resources, these results imply that the US system includes a group of universities that have a far higher level of publication output and citations than their European counterparts and, therefore, appear at the top of international rankings.

Our analysis shows two further characteristics. First, the strongest association between budget and research output is via additional resources per staff, rather than an increase in the number of faculty. Therefore, universities with more resources also have higher resources per unit of staff. This suggests that a key underlying mechanism explaining the observed patterns are providing attractive conditions to recruit talented academic staff, for example in the form of starting packages for newly hired professors (Stephan, Paula 2013).

Second, in both systems, the wealth of large doctoral universities is not built on the acquisition of project funds or on tuition fees – particularly in the US, where tuition fees are the main funding source for smaller and more education-oriented HEIs. On the contrary, universities with the highest level of resources are largely funded by core institutional funds, either from the state (in Europe) or from private donors (in the US).

Since the data are cross-sectional, our results cannot be interpreted in terms of causality, for example that a greater concentration of funding leads to a more than proportional increase in the number of scientific publications and of citations. Even more so, since the amount of resources acquired by universities is largely endogenous and influenced by the previous years' output (Stephan, Paula E. 1996). Rather, they should be interpreted as a demonstration of a structure in the university system where the amount of resources and publications (respectively citations) are tightly coupled through a super-linear relationship. While this is by no means new, our work, first, shows how strong and pervasive is this relationship across different university sizes and regions and, second, demonstrates super-linear scaling. The latter implies that measures of 'excellence' in the international rankings prime concentration of resources irrespectively of 'quality' (Paradeise and Thoenig 2013).

Possible generating mechanisms for this relationship proposed by the literature include size and scope economies in the production of scientific output (Koshal and Koshal 1999), agglomeration effects as demonstrated for cities (Bettencourt 2013) and attention-based preferential attachment (Newman, M. E. J. 2004). A further mechanism that is compatible with our data is sorting of human resources, where highly productive scientists have a preference for the 'best' places in terms of research quality, while in their hiring behavior universities attempt at maximizing 'excellence' by investing more resources in few highly productive people (Stephan, Paula 2013).

The latter remark emphasizes that a key component of these processes is the existence of a universal (context-free) and measurable definition of 'excellence' that might different from (context-related) quality (Paradeise and Thoenig 2013). Such a measure, like the one conveyed by international rankings, is not necessarily 'objective', but nevertheless drives the behavior of the actors in the science system, including policy-makers, university managers and scientists themselves, through subtle and largely invisible sociological mechanisms (Sauder and Espeland 2009). Furthermore, our analysis of financial data shows how these measures are coined to the position of a small set of highly-funded US universities and, therefore, by reproducing the same social norms throughout the whole higher education system and across countries, contribute to maintain their long-term hegemony (Hazelkorn, E. 2009; Deem, Mok and Lucas 2008b).

Our results also demonstrate empirically that fighting for the top-positions in international rankings must be associated with the concentration of large amounts of resources in a few places. The analysis of the funding system suggests that this is associated with the long-term construction of specific institutional structures (Bonaccorsi 2007). In Europe, this was achieved only by two countries, i.e. UK with its longstanding tradition of concentrating resources in a few top-universities, and in Switzerland through the creation of two 'national' universities in a federal system.

We finally suggest some implication of our results for policy-making and for university management. First, the strength and pervasiveness of these effects implies that it hardly makes sense for university managers and policymakers to oppose to them and to attempt to recreate alternative *global* norms of quality; in a subtle way, resistance just strengthens the dominance of the 'excellence' measure defined by international rankings (Sauder and Espeland 2009).

Second, maybe surprisingly, our results suggest that the battle for international rankings should not be the main concern of policymakers and of university managers for two good reasons: first this process is driven by largely endogenous mechanisms and, at the least in the short and medium term, there is important inertia that makes it difficult to substantially change the amount and distribution of resources; second, even in a well-funded system like the US one, this concerns only a handful of universities that account for a tiny proportion of higher education activities, particularly for what concerns education and the contribution to society and economy. Policy-makers and university managers should rather concentrate on enhancing the specific contributions individual HEIs might provide to their regional and national contexts (Paradeise and Thoenig 2013).

A look to the US higher education policy is, in this respect, instructive. The battle for world-class universities used to be mostly a concern of a few university presidents and for rich private donors. In turn, federal and state policies were mostly concerned with the provision of access to higher education through a highly diversified set of HEIs, while most HEIs in the system are fighting for 'regular' students and for their local rooting (reference). Our results suggest that imitating the 'top private' part of the system might bear more dangers than benefits both in the US and in Europe (Borden, Coates and Bringle 2018; Hazelkorn, Ellen and Gibson 2018).

6 Acknowledgements

The authors acknowledge support from the European Union through the project, Research infrastructures for the assessment of science, technology and innovation policy (RISIS) (Grant agreement no: 313082). Thanks are due to the European Commission, ETER project, and to the CWTS, University of Leiden, for providing access to the data. Support from Dr. V. Veglio in the preparation of the dataset and support from Prof. F. Bartolucci in extending the code for robust regression are acknowledged. Comments from Prof. Paula Stephan from Georgia State University are also acknowledged, as well as from an anonymous referee.

7 References

Abramo, G. & D'Angelo, C. A. (2016). A farewell to the MNCS and like size-independent indicators. *Journal of Informetrics, 10(2)*, 646-651.

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A. & Sapir, A. (2010). The governance and performance of universities: evidence from Europe and the US. *Economic Policy, 25(61)*, 7-59.

Barabási, A. & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286(5439)*, 509-512.

Bartolucci, F. & Scaccia, L. (2005). The use of mixtures for dealing with non-normal regression errors. *Computational Statistics & Data Analysis, 48(4)*, 821-834.

Bettencourt, L. M. (2013). The origins of scaling in cities. *Science (New York, N.Y.), 340(6139)*, 1438-1441.

Bonaccorsi, A. (2007). Explaining poor performance of European science: institutions versus policies. *Science and Public Policy, 34(5)*, 303-316.

Bonaccorsi, A., Cicero, T., Haddawy, P. & Hassan, S. (2017). Explaining the transatlantic gap in research excellence. *Scientometrics, 110(1)*, 217-241.

Borden, V., Coates, H. & Bringle, R. (2018). 15 Emerging perspectives on measuring and classifying institutional performance. *Research Handbook on Quality, Performance and Accountability in Higher Education,* , 189.

Cohen, A. M. (2007). *The shaping of American higher education: Emergence and growth of the contemporary system* John Wiley & Sons.

Daraio, C., Bonaccorsi, A., Geuna, A., Lepori, B., Bach, L., Bogetoft, P., Cardoso, M. F., Castro-Martinez, E., Crespi, G., Lucio, I. F. d., Fried, H., Garcia-Aracil, A., Inzelt, A., Jongbloed, B., Kempkes, G., Llerena, P., Matt, M., Olivares, M., Pohl, C., Raty, T., Rosa, M. J., Sarrico, C. S., Simar, L., Slipersaeter, S., Teixeira, P. N. & Van den Eeckaut, P. (2011). The European university landscape. *Research Policy, 40(1)*, 148-164.

Deem, R., Mok, K. H. & Lucas, L. (2008a). Transforming higher education in whose image? Exploring the concept of the ‘world-class’ university in Europe and Asia. *Higher Education Policy, 21(1)*, 83-97.

Deem, R., Mok, K. H. & Lucas, L. (2008b). Transforming higher education in whose image? Exploring the concept of the ‘world-class’ university in Europe and Asia. *Higher Education Policy, 21(1)*, 83-97.

Geiger, R. L. (1993). *Research and relevant knowledge: American research universities since World War II* Oxford: Oxford University Press.

Hazelkorn, E. (2009). Rankings and the battle for world-class excellence: institutional strategies and policy choices. *Higher Education Management and Policy, 21/1*.

Hazelkorn, E. & Gibson, A. (2018). 18 The impact and influence of rankings on the quality, performance and accountability agenda. *Research Handbook on Quality, Performance and Accountability in Higher Education,* , 232.

Katz, J. S. (1999). The self-similar science system. *Research Policy, 28(5)*, 501-517.

King, D. A. (2004). The scientific impact of nations. *Nature, 430(6997)*, 311-316.

Koenker, R. & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives, 15(4)*, 143-156.

Koshal, R. K. & Koshal, M. (1999). Economies of scale and scope in higher education: a case of comprehensive universities. *Economics of Education Review, 18(2)*, 269-277.

Leitao, J. C., Miotto, J. M., Gerlach, M. & Altmann, E. G. (2016). Is this scaling nonlinear? *arXiv Preprint arXiv:1604.02872,* .

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 5200-5205.

Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46(5)*, 323-351.

Nomaler, Ö, Frenken, K. & Heimeriks, G. (2014). On scaling of scientific knowledge production in US metropolitan areas. *PloS One, 9(10)*, e110805.

Paradeise, C. & Thoenig, J. (2013). Academic Institutions in Search of Quality: Local Orders and Global Standards. *Organization Studies, 34(2)*, 189-218.

Partha, D. & David, P. A. (1994). Toward a new economics of science. *Research Policy, 23(5)*, 487-521.

Peterson, G. J., Presse, S. & Dill, K. A. (2010). Nonuniversal power law scaling in the probability distribution of scientific citations. *Proceedings of the National Academy of Sciences of the United States of America, 107(37)*, 16023-16027.

Sauder, M. & Espeland, W. N. (2009). The discipline of rankings: tight coupling and organizational change. *American Sociological Review, 74(1)*, 63-82.

Stephan, P. (2013). The endless frontier: Reaping what bush sowed? In A. B. Jaffe & B. F. Jones(Eds.) *The changing frontier. Rethinking Science and Innovation Policy* (pp. 321-370). Chicago: Chicago University Press.

Stephan, P. E. (1996). The Economics of Science. *Journal of Economic Literature, 34(3)*, 1199-1235.

van Raan, A. F. (2013). Universities scale like cities. *PloS One, 8(3)*, e59384.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E., Tijssen, R. J., Eck, N. J., Leeuwen, T. N., Raan, A. F., Visser, M. S. & Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology, 63(12)*, 2419-2432.

Waltman, L. & van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics, 9(4)*, 872-894.

Weerts, D. J. & Ronca, J. M. (2012). Understanding differences in state support for higher education across states, sectors, and institutions: A longitudinal study. *The Journal of Higher Education, 83(2)*, 155-185.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society, ,* 817-838.