# Policy Instruments Matter!

## How governments' choice of policy mix shapes higher education performance in Western Europe

*Giliberto CAPANO*
*Scuola Normale Superiore*
giliberto.capano@sns.it

*Andrea PRITONI*
*Scuola Normale Superiore*
andrea.pritoni@sns.it

*Giulia VICENTINI*
*Scuola Normale Superiore*
giulia.vicentini@sns.it

**ABSTRACT**

*Over the last thirty years, many national higher education systems (HESs) in Europe have undergone structural changes following domestic pressure or international prescriptions. These changes have mainly been intended to enhance the overall university performance – conceived as students' access, quality of teaching and excellence in research. Almost all countries have decided to address these changes by adopting a similar policy design, fostering more organizational autonomy and differentiation, and promoting greater managerial steering. However, despite similar policy patterns being replicated nearly everywhere, indicators of performance still reveal remarkable variation across Western European countries. This leads to our main research question: what are the determinants of performance improvement in higher education? Thus far, neo-liberal economists and policy analysts have stressed the role of institutional autonomy, competitive funding mechanisms and the assessment of the quality of research and teaching in improving HESs' performance; however, this explanation seems to be an over-simplification of reality. Assumed causal explanations need to be refined specifically because similar methods have produced different outputs in different countries. We argue that differences in performance across national HESs depend on the mix of different types of policy instruments. We test this expectation with respect to eleven HESs in Western Europe: Denmark, England, Finland, France, Greece, Italy, the Netherlands, Ireland, Norway, Portugal, and Sweden. Utilizing a large dataset containing all the changes in policy instruments undertaken in the last 25 years, we turn to qualitative comparative analysis (QCA) to unravel conjunctural causation.*

Paper presented at 3rd *International Conference on Public Policy* (ICPP3) June 28-30, 2017 – Singapore, panel T02P17–*Policy Tools for Environment and Social Policies*

## 1. Introduction

According to the mainstream literature related to governance change in higher education (HE) (Huisman 2009; Paradeise *et al*. 2009; Shattock 2014; Capano, Regini 2014), governments have redesigned governance systems to make HE institutions more accountable by intervening with the introduction of rules governing the allocation of public funding and tuition fees, the recruitment of academics, and the evaluation and accreditation of institutions. To accomplish this, these countries have turned to a similar policy formula (the so-called 'steering at a distance' governance arrangement).

However, according to contrasting results in the literature, there is no clear evidence regarding whether and how the new governance template has delivered satisfactorily results. In this paper, we address this gap by assuming a policy instrument perspective, meaning that the actual national interpretation that each country has given to the common policy template in reforming HE governance can be assessed by focusing on the specific combinations of policy instruments that have been adopted at the national level. Thus, we assume that the way specific policy tools are set together determine the policy's performance. This perspective obviously assumes the relevance of combinatory causality and thus tries to overcome the intrinsic limits of variable-oriented research strategies while trying to rationalize the case-oriented perspective. We have pursued this research strategy by collecting data on the regulatory changes in HE in 11 countries over the last 25 years and then extracting and coding all the policy instruments adopted by those countries.

Our main goal is thus exploratory, as our intention is to demonstrate the degree to which the research design we have followed can be promising in filling the existing knowledge gap concerning the real effects of governance reforms in public policy.

The paper is structured as follows. In the second section, we present our policy instrument framework. We explain why it seems to be generally useful to study performance of governance shifts and how those shifts apply to the case of HE. We also reveal the reasons we chose qualitative comparative analysis (QCA) to develop our analysis. The third section introduces the research design and focuses on our justification of the case selection and timespan. We explain the operationalization of QCA outcomes and conditions and the process of data collection. Section four then presents some descriptive statistics and basic graphs concerning the policy instruments in use in the considered countries over recent decades. The fifth section presents the results of the QCA analysis, revealing the combination(s) of policy instruments that lead to better performance in HESs, both in teaching and research. Finally, the conclusion summarizes our preliminary results by addressing the applicability and usefulness of our dataset and theoretical and methodological framework in the context of future research in the field

and the broader issue of how to grasp methods of working on governance arrangements and their related policy mixes.

## 2. Governance arrangements and systemic performance in HE: an instrumental perspective

### *2.1 Governance reforms in HE*

Many scholars have underlined the apparent convergence towards the steering-at-a-distance mode in HE in recent decades (Braun and Merrien 1999; Paradeise *et al*. 2009; Huisman 2009; Shattock 2014). This governance arrangement is characterized by mixing the following instruments together: financial incentives to pursue specific outputs and outcomes in teaching and research**,** student loans**,** accreditation**,** ex-post evaluation conducted by public agencies, contract benchmarking**,** and provisions by the law for greater institutional autonomy (Gornitzka *et al*. 2005; Lazzaretti and Tavoletti 2006; Maassen and Olsen 2007; Trakman 2008; Capano 2011).

Studies in recent years have demonstrated that this convergence is certainly working to support general principles (more institutional autonomy, more evaluation, more competition), while the concrete ways through which the policies are made seem to be quite diverse. However, it should be noted that in terms of policy performance, the effects of these governance shifts have not been clearly assessed. In fact, the literature on higher education systems' (HESs') performance has mainly focused on a few aspects as key determinants of policy success (or failure): the mechanism of funding (Liefner 2003), the total amount of public funding (Winter-Ebmer and Wirz 2002), full institutional autonomy (Aghion *et al*. 2008), partisan dynamics (Ansell 2008), stratification (Willemse and De Beer 2012), or the type of loan system adopted (Flannery and O'Donoghue 2011). However, this strategy simply assesses whether certain variables have the power to influence the probability of the outcome changing as expected on average at the population level, regardless of contexts and often of interactions. These types of explanatory results seem weak and generally risk remaining very superficial. The salient point here is that focusing on a single dimension to assess the performance capacity of a governance arrangement is quite misleading. For example, the effects of shifting from a centralized governance system to a one in which universities are more autonomous cannot be analysed without contextualizing the change within its specific configuration, that is, by considering the other relevant dimensions (for example, how universities are funded, the system of degree accreditation, whether a national research evaluation assessment is present, etc.).

All in all, despite the significant governance shifts in HE, there is currently a lack of knowledge regarding both the actual nature of these changes in terms of existing combinations of adopted instruments and the policy performance of the new governance arrangements.

Thus, the explanatory gap calls for a more detailed description of the content of governance arrangements in terms of adopted measures and a more detailed perspective of the ways in which different policy instruments are combined.

### *2.2. Governance arrangements as a set of policy instruments*

Policies are steered by specific governance arrangements composed of rules, instruments, actors, interactions, and values (Capano, Howlett, Ramesh 2015). The implicit assumption of the governance literature is that different governance modes or arrangements present different results in terms of policy outcomes. However, empirical evidence on this issue has been lacking, especially because the main analytical focus in public policy has been the process of changing governance arrangements in terms of their content with respect to the actors involved, the distribution of power, and the adoption of 'new' policy instruments. Thus, there has not been enough focus on the policy results of these governance shifts in mainstream public policy.

However, there is an increasing awareness that pure types of governance arrangements do not actually work; instead, the main principles of coordination (hierarchy, market and network) are combined in various ways. All governance arrangements are basically hybrids and are characterized as working through policy mixes, that is, policy instruments belonging to 'different' instrument categories or pertaining to different policy paradigms/beliefs/systems/ideologies (Howlett 2005; Ring, and Schroter-Schlaack 2010; Capano, Rayner, Zito 2012). Thus, the existing set of adopted policy instruments can be conceptualized as specific portfolios, settings, and combinations from different types of policy instruments and bearing different working logic (Jordan *et al*. 2012; Schaffrin *et al*. 2014; Howlett and del Rio 2015).

However, how can the content of these policy mixes be described and understood, and how can their policy performance be assessed? In an attempt to fill these theoretical and empirical gaps, we adopt a bottom-up perspective by focusing on the basic unit of any governance mode – the policy instruments that can be adopted – and their possible combinations. This approach seems quite realistic; policy instruments are the operational, performance-related dimensions of governance arrangements.

Accordingly, we operationalize systemic governance arrangements in terms of adopted policy instruments and thus as specific sets of techniques or means by which

governments try to affect the behaviour of policy actors to direct them towards the desired results (Linder and Peters 1990; Vedung 1998; Howlett 2000; Salamon 2002).

There are numerous classifications by which policy tools can be ordered based on different criteria of analytical distinction, from coercion to the type of governmental source adopted (Ingram and Schneider 1990; Phidd and Doern 1983; Vedung 1998; Howlett 2011). All these typologies suggest different families of instruments. Our research framework focuses on the capacity of policy instruments to induce specific behaviours; thus, we need to consider the nature of the instruments and examine the different ways through which they induce action towards the expected result. In conducting this examination, the classical theorization of Vedung (1998) is useful. When focusing on the nature of substantive policy instruments, Vedung grouped those instruments by the basic inducement on which they relied to foster compliance.

By following this perspective, we can delimit four distinct families of substantial policy instruments that are have different – non-overlapping – capacities to induce behaviours: *expenditure, regulation*, *information* and *taxation*[1]. Each family bears a specific inducement. Expenditure induces remuneration, regulation induces behaviour control, information offers persuasion, and taxation – depending on the way it is designed – can engender behaviour control as well as remuneration. Notably, all four families of substantial tools can be employed by applying different methods of coercion that are dependent on how free individuals are to choose alternatives. Taxation can be quite coercive when a general tax increase is established, but it can also have a low degree of coercion when many targeted tax-exemptions exist. Regulation can be quite strong or weak according to the type of behavioural prescriptions that are provided for. Expenditure can lack coercion in the case of subsidies but be very demanding when delivering targeted funding. Information can be coercive when compulsory disclosure is imposed or lack strong coercion when monitoring is applied.

The four types of policy instruments we have decided to consider, as well as the types proposed by other policy instrument classifications, represent very general instrumental principles (which need to take specific forms to be practically applied). Thus, according to Salamon (2002), the shape through which the substantial instrument is designed to deliver the expected result is the important factor in terms of policy impact and potential performance. For every type of substantial policy instrument, there are

[1] In our perspective, taxation can be considered an autonomous substantial instrument. We are aware that in other typologies, taxation is the chief component of broader instruments. Phidd and Doern (1983) consider taxation to be a means of regulation (as it implies high coercion), while Hood (1983), following the same reasoning, holds partially to the 'authority' type (i.e., user charges) and partially to the treasury type (i.e., tax exemptions, tax expenditures). We believe that expenditure and taxation have different political and economic characteristics and present different ways of inducing or restraining institutional and individual behavior (Woodside 1983).

different methods of delivery that offer actual outlets through which those substantial instruments can affect reality. In addition, these instrument shapes should be considered the basic analytical unit when assessing how policies are made and thus how governance arrangements actually work in terms of policy performance.

Accordingly, the important factors in detecting the adoption of the expenditure instrument are the various forms through which expenditures can be delivered, such as grants, subsidies, loans, lump sum transfers, targeted transfers, etc. Regulation can be designed by imposing a specific behaviour, enlarging the range of opportunities, or establishing specific public organizations. Information can take the shape of neutral administrative disclosures, monitoring, diffusion, etc. Taxation can be delivered through fees, user charges, exemptions, etc.

Each of these instruments' shapes carries specific potential effects that cannot be measured alone because they should be considered in relation to the other tools that compose the actual set of adopted policy instruments. Understanding the distinct shapes that various types of substantive policy instruments can take when delivered is essential to grasping how governments change the instrumental side of governance arrangements over time[2]. There are three dimensions in which this distinction is helpful.

The first dimension is descriptive. By focusing on the different shapes of policy instruments, the usual description can acquire a more detailed reconstruction of shifts in governance in terms of more or less market, more or less hierarchy, etc. The second dimension is analytical. Due to the focus on the basic shapes through which substantial instruments are delivered, the concept of policy mix can be operationalized in a very effective and realistic way, and thus the eventual relevant differences can be assessed in terms of policy setting. The third dimension is clearly explanatory. If policy performance is assumed to be based on the adopted set of policy instruments, a detailed operationalization of the substantial instruments should uncover which combination of instruments actually works.

This perspective seriously considers the suggestion of those scholars who have observed how the actual set of adopted policy instruments are the consequence of a diachronic accumulation. Thus, we need to analyse the full package instead of a single type of policy instrument (Hacker 2004; Pierson 2004; Huitema and Meijerink2009; Tosun2013; Schaffrin *et al*. 2014).

[2] Here, we follow the suggestion of Ingram and Schneider (1990: 522, n.5) conceptually and in the operationalization presented below. Thy stated that 'most tools can be disaggregated into relatively small units and each unit then scored in terms of all behavioral dimensions of interest to the investigator. Even the smallest units, such as a single statement, may score "high" on more than one behavioral dimension. The units and their scores can then be re-assembled to produce a multi-variate characterization of the original policy tool'.

Therefore, from a theoretical perspective, the decision to focus on the instrumental side of governance arrangements to measure their policy performance encourages an explanation using combinatory logic: when the expected effect, namely, policy performance, is assumed to depend on the combination of multiple conditions (i.e., specific settings of policy instruments), then a complex causality principle is at work. Precisely for this reason, we decided to turn to QCA[3]. In fact, the main difference between QCA and other more quantitative research methods lies in the idea of causality underpinning the approach (Ragin 2008; Rihoux and Ragin 2009; Schneider and Wagemann 2012). While methods such as statistical analysis tend to imply mono-causality and focus on the estimate of each independent variable's separate effect on the variation of the dependent variable, the idea of causality in QCA is fundamentally characterized by equi-finality[4], conjunctural causation[5], and asymmetry[6]. Therefore, unlike quantitative methods, QCA aims at unravelling multi-causal rather than mono-causal explanations, focuses on combinations of conditions rather than on single variables, and does not assume that a unique solution (or equation) accounts for both the occurrence and non-occurrence of a particular outcome. Since we are interested in verifying which instrument combinations led to improvements in HESs' performance, QCA appears to be particularly useful[7].

[3]QCA represents a relatively new research approach (Ragin 1987; 2000; 2008; Rihoux and Ragin 2009; Schneider and Wagemann 2012); in recent years, QCA has drawn increasing attention within the social sciences (Wagemann and Schneider 2010), and some scholars consider QCA to already be a 'mainstream method' in political and sociological research (Rihoux *et al.* 2013).

[4] The idea of causation in QCA is equi-final in the sense that more than one causal pattern can lead to the outcome (Ragin 1987).

[5] The idea of causation in QCA is characterized by conjunctural causation in the sense that specific combinations of different conditions lead to the outcome (Rihoux and Ragin 2009).

[6] The idea of causation in QCA is asymmetric in the sense that there is no particular relationship between causal patterns leading to the presence of the outcome and the absence of the outcome. Conditions explaining the presence of the outcome are silent with respect to the absence of the outcome, and vice versa (Schneider and Wagemann 2012).

[7]There are different versions of QCA. While the older variant (Ragin 1987) requires a dichotomization of the variables and is based on Boolean algebra, the more recent variant (Ragin 2000; 2008) also allows for values between 0 and 1. These so-called 'fuzzy values' describe the degree of membership of a given case in a particular set (Schneider and Wagemann 2006, 752). The older variant is currently known as crisp-setQCA (csQCA); the more recent variant – which was proposed by Ragin himself as an answer to the criticisms csQCA had received – is currently known as fuzzy-setQCA (fsQCA). Moreover, there are two further versions of QCA: multi-valueQCA (mvQCA) and timeQCA (tQCA) (Schneider and Wagemann 2012). However, these versions are far less diffused than csQCA and fsQCA (Rihoux *et al.* 2013; Marx *et al.* 2014). Accordingly, in this paper, we turn to fsQCA in order to empirically test our theoretical arguments.

## 3. Research design

### *3.1 Case selection and timespan*

As mentioned above, the paper is based on a specific dataset policy tools used in 11 Western European countries (Denmark, England, Finland, France, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Sweden) over the last 25 years.

The country selection was intended encompass the complete population. At first, we intended to cover all the pre-2004 enlargement EU countries. Unlike Eastern European countries, Western Europe has been exposed to the same EU dynamics with respect to the modernization of governance in HE as well as other common goals regarding research and labour policies. However, we were forced to exclude Belgium from our analysis because of the strong differences between Flemings and Walloons (which also affected HESs). For similar reasons, we excluded Germany because of the federalist reform (in 2006) that transferred the competence of HE to the Landers, making such making Germany difficult to compare with the other countries under consideration. Finally, the impossibility of completing the extended coding in due time forced us to also temporally exclude Spain and Austria, but these two countries will be included in the successive versions of the paper. However, we included a non-EU country, Norway, as it presents many similarities with other Scandinavian countries in terms of HE reforms. Ultimately, the 11 countries considered allow to cover all the historically rooted types of university governance (Clark 1983; Braun and Merrien 1999; Huisman 2009; Paradeise *et al*. 2009; Shattock 2014).

In all the selected countries, the HE systems have undergone structural changes over the last three decades. Accordingly, we decided to start our analysis approximately at the end of the 1980sto encompass all the major changes that involved HE systems over the last 25 years. Some of these reforms were implemented during the 1990s, while many others occurred or developed during the new millennium as a consequence of the Bologna Process. This time span is justified by the need to consider a period that is characterized by huge legislative transformations in all the considered countries. Obviously, each country presents its own reform 'starting point' in the field, which means that some of countries had already produced relevant legislation by the beginning of the 1990s, while others started much later.

*3.2 Operationalization*

As already explained in the theoretical section, we assume that differences in HESs' performances follow differences in the combinations of the adopted policy tools. The operationalization of the outcome (namely, the system's performance) was quite easy. Among the various possible indicators of performance – such as access, academic recruitment and career, and third mission – we decided to focus only on teaching and research, which still ultimately represent the main tasks of every HE institute.

The most common indicator of teaching performance is the percentage (%) of people holding a tertiary degree. Accordingly, we decided to look at the percentage of adult people aged 30-34 holding a tertiary degree in each country. These data can be easily downloaded from the archives of international organizations or study centres, such as OECD or Eurostat[8].

In terms of the research dimension, the literature suggests indicators such as the percentage of citations according to ISI, the ratio between citations and researchers, the percentage of EU funding obtained per country, and the percentage of external research funding obtained with respect to the public funding (Meek and Van der Lee 2005; Blockmans *et al*. 2014; De Rijcke *et al*. 2016). We focused on academic rankings, which are based on some of the indicators listed above. As we know, rankings are provided by numerous institutions, research centres, and (academic) journals. However, all rankings present some weaknesses in their methodological approach. Additionally, timing seems to be the biggest problem for our goals. In fact, it is hard to find reliable rankings published before the last decade[9]. Relatively older rankings have been provided by QS World University Rankings (2004) and Academic Ranking of World Universities (ARWU 2003), which are regarded as the most influential and widely observed university measures together with the Times Higher Education Supplement's ranking.

The QS World University Rankings is the only international ranking to have received approval from the International Ranking Expert Group (IREG). However, it is criticized for allocating undue weight to subjective indicators and for having results that fluctuate substantially. By contrast, ARWU – an annual publication originally promoted by the Shanghai Jiao Tong University – is praised for its objective methodology but draws some condemnation for focusing narrowly on raw research power, and undermining the

[8] Another widely used indicator of performance is the % of the population aged 25-64 holding a tertiary degree. However, we opted for the increase in the percentage of 30-34-year-old adults attaining a tertiary degree between 1995 and 2015 as the best indicator for teaching performance because this indicator is more (rapidly) sensitive to policy change. This indicator's greater sensitivity is due to its focus on a small cohort of people rather than considering all adults from 25 to 64 years old.

[9] Times Higher Education World University Rankings began publication in 2010, and CWTS Leiden Ranking began in 2007.

humanities and quality of instruction. After due consideration, we finally opted for ARWU. Specifically, we decided to rely on the ranking that presented the top 300 HE institutions in the world because it represents the intermediate level and helps us to better discriminate between the different countries under consideration (in terms of the number of HE institutions included in the ranking) compared to all the other rankings ranging from the top20 to the top500.

The operationalization of the conditions of the next QCA, namely, the policy instruments, required greater theoretical reflection (see section 2.2) and greater effort in the data gathering, as we will explain below. More precisely, we have decided to operationalize the four families of substantial policy tools – regulation, taxation, expenditure, and information– while considering a long list of items (43 in total) that are presented in Appendix A. In this way, we have tried to capture all the possible shapes that substantial policy instruments can take in HE policy. We also avoided constructing categories that were too exclusive, which would have made the data collected in different countries difficult to compare.

### *3.3 Data collection, dataset construction and coding*

This paper represents the first (and rather preliminary) result of a tremendous research effort. By following the lines of our theoretical framework – which focused on the different combinations of the existing set of adopted policy combinations – we collected, analysed and coded all pieces of legislation regarding HE in all eleven countries under analysis from late1980s onwards. Every act, every decree, and every formal disposition entered our dataset; hundreds of official documents and thousands of pages of legislation have been carefully scrutinized and hand-coded while seeking both substantial and procedural policy instruments. The coding procedure developed in two steps: first, we broke down every piece of legislation into its main issues; second, we attributed each of those issues to one of the categories in which we classified the policy instrument repertoire in HE (see Appendix A).

In the first step, the research strategy has been twofold. With respect to Italy, France and both English-speaking countries – England and Ireland – the analysis has been conducted 'in house', meaning that the three authors of this paper are responsible for entering the Italian, French, English and Irish pieces of legislation into the dataset[10]. Linguistic barriers rendered similar direct coding impossible for the other seven countries – Denmark, Finland, Greece, the Netherlands, Norway, Portugal and Spain. Therefore,

[10] Regarding this, both the 'Eurydice dataset' and the various OECD 'Education at a glance' reports that have been published yearly since 1996 have represented a natural starting point.

we contacted a highly reputable country expert for each case in order to have a perfectly comparable list of legislative provisions regarding HE for those countries as well.

The attribution of all policy instruments to the appropriate category was again conducted by the authors. This second step of the coding procedure was developed as follows: first, each issue of each legislative provision in each country was coded separately by each author; second, contradictory cases – i.e., policy instruments placed in different categories by two or more coders (approximately 15% of the whole sample) – were solved jointly in a subsequent stage.

## 4. Descriptive statistics

We present some descriptive statistic here to offer a general picture of how the examined countries have intervened in HE. Table 1 shows the distribution of the countries' choices among the four families of policy instruments. As expected, regulation has been used extensively but with significant differences among the countries. Expenditure emerges as the second most preferred type of instrument; it is consistently adopted by those countries that have a relatively lower rate of regulation. The countries with the largest variety of instruments adopted are England and, surprisingly, Portugal and Finland.

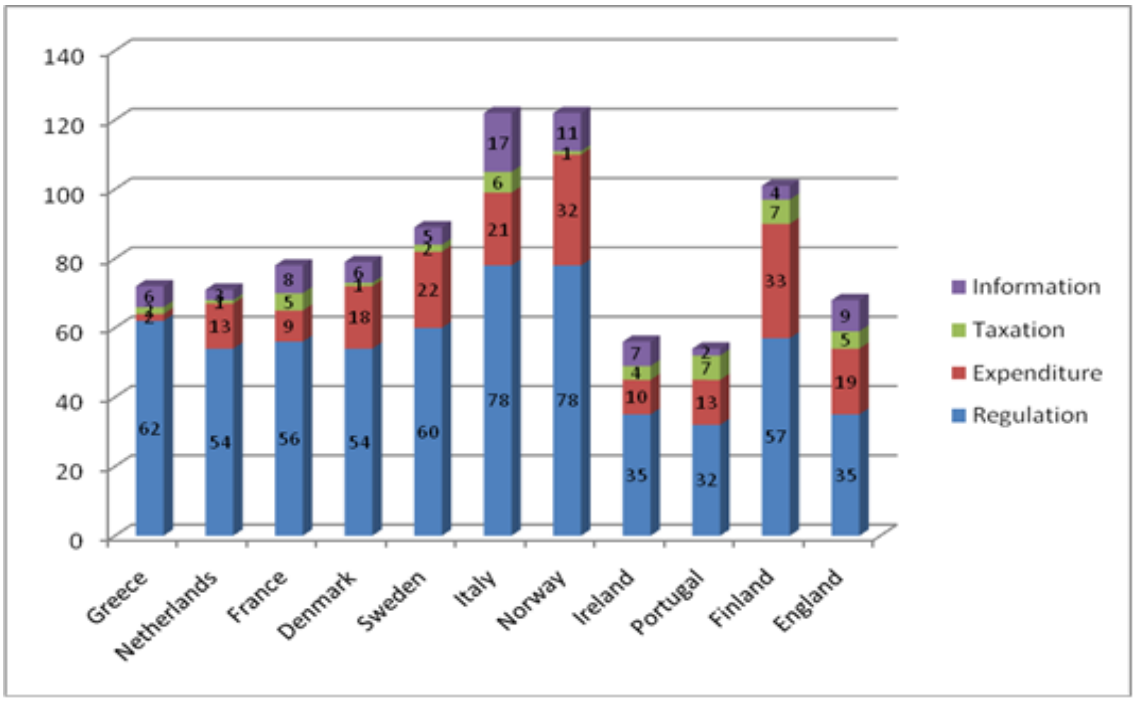
Table1 *Shapes of substantial instruments adopted in higher education legislation (1990-2014)*

| **Country** | **Reg. N** | **Reg. %** | **Exp. N** | **Exp. %** | **Tax. N** | **Tax. %** | **Info. N** | **Info. %** |
|---|---|---|---|---|---|---|---|---|
| Greece | 62 | 86.1 | 2 | 2.8 | 2 | 2.8 | 6 | 8.3 |
| Netherlands | 53 | 75.7 | 13 | 18.6 | 1 | 1.4 | 3 | 4.3 |
| France | 56 | 71.8 | 9 | 11.5 | 5 | 6.4 | 8 | 10.3 |
| Denmark | 54 | 68.3 | 18 | 22.8 | 1 | 1.3 | 6 | 7.6 |
| Sweden | 60 | 67.4 | 22 | 24.7 | 2 | 2.2 | 5 | 5.7 |
| Italy | 78 | 63.9 | 21 | 17.2 | 6 | 4.9 | 17 | 14 |
| Norway | 78 | 63.9 | 32 | 26.2 | 1 | 0.8 | 11 | 9.1 |
| Ireland | 35 | 62.5 | 10 | 17.9 | 4 | 7.1 | 7 | 12.5 |
| Portugal | 32 | 59.2 | 13 | 24.1 | 7 | 13 | 2 | 3.7 |
| Finland | 57 | 56.4 | 33 | 32.7 | 7 | 6.9 | 4 | 4 |
| England | 35 | 51.5 | 19 | 27.9 | 5 | 7.4 | 9 | 13.2 |
| **Mean** | **54.6** | **66.1** | **17.5** | **20.6** | **3.7** | **4.9** | **7.1** | **8.4** |

Figure 1 presents an assessment of the total number of decisions taken in the analysed period as well as the different intensities of the interventions. Italy and Norway have intervened in a very consistent way in redesigning HE governance, while Ireland

and Portugal appear to have been less intrusive. Examining the most preferred shapes for each family of substantial instruments is also interesting.

Figure 1 *Policy mixes in higher education legislation*




With regard to regulation, Figure 2 reveals that the two most adopted instruments are concerned with presenting more opportunities to universities both in terms of the content of curricula and in terms of institutional governance (and this seems quite coherent with the common template pursued by all the examined countries: the 'steering at a distance' governance model)

Figure 2. *An eye on regulation: which instruments are utilized the most (%)?*

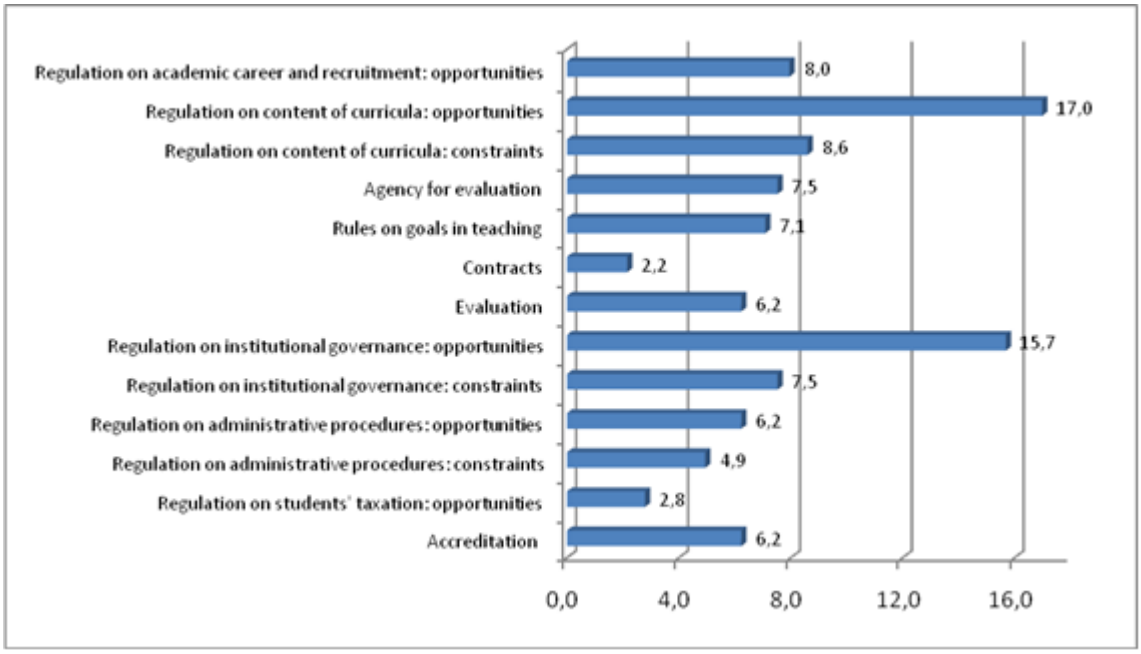
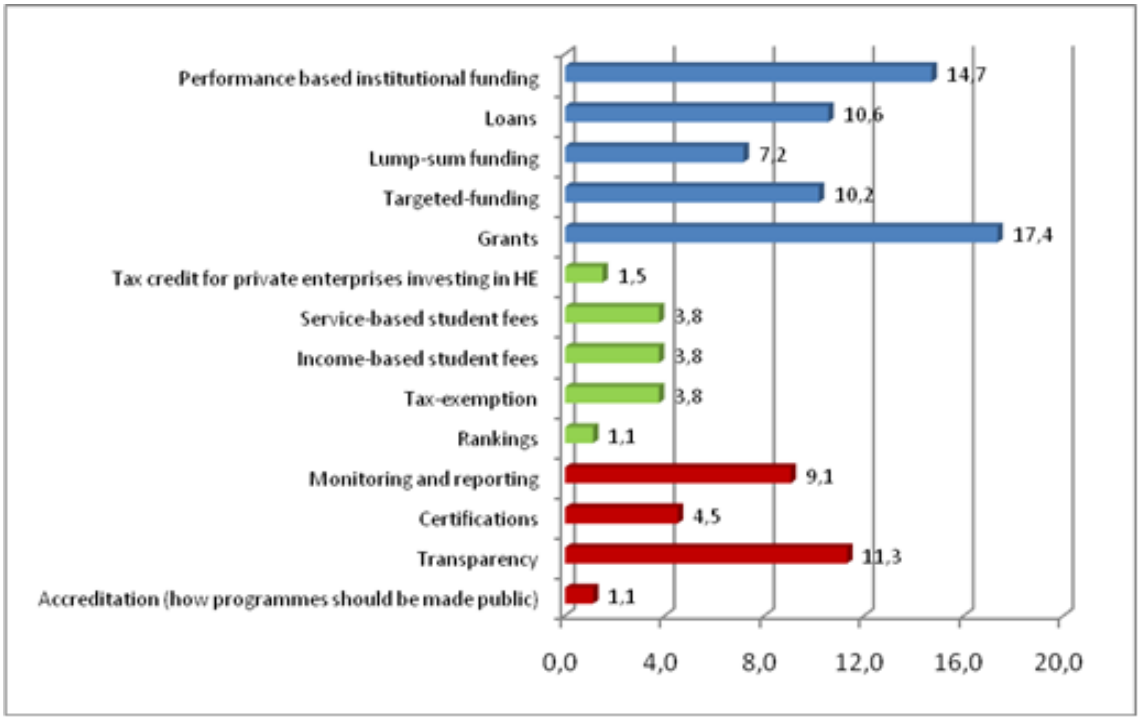



Figure 3 indicates which shapes are the most used from the other three families of substantial instruments. Apparently, this general picture also confirms the trend towards the common template but with some relevant specifications. In fact, expenditure has been delivered not only through grants but also through performance and targeted funding and loans. Thus, the most expected expenditure instrument in a pure 'steering at a distance' governance model, lump sum allocation, has been less adopted than expected. Therefore, governments have preferred a more coercive way to allocate public funding (together with an emerging attitude favouring the inclusion of families in paying HE loans) and thus demonstrate the will to maintain a certain degree of control over the behaviour of universities.

Figure 3 *An eye on the other families of substantial policy tools (expenditure, taxation, information): which instruments are utilized the most (%)?*




Overall, the descriptive picture reveals that we need to better understand actual events in the analysed countries when those governments change their set of adopted policy instruments in HE. In fact, the different distributions of substantial instruments and their possible shapes reveal that there has been significant variance when mixing together the same substantial instruments and their shapes.

## 5. What leads to better HES performance? A (double) configurational analysis

### *5.1 From policy instruments to policy mixes: a step-by-step process*

The outcome of our effort in data collection is a huge and unprecedented dataset whose main findings have been presented in the previous section. However, performing a QCA on eleven cases with dozens of different conditions (i.e., policy instruments) would clearly have made no sense due to problems of limited diversity of the data (Schneider and Wagemann 2012); thus, we turned to theoretical reflections to hypothesize a set of possible combinations to be tested in our QCA.

First, we started with eight potential configurations of conditions – six oriented towards teaching outcomes and two oriented towards research outcomes – that populated

the literature on HESs' performance to date. Those combinations of policy instruments mainly build on the literature that has emphasized the adoption of similar policy instruments to follow the common template of the 'steering at a distance' model. The combinations also build on the empirical evidence and contradictions emerging from those variable-oriented studies that focused on the determinants of performance in HE. Regarding the 'steering at a distance' literature, we have already emphasized the substantial literature underlying the ways governments have been changing governance in HE by improving institutional autonomy (and its dimensions, such as budgetary autonomy, degree of freedom in curricular content and autonomy in recruiting academic staff), quality assurance, accreditation, teaching and research assessments, monitoring, varieties of funding mechanisms (Gornitzka *et al*. 2005; Cheps 2006; Lazzaretti and Tavoletti 2006; Maassen and Olsen 2007; Trakman 2008; Huisman 2009, Shattock 2014, Capano, Regini, and Turri 2016).

This literature clearly addressed our choice to consider those shapes for each family of substantial policy instruments that seemed to better represent the operative dimensions of those main categories of government intervention. This choice has been reinforced by those contributions that have focused on the real effects of performance-based funding on institutional autonomy and on the degree of centralization of the governance system. This literature has produced contrasting empirical evidence and has thus suggested a strategic dimension that we should take into consideration. For example, with respect to performance and targeted funding as a cause of high graduation and research rates, the relevant studies show the weak performative capacity of these instruments. Most of these studies focused on the United States in which many states introduced performance criteria to determine the allocation of extra resources beginning at end of the 1970s (Rabovsky 2012; Tandberg and Hillman 2014; Volkwein and Talberg 2008; Rutheford and Rabovsky 2014). This evidence is quite contradictory when compared to the widespread use that the governments of our analysed countries have made of these types of instruments. Thus, we have considered all targeted expenditure tools as relevant in our combinations.

Regarding institutional autonomy, there is contradictory evidence emerging from reputable studies. For example, in their comparison between the EU and the US, Aghion and colleagues (2010) found that high institutional autonomy (and a competitive environment) positively correlates with high performance both in educational attainment and in research. By contrast, Braga and colleagues (2013) show that high institutional autonomy negatively impacts the level of educational attainment.

With respect to the level of the centralization of the governance system, Knott and Payne (2004) considered systemic centralization in US states to be high, intermediate, or low depending on the scope of the decision-making powers held by state boards. They

tested systemic centralization as a condition affecting an array of resource and productivity measures, including the size of university revenues and the number of published articles. They concluded that state governance matters and that flagship universities are penalized by centralization, which reduces their total revenues, research funding, and number of published articles. However, centralization was also found to reduce tuition revenues and – presumably – the cost of enrolment. The worst overall performance occurs under mild centralization. However, this reading contrasts with the qualitative study on five US states conducted by Richardson and Martinez (2009). They argued that universities in centralized systems might perform better than those with decentralized designs with respect to access and graduation rates.

Thus, according to the abovementioned literature, and assuming that governments have tried to pursue the 'steering at a distance' model within their own national identity, we have tried to combine these shapes in a way through which 8 different subcategories of this model have been established.

As shown in Table 2, we have first designed three possible types of coherent 'steering at a distance' policy mixes to address teaching. Two of them (***a, b***) are differentiated based on the type of shapes with respect to expenditure, information and taxation while maintaining the same three regulation shapes typical of 'steering at a distance'. The third one, ***c***, is characterized by four regulative dimensions (by including accreditation and regulation in goals) together with two expenditure dimensions.

Type ***d*** is composed of a significant number of financial tools in order to describe a mix in which the system is addressed specifically through these types of instruments. Type ***e*** is imagined as a policy mix in which competition is the mechanism activated through the use of numerous information tools and which charges students with a significant financial responsibility in funding the system. Mix ***f*** is composed of those shapes that have scarcely been used but have been considered theoretically interesting by many scholars in addressing the way in which systemic governance works. In our research, we have selected two mixes by assuming from the evidence emerging in the comparative perspective that there has been a significant push towards more competition. Thus, the two proposed mixes are characterized by being highly competitive (mix ***g***) or moderately competitive with the government placing some significant constraints on the actions of universities (mix **h**).

We also decided to add a configuration consisting of the most used policy instruments over the course of our time span: the two most utilized regulatory instruments and the most utilized instruments in the 'expenditure', 'taxation' and 'information' categories. We then decided to test the explanatory power of this configuration with regard to both teaching and research (mixes ***i*** and ***ii***). This in turn led our first set of policy mixes in ten configurations to be empirically tested, as Table 2 suggests.

Table 2 *Policy mixes: first step*

| *Policy mix* | *Policy mix: details* | *Outcome* |
|---|---|---|
| a. Steering at the distance – Teaching | R7 (more opportunities on curricula) + R9 (more opportunities on academic recruitment) + R17 (more institutional autonomy) + E7 (performance-based institutional funding) + T4 (service-based student fees) + I4 (monitoring and reporting) | Tertiary 30-34 yo. |
| b. Steering at the distance 2 – Teaching | R7 (more opportunities on curricula) + R9 (more opportunities on academic recruitment)+ R17 (more institutional autonomy) + E4 (Lump-sum funding) + T2 (Income-based student fees) + I2 (Transparency) | Tertiary 30-34 yo. |
| c. Steering at the distance 3 – Teaching | R1 (Accreditation - whereas the regulatory aspect is prevalent) + R7 (more opportunities on curricula) + R17 (more institutional autonomy) + R22 (Rules on goals in teaching) + E3 (Targeted-funding) + E7 (Performance based institutional funding) | Tertiary 30-34 yo. |
| d. Financial steering – Teaching | R1 (Accreditation - whereas the regulatory aspect is prevalent) + E3 (Targeted-funding) + E6 (Loans) + E7 (Performance based institutional funding) + I2 (Transparency) + I4 (Monitoring and reporting) | Tertiary 30-34 yo. |
| e. Competitive steering – Teaching | R9 (Regulation on academic career and recruitment: more opportunities) + E6 (Loans) + I2 (Transparency) + I3 (Certifications) + I4 (Monitoring and reporting) + I5 (Rankings) | Tertiary 30-34 yo. |
| f. Residual steering – Teaching | R1 (Accreditation - whereas the regulatory aspect is prevalent) + R9 (Regulation on academic career and recruitment: more opportunities) + R21 (Contracts) + R22 (Rules on goals in teaching) + E3 (Targeted-funding) + I1 (Accreditation - whereas the informative aspect is prevalent: how programs with accreditation should be made public) | Tertiary 30-34 yo. |
| g. Highly Competitive steering – Research | R4 (Agency for evaluation) + R9 (Regulation on academic career and recruitment: more opportunities) + R13 (Regulation on administrative procedures: more opportunities) + R17 (more institutional autonomy) + E7 (Performance based institutional funding) + T5 (Tax credit for private enterprises investing in HE) | ARWU-top300 |
| h. Moderately Competitive steering – Research | R2 (Evaluation) + R9 (Regulation on academic career and recruitment: more opportunities) + R12 (Regulation on administrative procedures: more constraints) + R16 (Regulation on institutional governance: more constraints) + E7 (Performance based institutional funding) + I4 (Monitoring and reporting) | ARWU-top300 |
| i. Most used instrumental shapes – Teaching | R7 (more opportunities on curricula) + R17 (more institutional autonomy) + E1 (Grants) + T1 (Tax-exemption) + I2 (Transparency) | Tertiary 30-34 yo. |
| ii. Most used instrumental shapes – Research | R7 (Regulation on content of curricula: more opportunities) + R17 (Regulation on institutional governance: more opportunities) + E1 (Grants) + T1 (Tax-exemption) + I2 (Transparency) | ARWU-top300 |

As previously stated, in this paper we turn to fsQCA to unravel policy combinations that led to improved performances in HE. One of the first steps in each fsQCA is the calibration of sets (both conditions and outcomes) (Ragin 2008; Schneider and Wagemann 2012). In this basic process – which should be as transparent as possible and which should be discussed in detail (Schneider and Wagemann 2010, 403) – it is particularly crucial to specify the qualitative anchors for full membership (1), full non-membership (0) and the point of maximum ambiguity (0.5)[11]. All instruments entering the ten configurations have been thus calibrated, and the choices made are summarized in Appendix B.

Once the instruments were calibrated, we ran many different QCAs to test which configurations represented an acceptable compromise between theoretical expectations and empirical reality. However, given that a QCA of six conditions on a sample of only eleven cases gave rise to many logical remainders, we also tested a few combinations of four policy instruments rather than six. Again, those combinations were built on the basis of the above literature and on the basis of some evidence raised by the test of the first list of ten configurations.

This all led to our second step, which consisted of testing the four configurations of conditions that we considered more theoretically convincing and empirically robust: three for 'teaching' and one for 'research' (see Table 3). Regarding teaching, the mix with 6 instruments, 'steering at a distance 1' (mix ***a***), examined an equilibrated combination of institutional regulation (in terms of more opportunities), regulation for evaluation (accreditation and presence of an agency of evaluation), and targeted funding (since the test revealed that the lump sum budget was shown to be substantially irrelevant whenever considered). The two mixes composed of 4 instruments are characterized by different dimensions of regulations (although they are similar in terms of expenditure). The difference seems to be limited to the more autonomous behaviour ensured to institutions that characterizes the 'steering at a distance 3' mix.

We have also selected a mix based on evaluation, as this dimension emerged in the first test as apparently the most relevant one.

[11] We made use of the direct method of calibration (Ragin 2008, 85). Once qualitative anchors have been chosen, the QCA software applies a logarithmic function and attributes fuzzy values to the remaining cases. By contrast, the indirect method requires cases to be initially grouped into set membership scores (for example: 0.8, 0.6, 0.4, 0.2). Using the fractional logit model, these preliminary set membership scores are then regressed on the raw data. The predicted values of this model are then used as the fuzzy-set membership scores (Schneider and Wagemann 2012).

Table 3. *Policy mixes: second step*

| *Policy mix* | *Policy mix: details* | *Outcome* |
|---|---|---|
| Steering at the distance **1** (6 conditions) – Teaching | R1 (Accreditation - whereas the regulatory aspect is prevalent) + R4 (Agency for evaluation) + R7 (Regulation on content of curricula: more opportunities) + R17 (Regulation on institutional governance: more opportunities) + E3 (Targeted-funding) + I4 (Monitoring and reporting) | Tertiary 30-34 yo. |
| Steering at the distance **2** (4 conditions) – Teaching | R1 (Accreditation - whereas the regulatory aspect is prevalent) + R17 (Regulation on institutional governance: more opportunities) + E3 (Targeted-funding) + E7 (Performance based institutional funding) | Tertiary 30-34 yo. |
| Steering at the distance **3** (4 conditions) – Teaching | R7 (Regulation on content of curricula: more opportunities) + R15 (more opportunities on students' admissions) + E3 (Targeted-funding) + T2 (Income-based student fees) | Tertiary 30-34 yo. |
| Evaluative Steering (4 conditions) – Research | R2 (Evaluation) + R17 (Regulation on institutional governance: more opportunities) + E3 (Targeted-funding) + E7 (Performance based institutional funding) | ARWU-top300 |

Again, we carefully confronted theories and empirical results to ascertain which configurations best resemble reality. To accomplish this, we particularly focused on the theoretical plausibility of identified solutions, on the (groups of) countries each solution accounted for, and on the values of consistency and the coverage of different final solution. At the end of this step-by-step process through which we have been able to drastically reduce the complexity of the descriptive picture presented in section 4, we arrived at two configurations of conditions – one explaining teaching performance and one explaining research performance – that fit very well with both theory and empirical reality. In other words, among the five possible aims of using QCA listed by Berg-Schlosser and colleagues[12] (Berg-Schlosser *et al*. 2008), at this stage, we focused on data summary in order to unravel patterns of variation among countries. In the following two subsections, we develop a QCA analysis for both teaching and research, presenting the two configurations of policy instruments that passed the (very long and demanding) selection conducted thus far.

[12] These five aims are as follows: *i)* to summarize data, *ii)* to check the coherence of the data with claims of subset relations, *iii)* to test existing theories and hypotheses, *iv)* to quickly overview the basic assumptions of the analysis, and *v)* to develop new theoretical arguments.

*5.2 Policy mix and teaching performance*

As mentioned, all the conditions (i.e., policy instruments) included in all the different configurations we tested have been calibrated. However, we also had to make choices with respect to outcomes. As previously discussed, we operationalized teaching performance by turning to the proportion of 30-34-year-old adults attaining a tertiary degree. This starting point is represented by Eurostat data, and the performance of different countries depends on the degree to which the percentage of adults with a tertiary degree increased between 1995 and 2016. However, using this method, we would not have considered the obvious factor that improving results is easier when starting from a very low value than when already at the top. Therefore, we slightly modified the data by differentiating the countries in three categories: countries below the mean of tertiary degree attainment (30-34 years of age) in 1995, countries over the mean but under one standard deviation over the mean, and countries over one standard deviation over the mean. Countries in the last category received an increase in their performance equal to +30%, whereas countries in the intermediate (over the mean but under one s.d. over the mean) category received an increase in their performance between +10% and +30% depending on the degree to which they were over the mean. Finally, countries under the mean did not receive any increase. Because of this complex process, Table 4 shows both the original Eurostat data as well as how countries can be ranked with respect to our measure of teaching performance from 1995 to 2016.

Table 4 *Teaching performance: comparison between 1995 and 2016*

| *Country* | *Eurostat 1995* | *Eurostat 2016* | *2016-1995* | *Teaching performance(2016-1995)* |
|---|---|---|---|---|
| Ireland | 25.1 | 52.9 | +27.8 p.p. | +31.8 percentage points |
| Sweden | 29.4 | 51.0 | +21.6 p.p. | +28.1 p.p. |
| Norway | 28.9 | 50.1 | +21.2 p.p. | +27.6 p.p. |
| England | 24.1 | 48.1 | +24.0 p.p. | +26.4 p.p. |
| Finland | 24.6 | 46.1 | +21.5 p.p. | +24.2 p.p. |
| Netherlands | 24.2 | 45.7 | +21.5 p.p. | +23.8 p.p. |
| Denmark | 29.6 | 47.7 | +18.1 p.p. | +23.5 p.p. |
| Greece | 20.7 | 42.7 | +22.0 p.p. | +22.0 p.p. |
| France | 22.2 | 43.6 | +21.4 p.p. | +21.8 p.p. |
| Portugal | 13.2 | 34.6 | +21.4 p.p. | +21.4 p.p. |
| Italy | 8.6 | 26.2 | +17.6 p.p. | +17.6 p.p. |
| *Mean* | *21.8* | *26.6* | *+4.8 p.p.* | / |

That said, all the conditions (i.e., policy instruments) and outcomes included in all different configurations we tested have been calibrated, and the different qualitative anchors (full membership, full non-membership and maximum ambiguity) are listed in Appendix B. However, since what follows represents the best configuration of conditions we chose for explaining teaching outputs, it is necessary to justify in detail all the choices made, which are shown in Table 5 below.

Table 5 *Calibration of sets: conditions and the outcome (teaching)*

| Set | Full membership (1) | Point of maximum ambiguity (0.5) | Full non-membership (0) |
|---|---|---|---|
| More opportunities on curricula (R7) | 12 | 6.5 | 1 |
| More opportunities on admissions (R15) | 3.5 | 1.5 | 0 |
| Targeted-funding (E3) | 3 | 1.5 | 0 |
| Income-based student fees (T2) | 2 | 0.5 | 0 |
| Tertiary 30-34 yo. | +29 p.p. | +23 p.p. | +18 p.p. |

The calibration of conditions and the outcome of assessing teaching performance needs to be justified at length, especially where the point of maximum ambiguity is concerned. Regarding R7 (more opportunities to universities on curricula), since many countries have repeatedly intervened with respect to this instrument of regulation, the point of maximum ambiguity has been determined in a way that rewards those countries that have intervened seven or more times. The same reasoning is used in calibrating the condition R15. In this case, some countries never intervened, while other countries have consistently redesigned the content of this instrument of regulation through which more autonomy is given to universities in selecting students. Thus, the threshold for maximum ambiguity can be discerned correctly. The instrumental shapes E3 and T2 have also been used a few times by certain countries. Thus, the calibration has been conducted with the intention of giving relevance to those countries that have intervened the most because this demonstrates specific attention to these sensitive and politically relevant instruments. We have based our calibration choices while taking into account the descriptive picture offered in the comparative literature on targeted and performance-based funding (Frolich, Schmidt and Roma 2010; Flannery D. & C. O'Donoghue 2011; Cheps 2015, EUA 2015), student fee systems (Heller and Callender 2013; European Commission 2016), and university autonomy with respect to curricula and admissions (Russo 2013; EUA 2017)

Regarding the calibration of the outcome, our first concern was to reward countries that already had high performances in 1995. With this in mind, our calibration places only one country – Italy – in full non-membership; three other countries are below the

ambiguity point, and seven are above it; only Ireland is included in full membership. In this way, we believe that we have properly appreciated national performance with respect to the considered outcome.

Once the sets have been calibrated, the second step of each QCA – both crisp-set and fuzzy-set – consists of the analysis of necessity relations, which should always be conducted before the analysis of sufficiency conditions (Schneider and Wagemann 2010, 404). With respect to necessity relations, as Table 6 demonstrates, no condition (or non-occurrence of a condition, which is indicated with the tilde ~) is necessary for the outcome[13].

Table 6 *Analysis of necessary conditions. Outcome: variation (1995-2015) in the % of 30-34 yo. adults attaining a tertiary degree*

| *Condition* | *Consistency* | *Coverage* |
|---|---|---|
| R7 | 0.658424 | 0.657407 |
| ~R7 | 0.539413 | 0.772124 |
| R15 | 0.727975 | 0.710407 |
| ~R15 | 0.431221 | 0.638444 |
| E3 | 0.709428 | 0.661383 |
| ~E3 | 0.409583 | 0.652709 |
| T2 | 0.452859 | 0.531760 |
| ~T2 | 0.673879 | 0.794171 |

Following the analysis of necessity, the empirical test of sufficiency set-relations between (combinations of) conditions and the outcome is conducted through the 'truth table'. The process proceeds as follows: *i)* the algorithm predicts the conversion of the data matrix into the abovementioned truth table; *ii)* single truth table rows are assessed on the basis of their consistency scores as to whether or not they can count as sufficient conditions for the result; and *iii)* if they can be considered sufficient conditions, they are included in the so-called 'Boolean minimization process'; otherwise, they are not. On this point, see Table 7.

[13] All the consistency thresholds are lower than 0.9, which is the value over which empirical evidence is considered to support the claim that a condition is necessary for the outcome (Schneider and Wagemann 2012, 278).

Table 7 *Truth table (teaching performance)*

| *R7* | *R15* | *E3* | *T2* | *Number* | *outcome* | *Raw consist.* | *PRI consist.* | *SYM consist.* |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 3 (27%) | 1 | 0.84 | 0.71 | 0.99 |
| 1 | 1 | 0 | 0 | 2 (45%) | 1 | 0.94 | 0.89 | 0.99 |
| 0 | 0 | 1 | 1 | 2 (63%) | 1 | 0.93 | 0.89 | 0.89 |
| 1 | 1 | 1 | 1 | 1 (72%) | 0 | 0.76 | 0.07 | 0.07 |
| 1 | 1 | 0 | 1 | 1 (81%) | 0 | 0.68 | 0.06 | 0.06 |
| 1 | 0 | 1 | 1 | 1 (90%) | 0 | 0.46 | 0.03 | 0.03 |
| 0 | 0 | 0 | 1 | 1 (100%) | 0 | 0.62 | 0.36 | 0.46 |
| 1 | 0 | 1 | 0 | 0 (100%) | 0 | / | / | / |
| 1 | 0 | 0 | 1 | 0 (100%) | 0 | / | / | / |
| 1 | 0 | 0 | 0 | 0 (100%) | 0 | / | / | / |
| 0 | 1 | 1 | 1 | 0 (100%) | 0 | / | / | / |
| 0 | 1 | 1 | 0 | 0 (100%) | 0 | / | / | / |
| 0 | 1 | 0 | 1 | 0 (100%) | 0 | / | / | / |
| 0 | 1 | 0 | 0 | 0 (100%) | 0 | / | / | / |
| 0 | 0 | 1 | 0 | 0 (100%) | 0 | / | / | / |
| 0 | 0 | 0 | 0 | 0 (100%) | 0 | / | / | / |

First, it should be noted that there are nine logical remainders. This means that not all combinations of conditions are characterized by at least one empirical case and that problems of limited diversity may arise. Therefore, solution formulas – which are complex, parsimonious and intermediate – are not interchangeable[14] and assumptions on that same logical remainder should be made. In this case, one generally ought to consider the intermediate solution in order to lower the risk of drawing the wrong conclusions about the automatic counterfactuals used in the parsimonious and complex solution (Jano 2016, 15).

Second, as Schneider and Wagemann (2012) suggest, only combinations of conditions that present a raw consistency higher than 0.75 should contribute to the minimization of the truth table algorithm. However, the 0.75 threshold should not be applied mechanically; for example, in this case, we decided to exclude the combination of conditions showing a raw consistency of 0.76 because it is very close to the threshold

[14] Indeed, in QCA, solution formulas differ on the basis of assumptions on logical remainders. The complex solution assumes that logical remainders would not produce the outcome. The parsimonious solution treats remainders as 'don't care', stimulating outcome values such that parsimony is obtained. Instead of these outcomes, the intermediate solution evaluates the plausibility of remainders in accordance with the researcher's simplifying assumptions based on theoretical or substantive empirical knowledge.

and is rather far from the other combinations with higher values. Thus, the intermediate solution formula is as follows:

Intermediate solution = R7*R15*~T2 + E3*T2*~R7*~R15

Table 8 *Solution formulas, consistency, coverage and cases covered (teaching)*

| | *Raw coverage* | *Unique coverage* | *Consistency* | *Cases covered* |
|---|---|---|---|---|
| R7*R15*~T2 | 0.602782 | 0.511592 | 0.851528 | Denmark (0.95; 0.56); Sweden (0.95; 0.93); Finland (0.80; 0.63); Netherlands (0.69; 0.60); Norway (0.69; 0.91) |
| E3*T2*~R7*~R15 | 0.312210 | 0.221020 | 0.926605 | England (0.73; 0.85); Ireland (0.73; 0.99) |

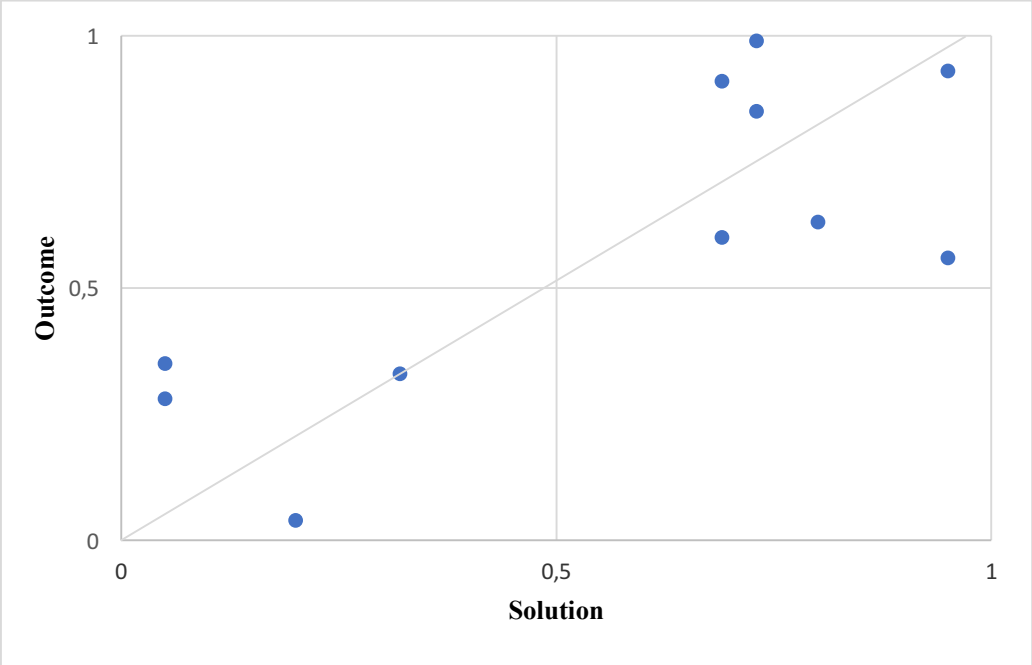
- Solution coverage: 0.823802
- Solution consistency: 0.865260

Theoretically, the solution above means that there are five countries whose performance improvement in teaching has been due to the joint occurrence of more opportunities in choosing students and the content of the offered degree course in the absence of any specific reference to income-based fees. This seems quite interesting because these countries diverge with respect to tuition fee policies (consistently present in the Netherlands in the comparative perspective while absent in the other countries). In all five countries, governments have introduced significant changes in institutional governance arrangements. This raises the question of the effectiveness of the conditions R7 and R15. Interestingly, the other combination – E3 with T2 – includes England and Ireland, countries in which institutional autonomy has historically always been high. In these countries, the government can also evidently expect a positive reaction when allocating targeted funding. Additionally, income-based fees (often based on the right to specific mean-tested grants) matter in these countries. It is interesting that these two conditions work together in the absence of the other two conditions, the regulative ones. This is not unexpected due to the historical autonomy that universities have enjoyed in these two countries with respect to admission and curricular content.

Overall, the consistency value of the intermediate solution is very good (0.87), and the coverage of the solution formula is more than satisfactory (0.82). There are no contradictory cases (in the lower-right quadrant) or 'deviant cases for coverage' (Schneider and Rohlfing 2013, 585) in the upper-left quadrant. By contrast, three cases (Norway, England and Ireland) – which are above the diagonal in the upper-right corner – are 'typical cases', whereas four other cases (Denmark, Sweden, Finland and the

Netherlands) are 'deviant cases in degree of consistency'. Finally, four cases (France, Greece, Italy and Portugal) in the lower-left quadrant – which are neither good examples of the solution formulas nor of the outcome – do not merit particular attention.

Figure 4 *Final XY Plot – Policy shapes (conditions) vs. teaching performance*




Overall, the configurational analysis seems to give robust support to the argument that different combinations of financial and regulatory policy instruments are relevant in increasing HESs' teaching performances. More autonomy for universities matters as does targeted funding. However, they matter in different contexts and according to different historical traditions in HE. Furthermore, the configuration of policy instruments that has been more effective in pushing towards an increase of 30-34-year-old tertiary degree-holders works in countries characterized by accountability and high institutional autonomy.

This is the case in England and Ireland where institutional autonomy is a historically rooted trait; this is also the case for other countries where consistent changes in the institutional governance of universities have been introduced by government in the last two decades specifically to grant more institutional autonomy. This raises a relevant question concerning the effectiveness of those institutional governance reforms introduced in Portugal, France and Italy.

*5.3 Policy mix and research performance*

Just as we developed a QCA on teaching outputs, in this subsection, we focus on research by presenting the configuration of conditions (i.e., policy instruments) we identified as the best for explaining how and why some countries increased the number of their universities listed in the ARWU Top300 Ranking, whereas others did not.

However, before justifying all the choices made for calibration, data on countries' research performance must be presented as our 'outcome' in this QCA. As mentioned above, we operationalized research performance through the variation in the number of universities listed in the ARWU Top300 Ranking between 2003 – the first year in which the ARWU Ranking was published – and 2015. On this, see Table 9 below.

Table 9 *Research performance: comparison between 2003 and 2015*

| *Country* | *2003* | *2015* | *Difference 2015-2003* |
|---|---|---|---|
| France | 12 | 15 | 3 |
| Norway | 1 | 3 | 2 |
| England | 27 | 28 | 1 |
| Portugal | 0 | 1 | 1 |
| Denmark | 4 | 4 | 0 |
| Ireland | 1 | 1 | 0 |
| Netherlands | 10 | 10 | 0 |
| Finland | 2 | 1 | -1 |
| Greece | 1 | 0 | -1 |
| Sweden | 9 | 7 | -2 |
| Italy | 11 | 8 | -3 |
| *Mean* | *7.1* | *7.1* | *0* |

Again, the first step in each fsQCA is the calibration of both conditions and the outcome; all choices are summarized in Table 10.

Table 10 *Calibration of sets: conditions and the outcome (research performance)*

| *Set* | *Full membership (1)* | *Point of maximum ambiguity (0.5)* | *Full non-membership (0)* |
|---|---|---|---|
| Evaluation (R2) | 3.5 | 2.5 | 1.5 |
| More opportunities in institutional governance (R17) | 9 | 5.5 | 3 |
| Targeted-funding (E3) | 3 | 1.5 | 0 |
| Performance based institutional funding (E7) | 4 | 1.5 | 0 |
| Variation in the no. of universities in ARWU-top300 | 1.5 | -0.5 | -1.5 |

The calibration chosen for R2 is based on the asymmetric distribution of countries' decisions. Therefore, we have penalized those countries that have scored 0 or 1 on this instrument, while we have appreciated those countries (4) that intervened more than 4 times on it. With respect to R17, we have coded numerous interventions on this instrument of regulation, which is expected since providing more opportunities for institutional governance has been the pillar (almost a mantra) in the waves of HE reform. Therefore, we have also more substantially rewarded those countries that have intervened abundantly on this instrument (2) and penalized those that have intervened rarely (2), while we have given a positive location to those between the ambiguity point and full membership (7). E7 has been characterized by a very asymmetric distribution, so we have strongly discriminated against those without any intervention compared to those who have been quite active. For E3, the same reasoning presented above for teaching apply. We have based our calibration on the empirical evidence offered by the existing studies. With respect to performance, targeted funding, and evaluation, we have based our choices on a cross-reading of Whitley and Gläser (2007), Frolick (2011), Hicks (2012), Cheps (2015), and Jonkers and Zacharewicz (2016), who have offered deep comparative empirical material to drive the calibration in a very conscious way. In addressing institutional governance and the relationship with research performance, we have capitalized on the substantial literature on this topic (among the others: Huismans 2009; Aghion *et al*. 2010; Shattock 2014; Dobbins and Knill 2014; Capano, Regini and Turri 2016).

Again, before proceeding with the analysis of sufficiency relations, necessity relations must be examined. As Table 11 demonstrates, no condition (or non-occurrence of a condition) is necessary for the outcome:

Table 11 *Analysis of necessary conditions. Outcome: variation (2003-2015) in the no. of universities in ARWU-top300*

| *Condition* | *Consistency* | *Coverage* |
|---|---|---|
| R2 | 0.802905 | 0.663808 |
| ~R2 | 0.300830 | 0.280464 |
| R17 | 0.695021 | 0.542950 |
| ~R17 | 0.493776 | 0.492754 |
| E3 | 0.800830 | 0.556196 |
| ~E3 | 0.363071 | 0.431034 |
| E7 | 0.703320 | 0.472803 |
| ~E7 | 0.383817 | 0.483029 |

Following the analysis of necessity, the empirical test of sufficiency set-relations between (combinations of) conditions and the outcome is again conducted through the 'truth table'. On this point, see Table 12.

Table 12 *Truth table (research performance)*

| *R17* | *R2* | *E3* | *E7* | *Number* | *outcome* | *Raw consist.* | *PRI consist.* | *SYM consist.* |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 2 (18%) | 0 | 0.28 | 0.14 | 0.14 |
| 1 | 1 | 0 | 1 | 2 (36%) | 1 | 0.94 | 0.93 | 0.93 |
| 0 | 0 | 1 | 1 | 2 (54%) | 1 | 0.97 | 0.95 | 0.99 |
| 1 | 1 | 0 | 0 | 1 (63%) | 0 | 0.46 | 0.27 | 0.27 |
| 1 | 0 | 1 | 1 | 1 (72%) | 1 | 0.99 | 0.98 | 0.98 |
| 1 | 0 | 0 | 0 | 1 (81%) | 1 | 0.94 | 0.87 | 0.98 |
| 0 | 1 | 1 | 1 | 1 (90%) | 0 | 0.32 | 0.07 | 0.07 |
| 0 | 0 | 1 | 0 | 1 (100%) | 1 | 0.99 | 0.99 | 0.99 |
| 1 | 1 | 1 | 0 | 0 (100%) | 0 | / | / | / |
| 1 | 0 | 1 | 0 | 0 (100%) | 0 | / | / | / |
| 1 | 0 | 0 | 1 | 0 (100%) | 0 | / | / | / |
| 0 | 1 | 1 | 0 | 0 (100%) | 0 | / | / | / |
| 0 | 1 | 0 | 1 | 0 (100%) | 0 | / | / | / |
| 0 | 1 | 0 | 0 | 0 (100%) | 0 | / | / | / |
| 0 | 0 | 0 | 1 | 0 (100%) | 0 | / | / | / |
| 0 | 0 | 0 | 0 | 0 (100%) | 0 | / | / | / |

As with teaching performance, there are again logical remainders. Unfortunately, with only eleven cases, logical remainders are largely unavoidable. Again, solution formulas – complex, parsimonious and intermediate – are therefore not interchangeable, and once more, we decided to consider the intermediate solution in order to lower the risk of making the wrong assumptions regarding the automatic counterfactuals used in the parsimonious and complex solutions (Jano 2016).

That said, the intermediate solution originating from the process of the minimization of the truth table algorithm is as follows:

Intermediate solution = E3*~R2*~R17 + E3*E7*~R2 + R17*~R2*~E3*~E7 + R2*R17*~E3*E7

Table 13 *Solution formulas, consistency, coverage and cases covered (research)*

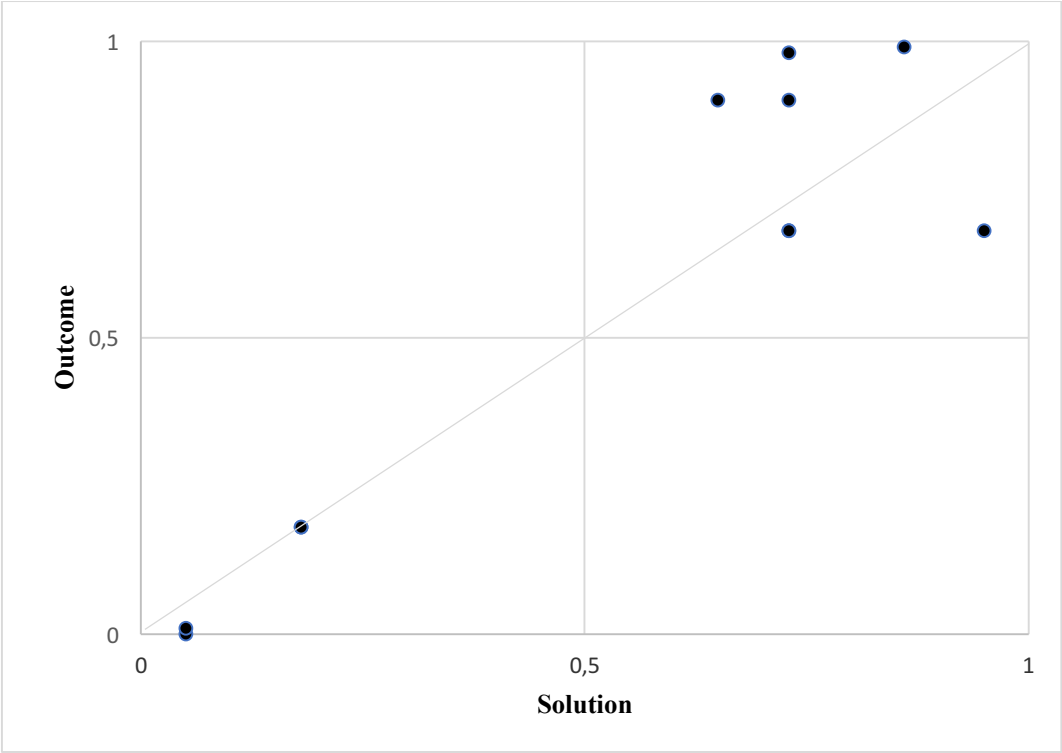
| | *Raw coverage* | *Unique coverage* | *Consistency* | *Cases covered* |
|---|---|---|---|---|
| E3*~R2*~R17 | 0.477346 | 0.108414 | 0.980066 | France (0.86; 0.99); England (0.73; 0.9); Ireland (0.73; 0.68) |
| E3*E7*~R2 | 0.415858 | 0.046926 | 0.886207 | Denmark (0.95; 0.68); England (0.73; 0.9); Ireland (0.73; 0.68) |
| R17*~R2*~E3*~E7 | 0.160194 | 0.122977 | 0.942857 | Netherlands (0.73; 0.68) |
| R2*R17*~E3*E7 | 0.239482 | 0.223301 | 0.942675 | Norway (0.73; 0.98); Portugal (0.65; 0.9) |

- Solution coverage: 0.870550
- Solution consistency: 0.921233

Overall, the consistency value of the intermediate solution is quite remarkable (0.92), and the coverage of the solution formula is very good (0.87). There are neither contradictory cases (in the lower-right quadrant) nor 'deviant cases for coverage' (Schneider and Rohlfing 2013, 585) in the upper-left quadrant. By contrast, four cases (England, France, Norway and Portugal) – which are above the diagonal in the upper-right corner – are 'typical cases', whereas three further cases (Denmark, Ireland and the Netherlands) are 'deviant cases in degree of consistency. Finally, the four cases (Finland, Greece, Italy and Sweden) in the lower-left quadrant – which are neither good examples of solution formulas nor of the outcome – do not merit particular attention.

Theoretically, the solution above is not only very effective in terms of both coverage and consistency, but it is also rather interesting because it reveals that performance in research can be driven by a small combination of factors but rely on a specific contextual situation. In sum, the configurational analysis seems to give substantial support to the argument that different combinations of financial and regulatory policy instruments are relevant to increasing HESs' research performances. Targeted funding (in the absence of other conditions) seems quite powerful in addressing HESs belonging to different traditions (like France on the one hand and England and Ireland on the other); direct choices by the government to target funding as well as performance-based funding (in absence of other two conditions) work in Denmark (but also again in England and Ireland), while the other two regulatory conditions also seem to be quite effective taken alone, as shown by the cases of the Netherlands, Portugal and Norway.

Figure 5 *Final XY Plot – Policy shapes (conditions) vs. Research performance*




These results reveal that the literature emphasizing that competition for funding and institutional autonomy are the main variables in explaining the research performance of universities is partially wrong or at least does not tell the entire story. In fact, from our

analysis, hierarchic driven instruments, such as targeted and performance funding, also appear to work well, and evaluation or autonomous institutional governance alone can stand as significant drivers of effective performance. This type of analysis provides both complexity and clarity.

## 6. Conclusions and future research

We have devoted this paper to addressing a general problem in analysing governance shifts in public policy by empirically focus on HE as an exemplary field. We have assumed that the complexity of understanding whether and how governance changes matter in terms of policy performance should be analysed through a detailed perspective that starts from the basic component of governmental actions and governance arrangements: policy instruments. We have proposed a classification of substantial instruments (expenditure, regulation, information and taxation) to grasp the complete spectrum of induced behaviour that can be addressed. Then, we have operationalized these substantial instruments according to different shapes, ways through which governments can adopt each of them. We have used this long list to code the instrument choices made in 11 European countries over the last 25 years in the governance arrangements of their HE system. Thus, thanks a complex procedure, we have presented all the available details, and we have found what seems to have been the best combination of policy instruments in terms of allowing better performance in teaching as well as research.

The results are quite interesting both for the study of performance in HE and, more generally, for the study of effects of governance shifts in public policy and in turn with regard to the portfolios/mix of adopted policy instruments.

Where the literature evaluating the performance of university systems is concerned, our empirical evidence shows that there is not a unique and linear causation process capable of producing the expected outcome. By contrast, that same outcome emerges from a specific configuration of conditions (policy instruments) that must be present or absent to work. This way of thinking and thus this reading of the way governance arrangements work is explained by different combinations of instrumental shapes that can carry similar effects and thus by the way these combinations should be contextualized in the proper manner. In this sense, the empirical evidence presented in this paper shows that the evaluative literature on HE performance should find a third method that lies between the variable-oriented research strategy and the dense description of case-study analysis in order to fully grasp what is important in terms of performance.

Furthermore, with respect to the literature that has focused on the bundle of changes that have been discussed in a comparative perspective towards a common

template, the ‘steering at a distance’ model, there have clearly been certain national paths that have merged when translating the common template focused on certain instruments over others. Why these paths, these specific combinations of instruments, have been chosen is not of interest in this paper, but it could be taken into consideration in the future for a better understanding and explanation of the process of governance shifts, their features, their drivers and their decisional outputs.

Finally, regarding the broader literature on policy instruments, this paper could be a contribute to better understanding how these types of mixes actually work and how often only a limited combination of them actually impact the outcome of interest, despite the large number of the adopted set of instruments,.

There are obviously different possible paths for further research starting from the approach we have presented in this paper. We reference to two of them here.

First would be the challenge of extending the research while including countries belonging to different geo-political contexts and with different legacies in governing HE systems.

Second would be the challenge of deepening the analysis to grasp the working rules of the shapes through which the substantial instruments are used. This would also mean working on the dimensions of the rules through which each shape is designed. We refer to the rules through which decisional powers and competences are attributed and accountability rules are fixed when the shapes are designed. For example, regarding the use of loans, the focus should be on the rules of eligibility, the rules of reimbursement, the actors responsible for deciding who can get the loan, and so on. Obviously, this analysis would be a very complex path, but it could be very interesting and promising for definitively grasping how instruments work in day-to-day policy dynamics.

We definitively know that policy instruments matter, and we should work in the future to better understand how they do their job at the micro-level.

## REFERENCES


Aghion P., Dewatripont M., Hoxby C., Mas-Colell A., Sapir A. (2008), *Higher Aspirations: An Agenda for Reforming European Universities*, Brussels: Bruegel.

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.

Ansell B.W. (2008), 'University Challenges: Explaining Institutional Change in Higher Education'. *World Politics*, 60(1): 189–230.

Berg-Schlosser D., De Meur G., Rihouxand B., Ragin C.C. (2008), 'Qualitative Comparative Analysis (QCA) As An Approach', in B. Rihoux and C.C. Ragin (eds.), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, (pp. 1–18), Thousand Oak: Sage.

Blockmans W., Engwall L, and D. Weaire (eds.) (2014), *Bibliometrics: Use and Abuse in the Review of Research Performance*; Portland: Portland Press.

Braun D. and Merrien F.X. (eds.) (1999), *Towards a New Model of Governance for Universities?*, London: Jessica Kingsley Publishers.

Braga M., Checchi D., and Meschi E. (2013), 'Educational Policies in a Long-Run Perspective'. *Economic Policy*, 28(73): 45–100.

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.

Capano G., Reyner J. and Zito, A. (2012), 'Governance from the Bottom up: Complexity and Divergence in Comparative Perspective'. *Public Administration*, 90(1), 56–73.

Capano G. and Regini M. (2014). 'Governance Reforms and Organizational Dilemmas in European Universities'. *Comparative Education Review*, 56(1), 73–103.

Capano, G., Howlett, M. Ramesh, M. (eds) (2015. *Varieties of Governance*. London. Palgrave.

Capano G., Regini M. and Turri M. (2016). *Changing Governance in Universities. Italian Higher Education in Comparative Perspective*, London: Palgrave MacMillan.

CHEPS. 2006. *The extent and impact of higher education governance reform across Europe.* Final Report to the Directorate General for Education and Culture Of European Commission. http://ec.europa.eu/education/.

CHEPS (2015), *Performance-Based Funding and Performance Agreements in Fourteen Higher Education Systems*, Twente, NL: Cheps.

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

De Rijcke S., Wouters P.F., Rushforth A.D., Franssen T.P. and Hammarfelt B. (2016), 'Evaluation Practices and Effects of Indicator Use—a Literature Review'. *Research Evaluation*, 25(2): 161–169.

Dobbins M. and Knill C. (2014), *Higher Education Governance and Policy Change in Western Europe*, Basingstoke: Palgrave Macmillan.

European Commission/EACEA/Eurydice, (2016), *National Student Fee and Support Systems in European Higher Education – 2016/17. Eurydice Facts and Figures*, Luxembourg: Publications Office of the European Union.

European University Association (EUA) (2015), *European Funding Observatory 2014*, Brussels: EUA.

European University Association (EUA) (2017), *University Autonomy in Europe III,* Brussels: EUA

Flannery D. and O'Donoghue C. (2011), 'The Life-cycle Impact of Alternative Higher Education Finance Systems in Ireland'. *The Economic and Social Review*, 42(3): 237–270.

Frølich N. (2011), 'Multi-layered Accountability. Performance-based Funding of Universities'. *Public Administration*, 89(3): 840–859.

Frolich N., Schmidt E. and Roma M. (2010). Funding Systems for Higher Education and Their Impacts on Institutional Strategies and Academia**:** A Comparative Perspective**.** *International Journal of Educational Management*, 24(1): 7–21.

Gornitzka A., Kogan M. and Amaral A. (eds.) (2005), *Reform and Change in Higher Education*, Dordrecht: Springer.

Hacker, S.J. (2004). Reform Without Change, Change Without Reform: The Politics of US Health Policy Reform in Comparative Perspective, in M.A. Levin and M. Shapiro (eds.), *Transatlantic Policymaking in an Age of Austerity: Diversity and Drift*, (pp. 13–63), Washington, DC: Georgetown University Press.

Heller D. and Callender C. (eds.) (2013), *Student Financing of Higher Education*: *A Comparative Perspective*, London: Routledge.

Hicks D. (2012), 'Performance-based University Research Funding Systems'. *Research Policy*, 41(2): 251–261.

Hood C. (1983), *The Tools of Government*, London: Macmillan.

Howlett M. (2000), 'Managing the 'Hollow State': Procedural Policy Instruments and Modern Governance'. *Canadian Public Administration*, 43(4): 412–431.

Howlett, M. (2005). What is a policy instrument? Policy tools, policy mixes and policy-implementation styles. In P. Eliadis, M. M. Hill, & M. Howlett (Eds.), Designing government. From instruments to governance (pp. 311–350). Montreal and Kingston: McGill-Queen's University Press.

Howlett M. (2011), *Designing Public Policies: Principles and Instruments*, London: Routledge.

Howlett M. and Del Rio P. (2015), 'The Parameters of Policy Portfolios: Verticality and Horizontality in Design Spaces and Their Consequences for Policy Mix Formulation'. *Environment and Planning C: Government and Policy*, 33: 1233–1245.

Huisman J. (eds.) (2009), *International Perspectives on the Governance of Higher Education*, London: Routledge.

Huitema D. and Meijerink S. (eds.) (2009), *Water Policy Entrepreneurs: A Research Companion to Water Transitions Around the Globe*, Camberley, UK – Northampton, USA: Edward Elgar Publishing.

Knott A. and Payne A. (2004), 'The Impact of State Governance Structures on Management and Performance of Public Organizations: A Study of Higher Education Institutions'. *Journal of Policy Analysis and Management*, 23(1): 13–30

Ingram H. and Schneider A.I. (1990), 'The Behavioral Assumptions of Policy Tools'. *The Journal of Politics*, 52(02): 510–529.

Jano D. (2016), 'Compliance with EU Legislation in the Pre-accession Countries of South East Europe (2005-2011): A Fuzzy-set Qualitative Comparative Analysis', *Journal of European Integration*, 38(1): 1-22.

Jonkers K. and Zacharewicz T. (2016). *Research Performance Based Funding Systems: a Comparative Assessment*, Brussels: Joint Research Centre, the European Commission's in-house science service.

Jordan A., Benson D., Zito A. and Wurzel R. (2012), 'Environmental Policy: Governing by Multiple Policy Instruments?', in J.J. Richardson (ed.), *Constructing a Policy State? Policy Dynamics in the EU*, (pp. 104–124), Oxford: Oxford University Press.

Lazzaretti L. and Tavoletti E. (2006), **'**Governance Shifts in Higher Education: A Cross National Comparison'. *European Educational Research Journal*, 5(1): 18–37.

Liefner I. (2003), 'Funding, Resource Allocation, and Performance in Higher Education Systems'. *Higher Education*, 46(4): 469–489.

Linder S.H. and Peters B.G. (1990), 'An Institutional Approach to the Theory of Policymaking: The Role of Conscious Choice in Policy Formulation'. *The Journal of Theoretical Politics*, 2: 59–83.

Marx A., Rihoux B. and Ragin C.C. (2014), 'The Origins, Development, and Application of Qualitative Comparative Analysis: The First 25 Years'. *European Political Science Review*, 6(1): 115–142.

Maassen P. and Olsen J.P. (2007), (eds.), *University Dynamics and European Integration*, Dordrecht: Springer.

Meek V.L. and Van der Lee J. (2005), 'Performance Indicators for Assessing and Benchmarking Research Capacities in Universities', in M. Yahaya and I. Abdullah (eds.), *Enhancing Quality and Strengthening Research Capacities in Universities: Polices and Best Practices*, (pp. 1–72), Kuala Lumpur: University Kebangsaan Malaysia,

Paradeise C., Reale E., Bleiklie I. and Ferlie E. (eds.) (2009), *University Governance. Western European Comparative Perspectives*, Dordrecht: Springer.

Phidd R. and Doern G.B. (eds.) (1983), *Canadian Public Policy: Ideas, Structure, Process*; Toronto: Methuen.

Pierson C. (2004), *The Modern State*, London: Routledge.

Rabovsky T. (2012), 'Accountability in Higher Education: Exploring Impacts on State Budgets and Institutional Spending Patterns'. *Journal of Public Administration Research and Theory*, 22(4): 675–700.

Ragin C.C. (1987), *The Comparative Method*, Berkeley: University of California Press.

Ragin C.C. (2000), *Fuzzy-set Social Science*, Chicago: The University of Chicago Press.

Ragin C.C. (2008), *Redesigning Social Inquiry: Fuzzy Sets and Beyond*, Chicago: The University of Chicago Press.

Richardson R.C. and Martinez M.C. (2009), *Policy and Performance in American Higher Education: An Examination of Cases Across State Systems*, Baltimore: The John Hopkins University Press.

Rihoux B. and Ragin C.C. (eds.) (2009), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, Los Angeles (USA): Sage.

Rihoux B., Alamos-Concha P., Bol D., Marx A. and Rezsöhazy I. (2013), 'From Niche to Mainstream Method? A Comprehensive Mapping of QCA Application in Journal Articles from 1984 to 2011'. *Political Research Quarterly*, 66(1): 175–184.

Ring I. and Schroter-Schlaack C. (2010). *Instrument Mixes for Biodiversity Policies*, Leipzig: Helmholtz Centre for Environmental Research.

Russo C. (ed.) (2013). *Handbook of Comparative Higher Education Law*, New York: Rowman and Littlefield.

Rutherford A. and Rabovski T. (2014), 'Evaluating Impacts of Performance Funding Policies on Student Outcomes in Higher Education'. *The Annals of The American Academy of Political and Social Science*, 655(1): 185–209.

Salamon L.M. (eds.) (2002), *The Tools of Government: A Guide to the New Governance*, Oxford: Oxford University Press.

Schaffrin A., Sewerin S. and Seubert S. (2014), 'The Innovativeness of National Policy Portfolios – Climate Policy Change in Austria, Germany, and the UK'; *Enviromental Politics*, 23(5): 860–883.

Schneider C.Q. and Wagemann C. (2006), 'Reducing Complexity in Qualitative Comparative Analysis (QCA): Remote and Proximate Factors and the Consolidation of Democracy'. *European Journal of Political Research*, 45: 751–786.

Schneider, C.Q, and I. Rohlfing. 2013. "Combining QCA and Process Tracing in Set-Theoretic Multi-Method Research." *Sociological Methods & Research* 42(4):559-597.

Schneider C.Q. and Wagemann C. (2012), *Set-Theoretic Methods for the Social Sciences*, Cambridge: Cambridge University Press.

Shattock M.L. (eds.) (2014), *International Trends in University Governance*, London: Routledge.

Tandberg D. and Hillman N. (2014), 'State Higher Education Performance Funding: Data, Outcomes, and Policy Implications'. *Journal of Education Finance* 39(3): 222–243.

Tosun J. (2013), *Risk Regulation in Europe. Assessing the Application of the Precautionary Principle*, New York: Springer.

Trakman L. (2008), 'Modelling University Governance'. *Higher Education Quarterly*, 62(1-2): 63–83.

Vedung E. (1998), 'Policy Instruments: Typologies and Theories', in M.L. Bemelmans-Videc, R.C. Rist, and E. Vedung (eds.), *Carrots, Sticks, and Sermons: Policy Instruments and Their Evaluation*, (pp. 21–58), New Brunswick, NJ: Transaction.

Volkwein J. and Tandberg D. (2008), 'Measuring Up: Examining the Connections Among State Structural Characteristics, Regulatory Practices, and Performance'. *Research in Higher Education*, 49(2): 180–197.

Wagemann C. and Schneider C.Q. (2010),'Qualitative Comparative Analysis (QCA) and Fuzzy-Sets: Agenda for a Research Approach and a Data Analysis Technique'. *Comparative Sociology*, 9(3): 376–396.

Whitley R. and Gläser J. (eds.) (2007), *Changing Governance of the Sciences. The Advent of Research Evaluation Systems*, Dordrecht: Springer.

Willemse N. and De Beer P. (2012), 'Three Worlds of Educational Welfare States? A Comparative Study of Higher Education Systems Across Welfare States'. *Journal of European Social Policy*, 22(2): 105–117.

Winter-Ebmer R. and Wirz A. (2002), 'Public Funding and Enrolment into Higher Education in Europe'. *IZA Discussion Papers 503*, Institute for the Study of Labor (IZA).

Woodside K. (1983), 'The Political Economy of Policy Instruments: Tax Expenditures and Subsidies', in M. Atkinson and M. Chandler (eds.), *The Politics of Canadian Public Policy*, (pp. 173–197), Toronto: University of Toronto Press.

## Appendix A – Classification of policy instruments

| Dimension | Variable | Policy instrument |
|---|---|---|
| *Regulation* | R1 | Accreditation (whereas the regulatory aspect is prevalent) |
| | R2 | Evaluation |
| | R3 | Assessment (how the procedure of assessment should be done) |
| | R4 | Agency for evaluation |
| | R5 | Agency for accreditation |
| | R6 | Regulation on content of curricula: more constraints |
| | R7 | Regulation on content of curricula: more opportunities |
| | R8 | Regulation on academic career and recruitment: more constraints |
| | R9 | Regulation on academic career and recruitment: more opportunities |
| | R10 | Regulation on students' taxation: more constraints |
| | R11 | Regulation on students' taxation: more opportunities |
| | R12 | Rules on administrative procedures (financial, legal, budget accountability): constraints |
| | R13 | Rules on administrative procedures (financial, legal, budget accountability): opportunities |
| | R14 | Regulation on students' admissions: more constraints |
| | R15 | Regulation on students' admissions: more opportunities |
| | R16 | Regulation on institutional governance: more constraints |
| | R17 | Regulation on institutional governance: more opportunities |
| | R18 | Rules to borrow funds |
| | R19 | Teaching assessment (how the results of teaching assessment should be obtained) |
| | R20 | Research assessment (how the results of research assessment should be obtained) |
| | R21 | Contracts |
| | R22 | Rules on goals in teaching |
| *Expenditure* | E1 | Grants |
| | E2 | Subsidies |
| | E3 | Targeted-funding |
| | E4 | Lump-sum funding |
| | E5 | Performance/quality based teachers' salary |
| | E6 | Loans |
| | E7 | Performance based institutional funding |
| | E8 | Standard-cost per student |
| | E9 | Per-capita funding |
| *Taxation* | T1 | Tax-exemption |
| | T2 | Income-based student fees |
| | T3 | Merit-based student fees |
| | T4 | Service-based student fees |
| | T5 | Tax credit for private enterprises investing in Higher Education |
| *Information* | I1 | Accreditation (how programmes with accreditation should be made public) |
| | I2 | Transparency |
| | I3 | Certifications |
| | I4 | Monitoring and reporting |
| | I5 | Rankings |
| | I6 | Teaching Assessment (how the results of teaching assessment should be made public) |
| | I7 | Research Assessment (how the results of research assessment should be made public) |

## Appendix B – Calibration of conditions and outcome(s)

| *Conditions/outcomes* | *Full membership (1)* | *Point of maximum ambiguity (0.5)* | *Full non-membership (0)* |
|---|---|---|---|
| Tertiary 30-34 yo. | 29 | 23 | 18 |
| ARWU-top300 | 1.5 | -0.5 | -1.5 |
| R7 | 12 | 6.5 | 1 |
| R17 | 9 | 5.5 | 3 |
| R6 | 6 | 2.5 | 1 |
| R9 | 6 | 3.5 | 2 |
| R4 | 5 | 2.5 | 2 |
| R16 | 4 | 2.5 | 1 |
| R22 | 3.5 | 2.5 | 1.5 |
| R1 | 3.5 | 2.5 | 1.5 |
| R2 | 3.5 | 2.5 | 1.5 |
| R13 | 3.5 | 2.5 | 1.5 |
| R12 | 3.5 | 2.5 | 1.5 |
| R11 | 3 | 1.5 | 0 |
| R14 | 3.5 | 1.5 | 0 |
| R15 | 3.5 | 1.5 | 0 |
| R21 | 3 | 1.5 | 0 |
| E1 | 6 | 3.5 | 2 |
| E7 | 4 | 1.5 | 0 |
| E6 | 3.5 | 2.5 | 1.5 |
| E3 | 3 | 1.5 | 0 |
| E4 | 4 | 2.5 | 1.5 |
| T1 | 1.5 | 0.5 | 0 |
| T2 | 2 | 0.5 | 0 |
| T4 | 3 | 1.5 | 0 |
| T5 | 1.5 | 0.5 | 0 |
| I2 | 4 | 2.5 | 1 |
| I4 | 3 | 1.5 | 1 |
| I3 | 3 | 1.5 | 0 |
| I1 | 1 | 0.5 | 0 |
| I5 | 1 | 0.5 | 0 |