

The Transformative Power of Evaluation on University Governance

Emanuela Reale and Giulio Marini

E. Reale (*)

Research Institute on Sustainable Economic

Growth IRCRES, National Research Council CNR, Rome, Italy e-mail: emanuela.reale@ircres.cnr.it

G. Marini

Centre for Global Higher Education, Institute of Education, UCL, London, UK

e-mail: giulio.marini@uniroma1.it

Abstract This chapter investigates how external evaluation affects university governance. The two research questions are: What makes evaluation a powerful instrument affecting the governance of universities? Do different evaluation instruments have different strengths in affecting governance? We assume that evaluation has several areas of potential effect on universities related to transformation of: hierarchical relationships between actors; the academic profession; management and performance. The chapter surveys relevant literature, followed by an outline of the conceptual framework, a presentation of the data used and the tests developed, and data analysis. We attempt to identify homogenous clusters of higher education institutions (HEIs) in terms of the extent to which evaluation influences their governance, before our concluding discussion of the analytical results and the conclusions.

Introduction

The aim of this chapter is to investigate how external evaluation contributes to the transformation of universities through effects on their internal governance. The inquiry is driven by two research questions: What makes evaluation a powerful instrument affecting the governance of universities? Do different evaluation instruments (evaluation of research, evaluation of teaching, quality assurance) have different strengths in affecting governance?

In the current study, we consider evaluation as having several areas of potential effect in the transformation of universities: (i) changes in the hierarchical relationships between actors, with possibilities of verticalization in the distribution of power, and reinforcement of the central power through the definition of strategies, and the use resource allocation and performance assessments; (ii) contributions to the reconfiguration of the academic profession and the formation of new elites; (iii) improvement in management and performance, through the rationalization of the use of available resources and the pursuit of efficiency and effectiveness.

The study investigates the first of the above three effects of transformation, arguing that evaluation does not bring about a concentration of power in the hands of the university's central governing bodies. Rather it is expected that an observable division of competences and influences between the actors involved, meaning between the top and middle managers, is linked to the type of evaluation used. Furthermore, we argue that research evaluation is likely to produce a deeper impact on university governance than other forms of evaluation, particularly the assessment of teaching.

The chapter is organized as follows. The first section presents the relevant literature on the effects of evaluation on university governance. The second section describes the conceptual framework, the third section data used and the tests developed. The fourth section provides the data analysis. The fifth section attempts to identify homogenous clusters of higher education institutions (HEIs) in terms of the extent to which evaluation influences their governance. The last section provides a discussion of the analytical results and the conclusions.

Relevant Literature

Governance is the ways in which decision-makers 'combine to solve collective problems' (Capano 2011; Kooiman 2003), and the steering of this process; thus decisions are formulated and implemented as a result of the interactions of the various actors involved in the organization (Capano 2011, 1625).

The study deals with the internal governance of universities, and in particular with the area of how universities are currently changing the organization and distribution of internal power. These changes are the result of a combination of three influences: (i) the governments pushing forward reforms, redesigning systemic governance; (ii) the increasing numbers of internal and external actors involved in the steering, and (iii) the use of new internal systematizations and instruments designed to govern university organizational and academic behaviors. All three of these influences should be considered for a full understanding of the transformation of internal university governance. In this study the intention is not to provide a general description of how the introduction of evaluation as a steering tool has impacted universities. Our particular focus is instead on how the

actors within the universities perceive the advent of evaluation and of the 'Evaluative State' (Neave 1998, 2012). From this, we also arrive at what kinds of effects this generates in terms of the distribution of decision-making power among the different levels of university government, specifically the central government (top) and middle levels.

The literature offers several elements contributing to this work. First, we refer to the extensive investigation of the unique nature of universities as organizations. Apart from certain differences in interpretation, there is substantial consensus that universities, at least in part, still remain as loosely coupled entities, 'incomplete organizations' or 'organized anarchies' (Weick 1976; Musselin 2007; Brunsson and Sahlin-Andersson 2000; Whitley 2012).

Second, we recognize that this distinctive character does not imply that the transformation of universities into tighter organizations has failed. Although the state reforms have achieved different levels of effect, and the processes are implemented through different steps and rates, all the European countries show a trend towards strengthening universities as organizations (Bleiklie and Kogan 2007; Paradeise et al. 2009). Recently, Seeber and colleagues (2014), using the same TRUE dataset as the current work, investigated the form of European universities in the dimensions of identity, hierarchy and rationality, which are the basic characteristics of complete organizations (Brunsson and Sahlin-Andersson 2000; Birnbaum 2004). The results show that although 'managerial' universities can be distinguished from traditional ones, these still cannot be considered 'completely complete' organizations, and probably cannot ever become such.

Third, evaluation is one of several instruments intended to enhance the command and control capabilities of government, meaning both the capabilities of the state government concerning universities and those of the internal governments themselves (Stame 2006). Evaluation is one component of a new steering-at-a-distance approach, featuring the redistribution of powers within the universities, the emergence of new actors influencing the universities' choices, and the increasing capability of the institutions themselves to plan, control and measure achievements. Nevertheless, some authors indicate that evaluation can have greater effects in terms of legitimization and establishing prestige among actors, than the desired ones as a tool for coordination and steering. Such counter-effects occur because of the complexities in assessing the academic profession, and the limited possibilities of influencing the research agendas of individuals (Whitley 2007).

Fourth, we observe that originally, evaluation developed as an independent steering instrument, mostly in the form of quality assessment (QA) and as *ex-post* research assessment. However in the European university context, the origin and drivers of change were different. Here, the Europe-

wide Bologna Process played a leading role in the growth of quality assessment, while research evaluation exercises have descended primarily from national political initiatives, with implementation by external agencies. The development of the evaluation and assessment instruments is often inspired by New Public Management principles, which aim at simultaneously incrementing the steering capabilities of the policy-maker and the autonomous responsibilities of the universities (Reale and Seeber 2013). Given this context, the implementation of the evaluation instrument is related to levels of autonomy granted to the institutions, and to the types of competitive funding systems in the different countries (Whitley 2007).

Furthermore, external evaluation impacts on governance through changing the behavior of the academics. In particular, empirical evidence has shown that middle managers tend to criticize the rules of competitive funding and evaluation, yet at the same time be attentive to them, 'reacting according to the requirements of the process' (Musselin 2013).

Fifth, Hansson (2006) outlines the passage from research evaluation as an 'anonymous and autonomous process of qualitative control based on peer review to a complex process involving several methods and tools', responding to a number of objectives and goals. The methods and approaches that have been introduced to research evaluation attempt 'to accommodate at least two, often contradictory, policy goals: the demonstration of accountability and productivity of the researchers and the research organization' (Hansson 2006). The changes in research evaluation have mainly been intended to achieve performance management in universities, without considering the social context or organization of the scientific work, simply pushing the organizations to exert controls on the individual researchers through 'constant monitoring of productivity' (Hansson 2006, 167).

A final stream of important literature concerns the organizational responses to internal and external evaluation, which have been explored from different institutional perspectives (Paradeise et al. 2009). However, the current study is not intended to examine how universities react to the reforms introducing evaluation as a steering instrument, or to internal reorganizations intended to implement evaluation as a means of governance. Our interest is to investigate the transformative effect of evaluation through the redistribution or concentration of power within the universities, at the different levels of internal governance, namely at the top and middle levels of management.

The distinction between external and internal evaluation is significant in our investigation. The former is a process of review carried out by a body not directly involved in the organization being evaluated, generally a government department or evaluation agency. This type of evaluation is generally considered as having advantages of objectivity, avoidance of influence from vested internal interests, and outside perspectives on organizational issues. Internal evaluation is a process of review carried

out by someone from the organization itself, for the organization's own ends, with or without the involvement of external peers. The advantage of this type of evaluation is that it is fully focused on the organization's internal priorities and purposes. The process of internal evaluation can serve as a step in preparation for an external evaluation, or can complement an external assessment in some other way. In spite of the distinction, both types of evaluation activities bring pressures to bear on the organization. External evaluation is directed at the different levels of the organization's internal government, and sets standards and performance objectives to be addressed. Internal evaluation anticipates, supports and prepares for the pressure deriving from external evaluation, thus concurring in producing the intended effects on the organization. Although the pressures exerted by the two types of evaluation can be analyzed separately, it is difficult to distinguish their relative contributions to the effects produced on the organization.

Finally, it is useful to point out that evaluation can be used for two different but quite interconnected aims: *summative* and *formative* (Scriven 1967; Taras 2005). Evaluation for summative aims is intended judge to what extent a specific goal or performance has been achieved. Evaluation for formative aims is devoted to learning lessons from the past, for future improvement. The literature generally describes summative and formative evaluation in terms of a dichotomy (Scriven 1967), but some authors point out the linkages between the two, since formative evaluation is a necessary step in preparing a good summative assessment (Taras 2005; Molas-Gallart 2012). The current study deals with evaluation primarily from the summative perspective. The different types of evaluation are seen as steering tools, linked to the emerging neo-liberal approach, although some formative purposes can also be considered to occur.

CONCEPTUAL FRAMEWORK AND EMPIRICAL CONTEXT

Proceeding from the literature, we require a conceptual framework to explain the way evaluation, and research evaluation in particular, are likely to affect the governance of universities.

One of our considerations is that the transformative power of evaluation is related to its specific uses. For this, we draw on the classification proposed by Molas-Gallart (2012), categorizing the uses of evaluation in the policy process. The starting point is the identification of the main activities linked to the policy process: resource allocation, activities for the pursuit of institutional tasks, and control over the correct use of resources allocated for the activities. Examining the linkages and potential contributions of evaluation to the policy process, the author suggests three main potential purposes or uses:

- *Distributive* uses, mainly referring to the allocation of the available resources to different beneficiaries, on the basis of the performance assessment or the merit of the individuals and groups;
- *Improvement* uses, where the focus is on the lessons that can be learned from past experience, using evaluation to disentangle the reasons and explanations for certain effects;
- *Controlling* uses, to scrutinize how individuals, organizations or groups have used the resources for carrying out the planned activities, and the extent to which they have achieved the expected results.

Distributive and controlling uses are primarily related to *summative* evaluation, while improvement purposes imply a *formative* approach to evaluation.

We assume that universities could apply external evaluation for all three uses in their internal decision-making processes, or could limit themselves to the two areas of summative purposes, leaving out the formative one. Given the definitions offered by Molas-Gallart, the current study is particularly interested in the use of external evaluation in a summative approach, for distributive and controlling purposes. The uses of external evaluation can thus affect the relevance of the internal actors in governance and management (rector, board, senate, central or ‘top’ administrators, middle managers) differently, depending on the type of evaluation, affecting the power they own, and create various impacts with either positive or negative values.

The relevant dimensions for the investigation are:

How the central administrative bodies perceive the influence of the external evaluation, carried out by the government or a specialized evaluation agency, on the specific uses of the evaluation. The distinction is *distributive uses* (effects on financial allocations, thus determining the budgets of institutions and academic units) and *controlling uses* (effects on persons, such as on the careers of academics, content of their teaching programs or research priorities).

How central government bodies perceive the impact of the different types of evaluation (evaluation of research, of teaching; quality assessment) on the decisions made within the university. Indeed, the literature suggests that that evaluation can have several purposes

and impacts in the case of universities (Dahler-Larsen 2012; Reale and Seeber 2013). The intention of is thus to deepen the extent to which the actors consider evaluation to be an instrument driving university decisions.

How the university's central government bodies on the one hand, and middle managers on the other, perceive the impact of evaluation on universities, distinguishing between *negative effects* (conflicts, bureaucracy, individual opportunistic behaviors) and *positive effects* (improving quality in teaching and research, transparency).

How the university's central government bodies on the one hand, and middle managers on the other, perceive the distribution of the decision-making power between the central government and middle management levels. The issues analyzed are those related to the distributive use of evaluation (setting the rules and procedures for the evaluation of faculties, institutes and other internal units) and the controlling use of evaluation (assessing the individual academic performance).

Given the assumptions and the dimensions for investigation, we elaborate three propositions. We can expect that:

1. The more that external evaluation is applied for distributive purposes about vital resources (e.g. funding; assignment of positions through recruitment or career advancement) the more the academics perceive negative effects;
2. The more that the university concentrates on the decision-making power at the internal government level (versus with middle managers), the stronger is the use of external evaluation for controlling purposes;
3. The more that research evaluation influences the activities carried out by the university or empowers the steering capacities of the central government bodies, the more it is likely to produce significant effects on governance, compared to other types of evaluation.

The three propositions will be also tested against other dimensions related to the characteristics of the universities involved in the TRUE survey and to the characteristics of the national higher education systems, searching for different effects might be linked to institutional and/ or national factors.

We expect to confirm the first proposition. Our reasoning is that the use of external evaluation for the allocation of resources is likely to produce perceptions of negative effects in all the academics involved in university governance, both at the central government and middle levels. This is in spite of the fact that different organizational responses to university evaluation have been observed at the level of the research groups (Reale and Seeber 2010).

However, we expect differentiated outcomes to the second proposition, depending on the university characteristics and ultimately on country factors. When the central governments of universities use evaluation for individual assessments or to achieve effects on research priorities or the content of teaching, the distribution of the decision-making power on issues related to the evaluation 'control tool' can be perceived differently by the central government and the middle managers, with the former adopting a steering-at-a-distance behavior and the latter adopting attitudes closer to 'shop-floor' preoccupations.

We expect to confirm the third proposition for all universities, concerning the effects of research evaluation on governance, regardless of institutional and national characteristics. Research determines the individual prestige and reputation lying at the very core of the academic world. Thus, when the central government perceives research evaluation as an important steering instrument with a strong influence on the university activities, the enactment of evaluation is also likely to influence the centralization of the decision-making power over the use of evaluation for control purposes.

The analysis acknowledges several limitations. The TRUE survey is a cross-sectional study, in which causal inference is highly problematic except when external measures are used. Thus we cannot establish strict causal relationships between the use of evaluation at the institutional level and the effects that it produces on the distribution of decision-making power, or on other positive or negative results. Moreover, the literature on both research evaluation and quality assurance indicates that a variety of factors, acting at both the institutional and national levels, can influence the results descending from evaluation, (Dahler-Larsen 2012:

20; Hammarfelt and de Rijcke 2015; Huisman et al. 2007). The direction of influence between the phenomena under investigation can also be questioned: does evaluation influence the concentration of decision-making power, or vice versa? The intention of the current study is to use the data obtained from the TRUE survey to test for the existence

of statistical associations (Pearson correlations) between the degrees of impact of several types of evaluation and the holding of true decision-making power in evaluation-related issues (whether by central or middle management). Such correlations could confirm the likelihood of the ongoing partition of decision-making power inside universities, although not informing about causalities. Further, the correlations can serve in assessing the use of evaluation for summative aims, indicating the degree of influence of evaluation on the governance of universities. In short, the aim of our investigation is not to identify dependencies between the different factors, but rather discuss how evaluation enters into the configuration of the universities and how the use of external evaluation for summative aims is likely to transform them.

EmPIriCAL PArAmeters

The study uses the dataset developed under the ‘ESF EUROCORE-TRUE Project - Transforming Universities in Europe’. The data are from a survey directed at different organizational levels within a sample of 26 universities in eight European countries (Germany, France, Italy, Netherlands, Norway, Portugal, Switzerland, United Kingdom; 696 answers, with a response rate of 48% overall, with no particular differences in terms of response rate by role, university and country). The sampling strategy provided for balanced representation by age and size and between general and specialized-technical universities. For each country, one highly reputed research university and one less reputed university were included. The survey used five different questionnaires for respondents in different positions: central government of the university (four questionnaires for rector/president, board member, senate member, central administrator), and middle managers (one questionnaire for deans and department heads). A number of the questions were repeated identically in the different questionnaires. ‘Shop-floor’ staff, meaning academics not active in government roles, was not asked to participate. In spite of this limitation, the dataset offers a large database on how academics involved in decision-making power perceive their roles and judge the influence of evaluation on university strategies and activities.

Many questions elicit perceptions and opinions, and we assume that these are valid. We use Cronbach’s alpha to estimate the reliability of the Likert scales used in the questionnaires. Table 5.1 presents the list of

Table 5.1 List of variables used in the analysis

<i>Label</i>	<i>Question wording</i>	<i>Respondents</i>	<i>Items analysed</i>
DMP	Please indicate the extent to which actors within your university have actual decision-making power for the listed issues	Central government and middle managers	Setting the rules and procedures for evaluations of units (faculties, institutes, etc.) Evaluating academic individual performance
EV-QA	How do you rate the effects of evaluation and quality assessment as regard to the following issues?	Central government and middle managers	Positive effects (Improving teaching quality, improving research quality, Increasing transparency, Improving strategic decision-making) Negative effects (Conflicts between managers and academics, Conflicts among academics, Academic activity subjected to more norms and rules, Opportunistic behaviors in teaching and research, Constraints on academic freedom)
EXTEVA	How would you rate the influence of external evaluation and quality assessment on your university in terms of ...	Central government	Control impact = Personal impacts (e.g. career of individual academics), Teaching programs (content impact), Research priorities Distributive impact = Financial impacts (e.g. budgets of institution and academic units) Evaluation of research quality/excellence Evaluation of teaching quality Accreditation
INFITEM	How would you rate the influence of the following items on the activities carried out by your university?	Central government	

TRUE survey variables used for the current study and the items considered in the analysis.

We enlarge the information set by adding some further variables intended to help reveal more about the universities' activities and the use of external evaluation in internal affairs.

A first group of variables concerns the universities themselves:

- Age of the university—The variable is dummied as 'old' (before 1900), 'middle' (between 1900 and 1960) and 'recent' (1961 onward). The cluster analysis considers the logarithm of the age.
- Size—This is measured by number of students enrolled and considers the following categories: 'small' (up to 10,000), 'medium' (10,000–24,999), 'large' (25,000 and more). The cluster analysis again considers the logarithm of the absolute numbers.
- The level of disciplinary specialization—The variable is expressed as a continuous index of disciplinary concentration from 0 = low specialization to 1 = high specialization.
- Managerialism and Collegialism - These two variables are constructed based on the perceptions of the TRUE survey respondents, who responded to questions asking them to rate the managerial culture and the collegial culture of their home universities, using a Likert scale from 1 (very low) to 5 (very high).¹

Another group of variables characterizes the country contexts of the universities (see Table 5.2). This information is useful in understanding what circumstances and specific policies concerning universities enable similar or different uses of evaluation in general, and even of particular types of evaluation.

The dimensions considered are:

- Higher education research and development (HERD)—Public investment in higher education as a percentage of GDP (source: EUROSTAT).
- New Public Management orientation of the country—The level of managerial orientation for the country university system. The categories are 'low', 'medium' and 'high', based on the scoring of national information derived from the literature (national charts in Paradeise et al. 2009).

Table 5.2 Country contexts for universities: degree of formal autonomy concerning evaluation; degree of general institutional autonomy; levels of NPM and HERD (% of GDP)

	<i>Internal evaluation systems for teaching*</i>	<i>Internal evaluation systems for research*</i>	<i>Take part in external quality assessment for teaching*</i>	<i>Take part in external quality assessment for research</i>	<i>Mean of institutional autonomy (70 issues)*</i>	<i>NPM**</i>	<i>***HERD</i>
UK	1	3	0	3	0.635	3	0.44
FR	2	3	0	0	0.367	1	0.47
IT	3	3	3	0	0.457	1	0.35
NL	1	2	0	1	0.409	3	0.62
NO	2	3	0	3	0.547	2	0.51
PT	1	3	0	0	0.498	2	0.50
CH	3	3	3	3	0.439	1	0.83
DE	3	1	2	1	0.452	1	0.51

Sources *Formal autonomy—own elaboration from TRUE project ‘Formal Autonomy’ dataset; **NPM—own elaboration from Paradeise et al. 2009; ***HERD—EUROSTAT (2011)

- Formal autonomy (general institutional autonomy)—The level of formal institutional autonomy for universities, as granted under national laws and regulations;
- Formal autonomy concerning evaluation—The level of formal institutional autonomy in terms of freedom to take part in external evaluations - of teaching, of research; the level of institutional autonomy in terms of freedom to establish internal evaluation systems—for teaching, for research.

The level of formal autonomy on the specific area of evaluation is derived from a TRUE descriptive dataset, completed in 2011. It covers four issues: (i) the freedom of the universities to set up internal evaluation systems for teaching and (ii) to set up internal evaluation systems for research; (iii) the freedom of the universities to take part in external quality assessment for teaching, and (iv) to take part in external assessments for research.²

The data from the TRUE survey do not permit consideration of different disciplines, of gender differences, or of the age or seniority of the respondents, and do not cover the opinions of academics at the

‘shop-floor’ level (with no governing responsibilities). Thus, the survey results can supply information on how the academics entitled to formal internal power perceive the role of the different forms of evaluation (research evaluation, teaching evaluation; quality assurance), but they do permit consideration of the influences that personal and professional characteristics are likely to produce.

desCriPtive And CorreLAtion AnALyses

This section presents the analyses of the data using descriptive statistics and correlations, illustrating the general characteristics of the sample of universities surveyed and providing empirical evidence of the positive and negative correlations between variables.

Descriptive Analysis

Table 5.3 shows the descriptive statistics of the respondents’ perceptions—both the top management (Top-M) and middle management (MM)—about the strength of effects produced by evaluation on central and middle management level.

A high percentage (about 61%) of respondents consider that the effect produced by external evaluation for distributive purposes is very high, and the same holds true concerning the effects of external evaluation impacting on control uses, although the percentage of respondents rating ‘high’ is not so large as in the former case. However, if we examine the different items that compose ‘control use’ we find that the influence of external evaluation on the specific area of setting research priorities is perceived as high and very high by 78% of the respondents.

Accreditation, evaluation of teaching quality, and evaluation of research quality and excellence are all considered as having a high influence on the activities carried out by the university. However, compared to the other two types of assessment, research evaluation achieves an exceptionally high status in the percentage of respondents’ perceptions. As to the effects produced by evaluation, it is interesting to note that top and middle management have largely the same positive perceptions about the capacity of evaluation and QA to improve efficiency and effectiveness. On the other hand, they show more diversified opinions about the capacity for generating negative effects. Still, for both management levels, the large parts consider that evaluation and QA have low impact in generating conflicts within the universities.

Table 5.3 Descriptive statistics of evaluation perceived by respondents

<i>Perceptions on evaluation</i>	<i>Level impacted by evaluation</i>	<i>Low* %</i>	<i>Some* %</i>	<i>High* %</i>	<i>Total %</i>
Effect and impact of evaluation: on Control	Top-M	23.2	26.7	50.1	100.0
Effect and impact of evaluation: on Distribution	Top-M	12.6	23.0	64.4	100.0
Influence of: Accreditation	Top-M	14.2	31.3	54.5	100.0
Influence of: Evaluation of teaching quality	Top-M	12.1	28.9	58.9	100.0
Influence of: Evaluation of research quality/excellence	Top-M	4.9	13.4	81.7	100.0
Evaluation - generating efficiency and efficacy	Top-M	23.6	13.7	62.7	100.0
	MM	25.6	9.7	64.6	100.0
Evaluation - generating conflicts	Top-M	42.3	24.4	33.2	100.0
	MM	44.3	22.9	32.8	100.0
Actual decision-making power about evaluating academic individual performance perceived by Central level	Top-M	43.3	34.1	22.6	100.0
	MM	38.8	40.0	21.2	100.0
Actual decision-making power about setting the rules and procedures for evaluations of units perceived by Central level	Top-M	9.5	30.9	59.6	100.0
	MM	10.2	32.8	57.0	100.0
Actual decision-making power about evaluating academic individual performance perceived by Middle level	Top-M	10.2	48.9	40.9	100.0
	MM	24.7	48.3	27.0	100.0
Actual decision-making power about setting the rules and procedures for evaluations of units perceived by Middle level	Top-M	13.7	36.2	50.1	100.0
	MM	18.9	42.9	38.3	100.0

* 'Low' recodes 'very low' and 'low' of the Likert scale; 'high' recodes 'very high' and 'high' of Likert scale

Source Own elaboration from TRUE dataset

Top management consider that both the central and middle levels hold high decision-making power (59.6 and 57.0% respectively) over setting the rules and procedures for the evaluation of the university units. Thus the central government does recognize that it holds strong

power over this issue, although they share it with middle management. Concerning the evaluation of individual performance, top management do not indicate a concentration of decision-making power at either the central or middle management level, although a higher number of respondents perceive the power held by the central government as lower than the one in the hands of the faculties.

Middle managers perceive that they hold a role of ‘some’ importance in both the decision-making processes about evaluation (for evaluation of units and individual performance), although more than 50% consider the decision-making power on setting the rules and procedures for the evaluation of units as highly concentrated at the central level.

Summing up, the central government and middle managers show a convergence in their perceptions of the concentration of decision-making power in the universities, as far as concerns the assessment of the units. This observation conforms with their perceptions of a high effect of external evaluation on distributive uses, and with their consideration that research evaluation produces higher effects on university activities than do other types of evaluation.

Correlations

Table 5.4 presents the main significant correlations, observed at the university level. The figures shown are the actual values from the correlation matrix, therefore a positive value always means a ‘the more, the more’ association while a negative value means a ‘the more, the less’ association. The average values presented are the scores assigned by the respondents within the universities (both the respondents from central government bodies and faculty/department levels).

We can make three preliminary observations. First, it is interesting that the three variables representing the characteristics of the respondent’s university (size, age, disciplinary specialization) are not correlated with the answers on evaluation issues, as observed in this study.³ This indicates that the perceptions of the respondents, involved in the university central government and middle management levels, are not influenced by certain key characteristics of the organizations where they work.

Second, one of the country-level features, the NPM orientation of the national university system and one institutional-level characteristic of universities, namely their managerial culture, are frequently associated in a significant way to the distribution of decision-making power and the

Table 5.4 Correlation matrix with level of statistical significance (italic)

	4	5	6	7	8	9	10	11	12	13	14	15	16	17
4	1													
5	0.192	1												
	<i>0.348</i>													
6	0.037	0.049	1											
	<i>0.859</i>	<i>0.810</i>												
7	-0.260	0.151	0.262	1										
	<i>0.200</i>	<i>0.460</i>	<i>0.196</i>											
8	-0.582**	0.246	0.122	0.693**	1									
	<i>0.002</i>	<i>0.226</i>	<i>0.554</i>	<i>0.000</i>										
9	0.620**	0.301	0.081	-0.222	-0.377	1								
	<i>0.001</i>	<i>0.135</i>	<i>0.692</i>	<i>0.275</i>	<i>0.058</i>									
10	0.670**	0.586**	0.208	0.068	-0.236	0.607**	1							
	<i>0.000</i>	<i>0.002</i>	<i>0.307</i>	<i>0.743</i>	<i>0.246</i>	<i>0.001</i>								
11	0.166	0.390*	0.079	-0.086	0.049	0.004	0.334	1						
	<i>0.419</i>	<i>0.049</i>	<i>0.701</i>	<i>0.676</i>	<i>0.811</i>	<i>0.986</i>	<i>0.096</i>							
12	0.417*	0.307	0.111	0.004	0.049	0.236	0.415*	0.638**	1					
	<i>0.034</i>	<i>0.127</i>	<i>0.589</i>	<i>0.985</i>	<i>0.814</i>	<i>0.246</i>	<i>0.035</i>	<i>0.000</i>						
13	0.591**	-0.081	-0.189	-0.302	-0.356	0.134	0.298	0.138	0.494*	1				
	<i>0.001</i>	<i>0.696</i>	<i>0.354</i>	<i>0.134</i>	<i>0.074</i>	<i>0.512</i>	<i>0.139</i>	<i>0.502</i>	<i>0.010</i>					
14	0.207	-0.014	-0.437*	-0.233	-0.019	0.015	0.011	0.030	0.037	0.413*	1			
	<i>0.311</i>	<i>0.948</i>	<i>0.026</i>	<i>0.252</i>	<i>0.927</i>	<i>0.943</i>	<i>0.958</i>	<i>0.884</i>	<i>0.859</i>	<i>0.036</i>				
15	0.414*	0.354	0.025	-0.249	-0.226	0.506**	0.388	0.175	0.124	0.108	0.062	1		
	<i>0.035</i>	<i>0.076</i>	<i>0.904</i>	<i>0.220</i>	<i>0.267</i>	<i>0.008</i>	<i>0.050</i>	<i>0.394</i>	<i>0.547</i>	<i>0.601</i>	<i>0.762</i>			
16	0.210	0.452*	0.045	0.042	0.200	-0.044	0.232	0.486*	0.352	0.224	0.250	-0.005	1	
	<i>0.303</i>	<i>0.020</i>	<i>0.827</i>	<i>0.838</i>	<i>0.326</i>	<i>0.832</i>	<i>0.254</i>	<i>0.012</i>	<i>0.078</i>	<i>0.272</i>	<i>0.217</i>	<i>0.981</i>		
17	0.481*	0.480*	0.083	0.038	0.018	0.168	0.554**	0.655**	0.646**	0.445*	0.063	0.027	0.700**	1
	<i>0.012</i>	<i>0.012</i>	<i>0.688</i>	<i>0.853</i>	<i>0.030</i>	<i>0.411</i>	<i>0.003</i>	<i>0.000</i>	<i>0.000</i>	<i>0.023</i>	<i>0.759</i>	<i>0.896</i>	<i>0.000</i>	

(continued)

Table 5.4 (continued)*Legend of variables*

1	Institutional size (continuous) [§] Institutional concentration (continuous) [§] Institutional age (continuous) [§]
2	Country NPM
3	Managerial culture
4	Collegial culture
5	Decision-making power (DMP) about setting the rules and procedures for evaluations of units Top-M DMP about evaluating academic individual performance Top-M DMP about setting the rules and procedures for evaluations of units MM
6	DMP about evaluating academic individual performance MM
7	DMP about evaluating academic individual performance MM
8	Influence of: Evaluation of research quality/excellence
9	Influence of: Evaluation of teaching quality
10	Influence of: Accreditation
11	Evaluation generating conflicts
12	Evaluation generating efficiency and efficacy
13	External evaluation impacting distributive uses
14	External evaluation impacting controlling uses
15	
16	
17	

[§] The variable is omitted from the table

* = $p < 0.05$

** = $p < 0.01$

Source: Own elaboration from TRUE dataset

use of evaluation. This indicates that evaluation from a summative perspective is linked to the strength of the managerial paradigm at both the national and the institutional levels, when other structural factors such as the size, heterogeneity of disciplinary fields and the age of the institutions are held equal. This observation is consistent with what is often suggested in the literature (Neave 1998, 2012).

Third, the likelihood of negative effects on university activities arising from evaluation (conflicts, bureaucratic load, opportunistic behaviors, constrains on academic freedom) is positively correlated to the implementation of accreditation (0.413*). At the same time, the negative effects resulting from evaluation are less likely to be perceived where there is a collegial culture, whereas where a managerial culture is present, there is more likelihood of the declaration that evaluation generates both controlling (0.480*) and distributive (0.452*) effects.

A further result concerns the self-reinforcing mechanisms that evaluation could produce. These emerge in the linkages between the different types of evaluation (research and teaching evaluation) on the activities of the universities. Both the distributive and the controlling functions of evaluation are significant and positively correlated (respectively 0.486* and 0.655*) with 'influence of evaluation of research quality/excellence'. Again, evaluation under a summative orientation shows a tendency to affect the allocation of financial resources and the performance of individual academics, as well as the content of teaching programs and the research priorities, thus becoming a policy instruments to steer the organization. The controlling use of evaluation is particularly correlated to all of these influences of evaluation, even in the area of teaching (0.646*) and accreditation (0.445*).

Data show that the distributive use of evaluation has a high and significant association with a managerial culture of the university (0.452*). Interestingly enough, no specific correlations emerge between the mentioned use and the distribution of the decision-making power within the universities.

A different picture emerges when the use of external evaluation for control purposes is concerned. The more external evaluation is used for controlling purposes, the more research evaluation and teaching evaluation have a strong influence on the activities of the universities. Further, the more a country has an NPM orientation in the university system, the more universities tend to allocate the decision-making power over both the assessment of individual performance and the evaluation of

units to faculties and departments (0.670** and 0.620** respectively). Interestingly, a strong negative correlation is found between the national degree of NPM and the decision-making power held by top management over evaluation of individual performance (-0.582**). Therefore, one can expect that the more the managerial orientation of the country, the less the universities centralize the power on decisions related to individual assessment.

These results suggest that the division of decision-making power inside the university follows very specific paths, with the concentration at the middle management level (as opposed to top management) being more clearly and strongly related to obtaining effects from evaluation. At the same time, a managerial orientation in university culture is likely to covariate with stronger decision-making power for the faculties, in the assessment of individual performance.

In other words, evaluation can serve as an instrument to be used by the university central government and the faculties, but the traditional weak hierarchy, which is characteristic of universities, means that when the specific use of external evaluation is present it could also serve in the redistribution of power between the different decision-making levels of the internal government. The hierarchical development of universities is probably fated to become stronger than in the past,⁴ but the internal organization of the university will not necessarily assume a truly 'vertical' configuration in the form of the assumption of further levels of power at the upper levels of internal government. These results are consistent with the observations that in more managerial universities, collegial culture increases above all when middle managers believe that evaluation has positive impacts (Marini and Reale 2015).

ComPARison of the EffeCts of iNstitutionAL And nAtionAL feAtures

The final step of the analysis investigates whether either the national dimension or institutional features emerge as important factors in influencing how external evaluation is likely to affect the governance of the universities. The statistical procedure is that of stage-by-stage hierarchical clustering, which, at any stage, is particularly suitable to understanding what cases (HEIs) appear to be more similar to each other.

We use the means at the institutional level from the original dataset of 697 respondents to identify the 26 cases. The aim is to observe how the universities group in clusters on the basis of: (i) national features; (ii) the characteristics of the individual organizations; and (iii) the variables regarding the way top- and middle-management academic personnel perceive the impact of external evaluation on their own higher education institutions. The cluster analysis serves to test three possibilities:

- i) The effect of external evaluation on governance of universities is mainly influenced by the national awarding of general institutional autonomy (AU_tot), and the degree of specific autonomy in evaluation issues (the four variables of autonomy in internal teaching and research evaluation, participation in teaching and research quality assessment). This first possibility would prove that European universities are still influenced by their nation-states in terms of the ways they might enact a more 'up-to-date' organizational aspect.
- ii) The governance of universities is influenced above all by their institutional features. This result would demonstrate that some organizational aspects of governance, particularly the distribution of decision-making power between the central and middle management levels, and the influence evaluation could have in the universities, can be explained primarily by the characteristics of the HEI itself (more established, larger, greater disciplinary specialization).
- iii) None of these two sets of constraints explain the governance of universities. This result could be understood as the capacity of European HEIs to change their governance models in response to evaluation, accepting greater or lesser extent of influences according to other variables not examined in the current study, such as the specific leadership.

At this stage of the analysis the aim is to see which universities are more similar in governance, based on the means of the perceptions at the institutional level (cfr. Annex 1). If on the contrary there are differences in opinions between the universities⁵ we wish to assess which forces might give rise to these observed institutional differences.

Reading Fig. 5.1 from left to the right, it is possible to see for each university, which is its most similar. The ‘red line’ in the figure results in

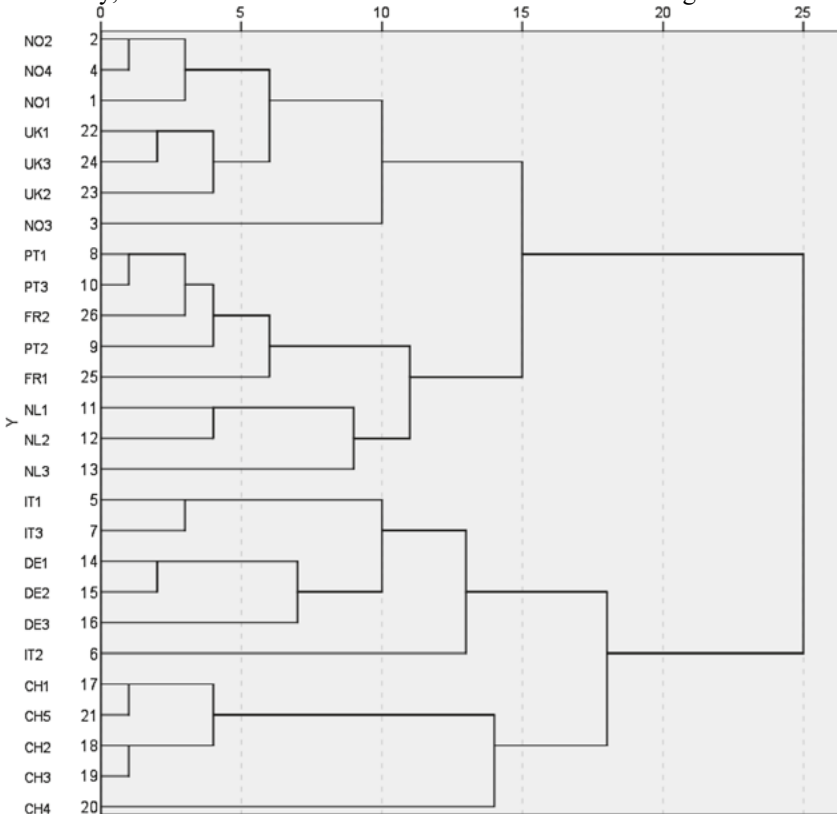


Fig. 5.1 Dendrogram of the hierarchical cluster analysis of all the variables introduced in Tables 5.1, 5.2 and 5.3. *Source* Own elaboration from TRUE dataset

the identification of four clusters, determined by some main discriminating variables. Table 5.5 lists only those six variables whose means by cluster show a substantial spread, meaning those primarily responsible for the formation of the four clusters.

The first group of universities is based in UK and Norway. In this cluster, the strength of top management’s decision-making power is low in the areas of evaluation aimed at both distributive and control uses, probably due to the fact that in national systems offering high autonomy

Table 5.5 Characteristics of the hierarchical : descriptors for four groups (selection of discriminating variables in constructing the clusters)

<i>Cluster^a</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
Country of HEIs	NO, UK	FR, PT, NL	DE, IT	CH
AU (institutional autonomy)_total	High	Medium	Medium	Medium
AU_teaching_internal	Low	Very Low	High	High
AU_research_external	Very High	Low	Medium	Very High
DMP at Top-M level (evaluation of individuals)	Low	Very Low	High	High
Influence of accreditation	High	Very High	High	Low
Evaluation producing efficiency and efficacy	High	High	Low	Low

^aThe list of variables presented in the table is less than the full set used in the analysis, concentrating on those that are mainly responsible for the clustering. For a larger display, see Annex 1

Source Own elaboration from TRUE dataset

in research, the evaluation is also developed at the individual university level. Evaluation has highly positive influences in producing efficiency and effectiveness; furthermore this is the only cluster to reveal a substantial distinction from the others in the sense of disciplinary concentration. Nevertheless, this might be due to the selection of the seven specific universities in Norway and the UK, since none of this sample is particularly specialized in terms of disciplines. In other words, although it is empirically true, the evidence does not seem sufficient to deduce that this institutional feature of disciplinary concentration plays a more important role than the national systemic features.

The universities in the second cluster (situated in France, Portugal, and the Netherlands) are characterized by very low formal autonomy in the area of internal teaching evaluation and a very low degree of top management decision-making power over individual evaluation. However, the universities in the cluster are very highly concerned with the importance of accreditation, probably due to nationally led evaluation of the curricula, as well as implementation of quality assurance. In these systems, evaluation of research appears to be less predominant in shaping the governance of the institutions.

Universities in the third cluster (German, Italian HEIs) have high institutional autonomy in internal teaching issues and, coherent with this, accreditation has a high importance. In contrast, evaluation is generally seen as creating less positive outcomes. In these two countries, the top management of universities holds a high decision-making power in evaluation of individuals, an effect that can be linked to the low autonomy in regards to external evaluation of research.

The fourth cluster, consisting only of Swiss universities, has a pattern of high autonomy in regards to external evaluation of research and of teaching, but a medium level of institutional autonomy (AU_total), while the top management hierarchies have high decision-making power on individual evaluation issues. Still these various characteristics for the cluster 4 universities do not seem to lead to any particular influence from accreditation, or even any positive consequences from evaluation.

Interestingly, there are no differences between the clusters either in terms of perception of managerial and collegial cultures, or in terms of distributive and controlling uses of the evaluation; none of the clustering can be explained by substantial differences in the variables that describe the individual universities, such as their age, size or disciplinary concentration.

To sum up, although there is a tightening process on European universities, intended to strengthen them as robust organizations, it is actually the national system that is foremost in driving the degree of effectiveness of the changing governance induced by external evaluation. In other words, the 'Evaluative State' (Neave 1998, 2012) is the most influential actor in triggering or dictating the pace of the change in internal governance due to evaluation, while the features of the individual universities appear to be less important. Thus, the transformation of governance towards the state of a 'complete organization' is detected in rough accordance with the degree and types of formal autonomy that each nation affords to its universities.

ConCLusions

This chapter investigates the extent to which external evaluation of universities is likely to transform the governance of the institutions impacting the distribution of the decision-making power. The study adopts a comparative perspective concerning 26 universities in eight European

countries. The analysis uses the results of the TRUE survey on the perceptions that those academics involved in the decision-making levels have about the impact of the evaluations on their universities' activities and governance. It also investigates these perceived effects in relation to several institutional characteristics of the universities covered by the survey, as well as to two key policy features of the national university systems, namely the level of formal autonomy granted and the NPM orientation in regards to universities.

The paper tested three propositions, of which the first is not confirmed by the data. This is the particular proposition that the academics active at the central government and middle management levels will perceive greater negative effects, as external evaluation increasingly impacts on decisions regarding the distribution of vital resources. The data instead confirm the second proposition, which is that when evaluation is used for controlling purposes, it has differing transformative effects on the levels of internal government (central, middle managers), distributing the decision-making power so as to mitigate the concentration in the hands of the central bodies. In other words, the data show that regardless the managerial or collegial orientation in the university, both hierarchical levels can have strong decision-making power over evaluation, a finding that is consistent with the literature on universities as 'non-complete' organizations.

As to the third proposition, research and teaching evaluation exert different influences between universities, contributing to shape their institutional configurations; the strength of the types of evaluation can be explained by the formal autonomy granted to the universities (at national level) over the types of evaluation themselves.

Turning back to the original research questions, the formal institutional autonomy granted to universities under the European country systems, as well as the NPM orientation of the country itself, emerge as characteristics shaping the way in which external evaluation is likely to transform the governance of universities, rather than any features at the institutional level.

Evaluation as a steering instrument is ostensibly geared toward maximizing or securing a minimum level of teaching and research quality. However, improvement of quality is not an activity that lends itself to the exclusive use of top-down steering. Instead, it responds better to some

kind of balance between top-down and middle-level influences, resembling 'soft' forms of power. Although the time and the purposes of external evaluation exercises are set by public authorities outside the higher education institutions, academics can influence the intended uses of evaluation through the roles they play in institutional government. Different evaluation 'regimes' also emerge depending on: (i) the types of evaluation employed (research evaluation, teaching evaluation, other forms of assessment), (ii) the national traditions (i.e. degree of New Public Management implementation in the different countries, and degrees of formal autonomy), and (iii) the ultimate use of the evaluation outcomes (both distributive and controlling uses). In these respects, universities in Europe resemble more a constellation of national configurations having certain levels of similarities than a unitary system, a fact that indicates the importance of research questions aimed at investigating if and how policies at supra-national level are able to transform universities, moving them toward integration.

Notes

1. The respondents were asked: 'To what extent do you agree with the following statements as regards this university: My university has a strong managerial culture; My university has a strong collegial culture.'
2. In response to the questions concerning the universities' freedom regarding the four evaluation issues, the case of no autonomy is indicated by the answer: 'No: this is required and government prescribes the process'; Low autonomy is indicated by 'No: this is required but university decides on methods that will be evaluated by government'; Some autonomy with limitations corresponds to the answer 'No: this is required but university decides on the methods'; High autonomy corresponds to the answer, 'Yes'.
3. For simplicity of presentation, these crosses are omitted from Table 4.
4. The data from the TRUE survey are insufficient for conclusive comparison to the characteristics of any preceding governance systems of the universities.
5. For instance the mean about the influence of accreditation on governance in university NO1 can result as different from the mean opinion in NL2.

**Annex: mAns At institutionAL Level used for the CLuster
AnALysis And resPeCtive mAns of the four CLusters**

HEI	INST CONCENTRATION	COUNTRY_NPM	Influence_ev_Research	Influence_Ev_Teaching	Influence_Accreditation	Managerial culture	Collegial culture	Distributive Power	Controlling Power	DMP_ev_ind_held by TopM according to_TopM	Ev_effs_TopM	DMP_ev_ind_held by MM	Log_age	Log_size
PT1	0.21	2	2.70	2.35	2.50	3.54	3.29	2.45	2.65	2.18	2.65	3.00	3.47	8.63
PT2	0.19	2	2.60	2.29	2.71	3.44	3.44	2.58	2.48	1.50	2.46	1.67	4.61	10.13
PT3	0.21	2	3.00	2.71	2.86	3.21	3.47	2.86	2.71	1.71	2.86	2.27	3.64	9.32
NL1	0.49	3	3.00	3.00	3.00	3.50	3.93	2.50	3.00	1.50	2.67	1.00	3.56	9.47
NL2	0.99	3	2.82	2.73	2.82	3.10	3.70	2.09	2.64	1.43	2.44	1.86	4.01	8.86
NL3	0.27	3	3.00	2.62	2.85	3.74	3.44	2.77	2.92	1.21	2.50	1.50	5.93	10.28
FR1	0.41	1	3.00	2.63	2.71	2.93	3.25	2.75	2.75	2.00	2.38	2.14	3.69	10.34
FR2	*	1	2.78	2.78	2.67	2.64	3.36	2.33	2.50	1.70	2.13	1.82	3.66	9.39
NO1	0.25	2	2.50	2.25	2.25	2.79	3.68	2.50	2.50	1.67	2.75	1.29	5.30	10.22
NO2	0.28	2	2.67	2.22	2.00	2.76	3.62	2.00	2.22	1.29	2.44	1.13	4.17	9.56
NO3	0.16	2	2.57	2.57	2.71	3.30	3.30	2.00	2.14	1.50	2.86	1.00	2.83	8.97
NO4	0.22	2	2.83	2.33	2.33	3.59	3.46	2.20	2.20	1.14	2.50	1.00	4.62	9.91
UK1	0.18	3	3.00	2.65	2.50	3.70	3.71	2.89	2.96	1.53	2.16	1.25	4.91	9.67
UK2	0.18	3	2.77	2.85	2.77	4.00	3.38	2.75	2.67	1.70	2.70	1.50	3.74	10.34
UK3	0.18	3	2.89	3.00	2.70	4.08	3.77	3.00	3.00	1.88	2.50	1.48	4.98	9.78

(continued)

HEI	INST CONCENTRATION	COUNTRY_NPM	Influence_ev_Research	Influence_Ev_Teaching	Influence_Accreditation	Managerial culture	Collegial culture	Distributive Power	Controlling Power	DMP_ev:nd_held by TopM according to_TopM	Ev:ffs_TopM	DMP_ev:nd_held by MM	Log_age	Log_size
IT1	0.19	1	2.72	2.16	2.11	3.51	3.55	2.50	2.25	1.74	2.73	1.80	3.50	9.75
IT2	0.21	1	2.94	2.63	2.38	3.30	3.83	2.31	2.31	2.00	2.63	2.14	6.83	11.40
IT3	0.99	1	2.86	2.38	2.35	3.60	3.77	2.71	2.52	2.13	2.45	2.00	4.65	10.08
DE1	0.30	1	3.00	2.63	1.57	4.00	3.39	2.57	2.57	2.20	2.43	2.25	4.96	10.03
DE2	0.22	1	3.00	2.65	2.38	3.22	3.50	2.59	2.64	1.92	2.31	2.00	6.32	9.83
DE3	0.18	1	2.47	2.40	2.40	3.06	3.60	2.43	2.14	2.25	1.79	2.25	3.71	9.77
CHI	0.22	1	2.47	2.02	2.33	3.13	3.51	2.33	2.26	1.81	2.16	1.80	6.31	9.12
CH2	0.36	1	2.81	2.43	1.63	3.49	3.85	2.40	2.26	2.00	2.24	2.44	5.05	9.25
CH3	0.49	1	2.89	2.67	2.38	4.33	3.33	2.44	2.89	2.00	1.75	2.50	4.95	8.47
CH4	0.49	1	2.83	2.50	2.17	3.90	4.00	2.67	2.83	2.50	2.17	2.00	2.71	7.68
CH5	0.33	1	2.73	2.65	2.29	3.32	3.88	2.29	2.35	1.91	2.26	2.00	6.16	9.16
CLUSTER 4	0.38	1.00	2.75	2.45	2.16	3.63	3.71	2.43	2.52	2.04	2.12	2.15	5.04	8.74
CLUSTER 3	0.40	2.13	2.86	2.64	2.76	3.26	3.49	2.54	2.71	1.65	2.51	1.91	4.07	9.55
CLUSTER 2	0.40	2.13	2.86	2.64	2.76	3.26	3.49	2.54	2.71	1.65	2.51	1.91	4.07	9.55
CLUSTER 1	0.21	2.43	2.75	2.55	2.47	3.46	3.56	2.48	2.53	1.53	2.56	1.23	4.36	9.78

ReferenCes

- Birnbaum, R. (2004). The end of shared governance: Looking ahead or looking back. *New Directions for Higher Education*, 127, 5–22.
- Bleiklie, I., & Kogan, M. (2007). Organization and governance of universities. *Higher Education Policy*, 20, 477–493.
- Brunsson, N., & Sahlin-Andersson, K. (2000). Constructing organizations: The example of public sector reform. *Organization Studies*, 21(4), 721–746.
- Capano, G. (2011). Government continues to do its job: A comparative study of governance shifts in the higher education sector. *Public Administration*, 89(4), 1622–1642.
- Dahler-Larsen, P. (2012). *The Evaluation Society*. Stanford University Press.
- Hammarfelt, B., de Rijcke, S. (2015). Accountability in context: Effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of Arts at Uppsala University. *Research Evaluation*, 24, 63–77.
- Hansson, F. (2006). Organizational use of evaluations: Governance and control in research evaluation. *Evaluation*, 12(2), 159–178.
- Huisman, J., Rebora, G., & Turri, M. (2007). The effects of quality assurance in universities empirical evidence from three case studies. In L. Purser, L. Wilson, & E. Froment (Eds.), *Introducing Bologna objectives and tools* (pp. 1–28). Stuttgart: Raabe.
- Kooiman, I. (2003). *Governing as Governance*. London: Sage.
- Marini, G., Reale, E. (2015). How does collegiality survive managerially led universities? Evidences from and European survey. *European Journal of Higher Education*, doi:[10.1080/21568235.2015.1070676](https://doi.org/10.1080/21568235.2015.1070676).
- Molas-Gallart, J. (2012). Research governance and the role of evaluation: a comparative study. *American Journal of Evaluation*, 33(4), 583–598.
- Musselin, C. (2007). Are Universities specific organizations? In G. Krücken, A. Kosmützky, & M. Torck (Eds.), *Towards a multiversity? Universities between global trends and national traditions*. New Brunswick (USA): Transaction Publishers.
- Musselin, C. (2013). How peer review empowers the academic profession and university managers: Changes in relationships between the state, universities and the professoriate. *Research Policy*, 42(5), 1165–1173.
- Neave, G. (1998). The evaluative state reconsidered. *European Journal of Education*, 33(3), 265–284.
- Neave, G. (2012). *The evaluative state, institutional autonomy and re-engineering higher education in Western Europe*. Great Britain: Palgrave Macmillan.
- Paradeise, C., Reale, E., Bleiklie, I., & Ferlie, E. (Eds.). (2009). *University governance: Western European comparative perspectives*. Dordrecht: Springer.

- Reale, E., & Seeber, M. (2010). Organisation response to institutional pressures in Higher Education: The important role of disciplines. *Higher Education*, 61(1), 1–22.
- Reale, E., & Seeber, M. (2013). Instruments as empirical evidence for the analysis of Higher Education policies. *Higher Education*, 65(1), 135–151.
- Scriven, M. (1967). The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), *Perspectives on curriculum evaluation*. Chicago: Rand McNally and Co.
- Seeber, M., Lepori, B., Montauti, M., et al. (2014). European universities as complete organizations? *Understanding Identity, Hierarchy and Rationality in Public Organizations*, *Public Management Review*, doi:10.1080/14719037.2014.943268.
- Stame, N. (2006). Governance. *Democracy and Evaluation, Evaluation*, 12(1), 7–16.
- Taras, M. (2005). Assessment: Summative and formative. *Some Theoretical Reflections*, *British Journal of Educational Studies*, 53(4), 466–478.
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21(1), 1–19.
- Whitley, R. (2007). ‘Introduction’. In R. Whitley, & J. Glaser (Eds.), *The changing governance of the sciences* (pp. 3–27). Dordrecht: Springer.
- Whitley, R. (2012). Transforming universities: national conditions of their varied organizational actorhood. *Minerva*, 50, 493–510.

Authors’ biogrAPhy

Dr. Emanuela Reale Social scientist, is Senior Researcher at the Research Institute on Sustainable Economic Growth IRCRES, National Research Council CNR of Italy. Her main areas of interest are higher education institutions and policy, governance of universities and public research organizations, R&D funding, research evaluation, and STI indicators. She was principal investigator or coordinator in several national and European projects dealing with higher education and evaluation. She was Vice President of the Italian Evaluation Association- AIV in 2009–2013, and Vice President of the European Forum for Studies on Policies for Research and Innovation-EU-SPRI (2009–2015). She is Member of the Board of the Consortium of Higher Education Researchers CHER. She published and served as referee in several international journals and books.

Giulio Marini is Research Associate at Centre for Global Higher Education (CGHE, www.researchcghe.org) at Institute of Education, UCL. He has a Ph.D. in methodology for social sciences from Sapienza University of Rome. Previously he was post-doc at Scuola Normale Superiore (Italy), Cipes (Portugal), and the Research Institute on Sustainable Economic Growth of the National Research Council (Italy). His main interests are in academic labour market, human resource management in higher education, and governance of higher education. His last publication appeared in *Higher Education*.