

NO GOOD RECOMMENDATION WITHOUT A MECHANISTIC EXPLANATION?

The case of higher education governance

Alessia Damonte

Università degli Studi, Milano

alessia.damonte@unimi.it

Giliberto Capano

Scuola Normale Superiore, Pisa

giliberto.capano@sns.it

Abstract. In higher education, reforms have long been driven by the theory that system performance depends on governance design; yet it remains far from clear which arrangements can actually deliver results, as shown in the analysis of various streams of research devoted to assessing performance in higher education. We reason that such a question can be better answered if research aims for a mechanistic explanation and operationalizes it to avoid the shortcomings of both ‘variable-oriented’ and ‘case-oriented’ strategies. We therefore develop a ‘diversity-oriented’ mechanistic framework that explains differences in performance by differences in policy tool mixes, which we define as governance regimes. This set of policy tools is meant as a configuration of properties of delivery vehicles, decision-making design, and accountability design. Such an explanatory focus has many advantages: policy tools are manipulable, as they depend on political and administrative decisions; moreover, they are efficient causes, as they trigger mechanisms at the individual level that directly account for both individual and institutional behavior and, hence, performance. Tool-based explanations therefore can more easily allow for policy learning and transfer than can ‘remote’ constitutional, historical, or cultural accounts.

Keywords. Explanation; Governance regimes; Higher education; Mechanisms; Policy tools.

1. INTRODUCTION

Over the last thirty years, the performance of higher education systems (hereafter, HESs) has become a prominent issue in a number of national and international agendas. Better education for a higher share of the population, excellence in research, and productive contributions to societal or regional development have all been identified as desirable goals that universities can and should meet. Improvements in all these dimensions have been programmatically promoted by many types of governments – through more or less extensive re-design of rules, subsidies, and fees, as well as through the re-distribution of decision-making powers and control rights. In essence, existing reform strategies have framed HES performance as a problem of poor systemic governance design – the solution to which lies in a shift toward “better” overall arrangements by changing what, with Howlett (2002, 2011), can be labeled ‘substantive’ policy tools (i.e., regulation, expenditure, taxation, and information) and ‘procedural’ policy tools (i.e., rules concerning decision-making and accountability).

In Europe, these developments prompted efforts to converge on a common template (van Vught 1989; Neave and van Vught 1991; Kickert 1997). Such efforts have been especially visible in the reforms pushing the “continental” governance model – characterized by hierarchical coordination through state-centered policies; no institutional autonomy; the powerful, all-pervasive authority of academic guilds; and faculties and schools as “confederations of chair-holders” (Clark 1983) – toward the model of English-speaking countries. These reforms have fostered organizational autonomy and differentiation among universities via the delegation of regulatory and expenditure power, broader freedom of choice for students and their families, and incentives to develop public and private partners – while institutional leadership and managerial steering capacity have been prized over collegiality in decisions (Gornitzka et al. 2005; Lazzaretti and Tavoletti 2006; Maassen and Olsen 2007; Trakman 2008; Huisman 2009; Paradeise et al. 2009; Capano and Regini 2014). In all of these transformations, the steering position of governments has been nevertheless maintained as the power to induce proper strategies via national standards, procedures for monitoring and evaluation, and criteria for financial rewards (Enders, De Boer, and Weyer 2013).

Such institutional solutions to HES performance have also been advocated at the European and international levels (European Commission 2006; Oecd 2008; World Bank

2008; Capano and Piattoni 2011). Far from being a European peculiarity, the mantra of ‘reforming governance’ has also characterized recent HES reforms on multiple continents. Interventions targeting HES governance have been the enduring hallmark of governmental agendas in Asia and Latin America (Mok 2010; Brunner and Villalobos 2014), as well as in key Anglo-Saxon countries, where governments have increased their intervention and regulation despite a tradition of institutional autonomy (El-Khawas 2005; Pick 2006; Jones 2009; McLendon and Hearn 2009; Schuetze, Bruneau, and Grosjean 2012; Capano 2015).

After more than three decades of reforms and ongoing adjustments, however, the effectiveness of HESs remains in question. Student enrolment and completion rates are seldom satisfactory; quality teaching has not been ensured; and universities’ developmental ‘third role’ has produced mixed results at best. More fundamentally, performance indicators have shown remarkable – and to some, alarming – variations both across and within countries (European Commission 2014; EUA 2014; OECD 2014). Moreover, there is no clear evidence that the new template can deliver results. Instead, it has been strongly criticized for the uneven redistributive effects of its implementation and, more generally, for its “false” narrative, as the claim of freedom would simply conceal a new subordination of universities to the government as agents to a principal (Lane and Kivisto 2008; Enders, De Boer, and Weyer 2013).

Underlying these debates, the basic question remains of *which governance arrangements generate good HES performance and which, instead, prove ineffective*.

In this article, we do not provide an answer to this question, which we believe can only be established empirically. Instead, we develop a reasonably viable framework for answering this question to avoid the shortcomings that characterize many previous analyses.

The argument runs as follows.

Section 2 investigates what defines a good explanatory model, arguing that such a model focuses on an aggregation of multiple factors (Mackie 1965) related to ‘mechanisms’ (Boudon 1981; Bhaksar 2008; Pawson 1989; Elster 1989). Mechanisms provide clear criteria for selecting the factors underlying positive and negative outcomes, while improving case comparability and model portability. Mechanisms cannot result in a viable model, however, unless a research strategy is chosen that specifies what should be observed and assumed, how coarse the gauges can be, and what knowledge aim is pursued. Such specifications implicate a trilemma, leading to ‘variable-oriented’, ‘case-oriented’ and ‘difference-oriented’ strategies

as alternative rationales of probation (Ragin 1997, 2000). Section 2 also reasons that suitable answers to our research question are better delivered by a difference-oriented approach.

Section 3 applies the lens of good explanatory models to mainstream explanations of HES performance. Variable-oriented studies are recognized as capable of assessing the predictive power of special factors from mechanistic hypotheses; yet, such factors are often gauged by additive indexes, or by limited interactions. Therefore, it is ultimately difficult to determine which element of governance explains performance differences. Conversely, case-oriented analyses are especially sensitive to the unexpected effects of the local implementation of single factors, thus highlighting how context can alter the meaning of a governance innovation. However, the findings of such analyses are seldom trustworthy or related to performance.

Section 4 builds on this recognition and develops an innovative framework for answering our question that is consistent with the ‘difference-oriented’ approach. This framework is based on the tenet that HES performance depends on the behavior of key actors in the field, whose strategies are inevitably shaped by the mix of policy tools in use. It clarifies the mechanism linking substantive tools to performance by referring to Etzioni’s theory of compliance and Vedung’s understanding of substantive tools. It also agrees with Salamon (2002) in that tool effectiveness rests on the properties of special delivery vehicles, which in turn only become credible when backed by a proper set of procedural instruments (Howlett, 2011) – namely, by decision-making and accountability design.

Section 5 offers some concluding remarks.

2. THE SEARCH FOR GOOD EXPLANATIONS

The literature on causality suggests that good explanations depend on a sound causal claim that can be proven true. The first point therefore consists in establishing a sound claim; the second, in defining and conveying its key characteristics for operationalization; the third, in making the operationalization fit some special assessment strategy. These three points are far from independent, but they can be addressed as separate steps.

2.1. Establishing a sound claim

What makes a causal claim ‘sound’ has long been far from obvious. One of the main issues concerns the difficult relationship between, on the one side, an understanding of causality as true statements of general determination and, on the other, bountiful records of local exceptions to those general statements that venture beyond the bounds of triviality. At the extremes, these two understandings have long pitted advocates of causation as the effect of statistically evident laws against those to whom causation, if any, is not intelligible beyond perceived contingency, so that subjective local descriptions exhaust possible knowledge about the world (Pawson 1989). In between, a convincing framework that squares ‘type’ with ‘token’ understandings of causality resonates with Bhaksar’s philosophy of science (2008).

To Bhaksar, causality is a latent power that actual objects can bear to a little or a great extent because of some special property, and which unfolds into effects under special conditions. Knowledge about ‘type’ causality abstracts from actual conditions and understands a causal power as a probabilistic tendency of a factor to obtain. Nevertheless, knowledge of a tendency cannot account for the reasons that make this power to unfold in certain actual cases alone – that is, for substantive differences in effects across cases. This gap entails another type of knowledge, concerning those conditions that enabled, triggered or hindered the ‘actualization’ of the latent power in single local contexts. The driving question is always ‘why’; yet, the answer is more complex, and calls for the opening of the ‘black box’ of causation with instruments of logic.

Following Mackie (1965), actual causation becomes intelligible as a local array of INUS factors – i.e., of Insufficient-but-Necessary parts conjoined into an Unnecessary-but-Sufficient complex that accounts for a special outcome. Explanations are therefore ‘*post factum*’ statements linking a complex antecedent to some consequent occurrence within a given ‘causal field’. The relationship between potential causality and actual causation therefore becomes a mere matter of specification, or ‘degrees of abstraction’, on a continuum from, on the one hand, the general ‘gappy tendency’ to obtain of a single INUS factor across a large number of cases and, on the other, the detailed local configuration of all the INUS factors required to account for a single occurrence (Mahoney 2008). Different degrees of abstraction do not mean that the token–type duality resurfaces, only that attention is paid to the same phenomenon in fields with different scopes. This is true until differently specified

statements all follow from a similar causal theory of the ‘generative process’ or ‘mechanism’ (Boudon 1979; Bhaksar 2008; Pawson 1989; Elster 1989) responsible for an outcome.

Indeed, mechanistic theories can support both special explanations and general expectations. This, however, is not the only heuristic gain – especially when compared with standard, law-like approaches (Hedström and Ylikoski 2010). Mechanisms also provide a better rationale for extrapolating causal findings into new settings. More fundamentally, ‘generative’ thinking vindicates the asymmetric nature of causation, which otherwise may appear as an identity of *explanandum* and *explanans* on which a direction is artificially imposed. Besides, it can avoid the faulty ascription of causal meaning to predictors or otherwise spurious associations. For these gains to be warranted, however, the explanatory model has to identify and represent key features of the generative process correctly.

2.2. Conveying key characteristics

A ‘consensus’ definition describes a mechanism as a “system with multiple components, which interact to produce some overall phenomenon” (Fagan 2012:453). When all constituents are given in a field and activated, such systems ‘actualize’ potential powers, thus engendering local and general effects. Correct causal statements therefore are those that identify the constituents that are alone required for a generative process to actually occur.

As such, the definition indicates two intertwined requirements. A factor is a cause if it gauges a property

[C1] of the entities constituting the mechanism and

[C2] relevant to the generative process in a defined field.

In [C2], ‘relevance’ is a twofold requirement. On the one hand, it refers to [C2.a] a chosen scope of observation. The generative process takes local forms – sometimes quite idiosyncratic ones—such that the width of the field of analysis defines which special entities carry the properties that bring it about, therefore the domain wherein the specified causal statement holds true (Mackie 1965). On the other hand, relevance refers to [C2.b] the contribution that special properties make to outcome generation. In this sense, relevant causes are ‘difference makers’ (Lewis 2001), the absence of which is capable of transforming a case in which the mechanism holds into one in which it defuses. The requirement thus shifts the issue of good explanations from the ontological to the epistemological level, as the

contribution to the mechanism can be conceived counterfactually and appreciated empirically.

[C1] states that proper mechanistic causes relate to entities that [C1.a] lay at a 'lower' or 'deeper' analytical level than the one at which the outcome is observed and [C1.b] carry some property that allows for a key generative activity. As such, mechanistic causes satisfy the principles of 'modularity' (i.e., properties are individually manipulable) and 'invariance' (i.e., such manipulations affect the intensity of the generative process but do not transform the nature of the mechanism) (Woodward 1997). So defined, causes truly are 'priors' to an outcome (Craver 2007).

Mechanistic thinking therefore prevents the mis-attribution of generative power while justifying the direction of causation. At the same time, it exposes explanations to a paradox. The foundation of causality at lower levels of observation potentially enables an infinite regress, as each key activity in a mechanism can in turn constitute the outcome of a further mechanism that ideally requires a reference to deeper entities and properties. A strategy for circumventing the paradox disaggregates mechanisms into ultimate deterministic, law-like generalizations that uncaused entities could be assumed to follow (Salmon 1984, Elster 1989, Glennan 1996). However, such reductionism has been contested: generative processes rather appear as non-fundamental empirical relationships that cannot be boiled down to the mere aggregate of constituting entities and properties (Fagan 2012, Cartwright 1997, 2007). The alternative strategy thus 'bottoms out' explanations within each particular discipline by establishing assumptions regarding properties, entities, or activities that are not directly observed by the researcher but are nevertheless proven when related data are found to be consistent with such assumptions (Coleman 1990, Bhaksar 2008).

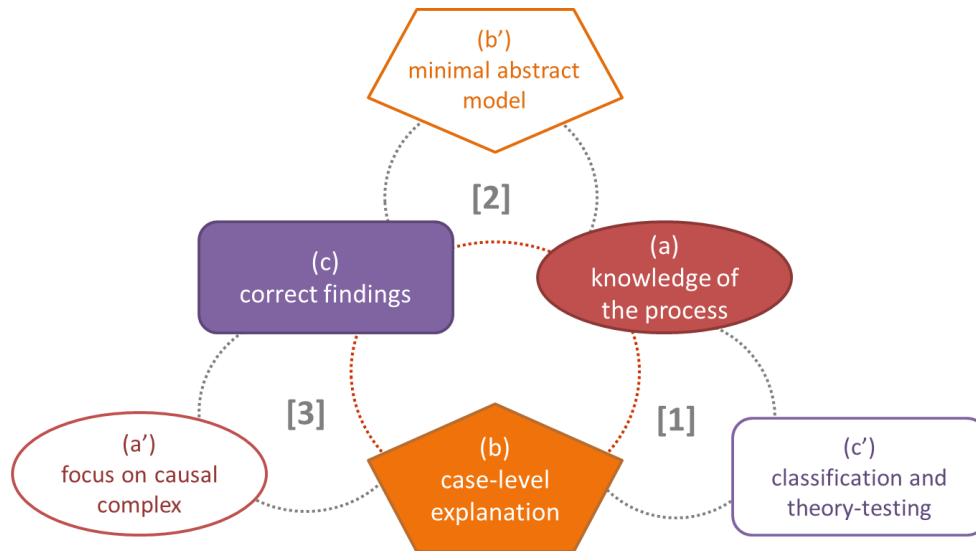
A stronger starting point in mechanistic modeling is therefore theory-driven (Hedström and Ylikoski 2010). Herein, the principles of modularity and invariance become epistemic devices for detailing a model, then proving it true (Woodward 1997, 2000; Pearl 2000).

2.3. Making hypotheses fit a probing strategy

In essence, the literature agrees that a strong explanatory hypothesis relates modular difference-makers to an invariant generative process. Such a hypothesis in principle allows for three gains: (a) knowledge of how the generative process led to the effect; (b) a case-level grasp of causation; and (c) findings correct enough for prescription. Unfortunately, these

three gains cannot be secured simultaneously by a single assessment strategy; instead, they define a trilemma in mechanism-oriented research strategies (fig.1).

Figure 1. The trilemma of assessing mechanistic hypotheses



Thus, techniques that (a) model the generative process explicitly – usually, as a chain of functional dependencies between activities or underlying properties – must choose.

If they maintain (b) a case-level focus, they have to abandon trustworthy inferences for (c') case understanding, classification, and theory-testing. This is what [1] 'case-oriented strategies' usually deliver (George and Bennett 2005; Hall 2006; Brady and Collier 2010; Mahoney 2012; Bennett and Checkel 2014). These strategies are driven by the puzzle of a special chain of causes and effects that generate outcomes in very specific contexts and are often singular in nature. However, the answer may be excessively focused on inductive theory creation; even when the analytic rationale follows the rigor of Bayesian tests for deciding whether a case is an instance of a theory, the representation of the generative process can be redundant, with little possibility of knowing which parts are essential for it to obtain. Additionally, the researcher enjoys remarkable latitude in constructing representations, and evidence is often collected under the assumption of case distinctiveness – a process that may produce results as convincing understandings of realities, yet seldom as cumulative knowledge or as trustworthy grounds for prescription.

If techniques maintain (c) the aim of strong inferences from as correct a model as possible, they have to give up the case-level focus for (*b'*) minimal abstract representations of causal structures. This is the case for [2] 'variable-oriented' strategies (Pearl 1988, 2000; Shadish et al. 2002). Analyses consider mechanisms to be observable structures of functional dependencies which trustworthily represent actual causation to the extent that the independence of explanatory factors – either as single vectors or as Markovian chains, either as such or after instrumentation – statistically holds. Omitted INUS factors are accounted as 'independent disturbances' unless otherwise proven. The mechanism is thus portrayed as the sum of dismembered functions, the number and aggregation of which only depend on the researcher's choice. As a consequence, valid results basically prove the 'gappy tendency' of single vectors to influence outcome occurrence across a sample, and can be trusted as predictors insofar as existing unknown conditions remain constant.

There is still a third methodological alternative. Strategies can maintain (b) the case-level focus on causation and (c) an ambition to achieve trustworthy results for prescriptions; so, instead of unfolding the generative process, the analysis has (*a'*) to 'bottom it out' as an unobserved chemical reaction, instead focusing on the compound that can account for its occurrence within a given population. This is the case of [3] 'diversity-oriented' strategies (Ragin 1987, 2000, 2008; Schneider and Wagemann 2012; Mahoney and Goertz 2006). These strategies are theory-driven, moving from a hypothesis concerning the unobserved generative process to the observable properties that can effect it. The hypothesis is therefore operationalized as an expectation that the outcome would have certainly occurred in a world in which all selected properties were given together. The driving question asks whether the generative hypothesis can account for differences in real-world outcomes, i.e., across actual cases, each conceived as an instance of the mechanism. Boolean inductions follow tests of variables' difference-making power inspired by Mill's canons of agreement and difference, restricted to the population under analysis and consistent with the tenet of non-contradictoriness. While an initial hypothesis can therefore be quite 'portable', solutions do not apply beyond the scope condition. Within this condition, however, solutions allow for lesson learning, as they ascribe generative power to particular complexes of necessary and sufficient conditions proven to explain single subpopulations, or cases, without contradiction.

The third set of strategies thus develops mechanistic hypotheses better suited to providing an answer to our initial question. However, this set is also least often applied for the probation of HES governance designs, as the following section will show.

3. THE SLIPPERY NATURE OF PERFORMANCE IN HES: CURRENT EXPLANATIONS AND THEIR LIMITS

Many studies of HES governance have been conducted in recent decades. These studies primarily belong to the fields of political science, public administration, the sociology of education, and public policy, as well as a few in economics. Beyond disciplines, however, two main streams can be recognized: that of statistical analyses assessing the general net effect of single interest factors on the performance of a sample of HESs; and that of qualitative studies emphasizing the idiosyncratic character of the relationship between complex governance arrangements and performance in single and multiple HESs.

3.1. Net-effect explanations

The first stream of research narrows explanations to a single determinant that accounts for variations in specific HES performance indicators. Here, determinants are identified as single policy tools, such as outlays; algebraic indexes, as in gauges of the organizational dimension of universities; or, less commonly, as interactions. We present below some examples drawn from this stream of literature.

The amount of public funding is usually considered a relevant determinant for achieving significant HES performance. Its correlation with access is based on the assumption that public subsidization re-balances the individually perceived trade-off between the marginal cost of accumulating years of education and the marginal benefits in terms of life-chances (Viaene and Zilcha 2013). This correlation has been tested, for example, by Winter-Ebmer and Wirz (2002) who, by analyzing 14 European countries using regression analysis, found a significant correlation between public funding and greater enrollment in higher education, although there is some evidence that the structure and organization of high school can influence enrolment in higher education. Moreover, they found that tuition increases have a significant counter-impact with respect to public funding. Thus, the reader grasps that at least

three factors impact enrolment in HES, but interactions among such factors are not disentangled.

Performance funding has been widely considered the cause of HES performance both in teaching and in research. On this topic, the most relevant studies are those from the US, where many states, beginning from the end of the 1970s, have introduced measures to allocate extra resources. Empirical evidence (Volkwein and Talberg 2008; Robovsky 2012; Tandberg and Hillman 2014)¹, however, shows that the impact has clearly been weak, if any, and not clearly independent of other factors such as changes in state governance or tuition policy. For example, Rutheford and Rabovsky (2014) analyzed the performance of 568 four-year degree-granting public institutions in fifty states (1993-2010); performance was associated with three indicators: six-year graduation rates, retention rates, and bachelor's degree production. The explanatory model focuses on specific channels of state funding policies (e.g., state funding that is directly linked to specific student outcomes; in the form of bonus incentives; or as a portion of base funding), controlling for duration, and thus is developed as a time-series cross-sectional analysis. Many control variables are included in the model (e.g., five regarding students; five regarding the features of institutions; and three concerning the political context). The result is that, in general, performance funding does not have a real effect on student outcomes, although a moderate positive effect emerges from those policies that allocate performance funding through base funding. However, other variables have a larger effect on performance (e.g., student profiles, institutional characteristics, and state political ideology).

Aghion et al. (2010) performed a complex analysis of the institutional determinants of HES performance. They explored developments in EU countries using a regression analysis of survey responses on the Shanghai ranking, while a more sophisticated difference-in-difference analysis examined US state systems with respect to utility patents produced by

¹ It should be emphasized that all these studies focused on the so-called first wave of performance funding in the US, when some states decided to allocate additional funding (as a small percentage of the overall amount of public funding) to HES. Subsequently, many of those states abandoned this policy (for many reasons). Beginning in 2009, a second wave of performance funding policy was initiated; some states introduced performance funding in the funding formula by focusing on output goals (percentage of retention and graduation) For example, since 2009, Tennessee has allocated all of its public funding through a performance funding formula. However, right now there are no reliable data on the effects of this new policy (Dougherty et al. 2014).

resident cohorts. Ultimately, the authors concluded that public expenditures at research universities do not increase HES performance in states that, together, are far from the technological frontier, offer low autonomy to public universities, and have little competition from private universities. Their US analysis is especially interesting, as it tests an array of interactions instead of single indexes; moreover, it applies a sophisticated strategy for instrumenting expenditures. Finally, the overall model scores a remarkable R-squared of 0.992. However, some doubts remain concerning the trustworthiness of the results, for which significance is not clearly reported, while model complexity and unclear test presentation may raise issues of over-fitting. The results may therefore be more suggestive than clear-cut – a flaw that has not inhibited their wide circulation and influence.

Braga, Checchi and Meschi (2013) provided a more elegant model of HES performance when investigating the impact of seven decades of reforms on educational attainment (defined in terms of completed years of education) in the EU. The mechanistic hypothesis is clearly present in their work, and variables are selected because they operationalize ‘channels’ that directly impact students’ preferences for investing in education. Thus, the model focuses on financial support to students, the selectivity of admission criteria, and institutional autonomy; and recognizes the possibility of differential effects for students with high and low performance by country, cohort and survey. Their findings show that financial aid is important but for students with poor family backgrounds, while selectivity is not meaningful. Moreover, the index of university autonomy is – in contrast to a common reading of Aghion *et al.* – negatively correlated with average educational attainment. Nevertheless, while financial aid is operationalized through a thorough index that credibly captures the direction of causation, institutional autonomy is gauged by the average of many specific indexes (e.g., budget, recruitment, organization, logistics, course organization, self-evaluation and development plans). Thus, each index is considered an independent and substitutable dimension of the concept, such that high scores on any index (i.e., planning) can compensate for low scores in any other (i.e., recruitment) – an assumption that even intuitively appears weak. Ultimately, their conclusions regarding the impact of institutional properties on educational attainment may be flawed due to measurement issues.

Another study of 48 countries showed that those HESs that better perform in all three university missions are above all those in which there is significant overall funding from public transfers and private donors (Williams, de Rassenfosse, Jensen and Marginson 2013).

This study was based on the ranking of selected countries according to 11 variables that are grouped under three labels: Resources (5 variables), Environment (4 variables) and Connectivity (2 variables). Nine outputs are grouped together. Then, after a qualitative ranking of countries for each label, at the aggregate level, regression results revealed that (again in contrast to Aghion et al.) research performance is higher in those countries where governmental expenditures on research as a percentage of GDP are higher, while performance in participation rates is higher when the total funding is higher. This study recognizes that there are interesting exceptions, such as countries with low funding but high overall performance, and that governance should be more effectively unfolded, as better performance could rest on different funding and institutional design specifications.

3.2. Understanding context

A second strategy focuses on instrumental change in few cases, often in single case studies. Differences are understood as idiosyncratic changes in HES governance that mirror the nature of broader national political and institutional systems. Thus, system performance can often be identified in the legal output that draws and enacts changes in governance. Attention shifts toward contexts, within which meanings, uses and functions of innovations are idiosyncratically molded by the characteristics of constitutional and political settings, legacies, veto points, and ideas. We present below some examples drawn from this stream of literature.

Regarding the evaluation of research, Capano and Turri (2015) emphasized how it has been introduced in many countries, but only in England and Italy have the results of this research exercise become a government criterion for allocating public monies; Musselin (2013) showed that, in France, the same tool sorted reputational effects, improving the position of some universities in the market for private funding.

With respect to cost-sharing, studies based on a few cases have shown that – unlike those rooted in statistical methods – there is not a straightforward negative correlation between increasing private sharing of the cost of HES and the level of demand (Johnstone 2000; Usher 2005). For example, Orr, Wespel, and Usher (2014) operationalized an analytical description of the systemic evolution (1995-2010) in nine countries to show that the rise of fees did not change the gender composition or socio-economic origin of enrollees.

Qualitative perspectives on institutional autonomy have observed how autonomous measures have been introduced or widened in all continental European countries (Huisman 2009; Paradeise et al. 2009; Dobbins and Knill 2014; Shattock 2014), yet implemented in quite different ways. Dutch universities have gained a greater degree of managerial autonomy than institutional autonomy (Enders, De Boer, Weyer 2013), while in Denmark universities enjoy a higher level of institutional autonomy, falling just short of their English peers (Wright and Oberg 2008; EUA 2011). In Italy, governmental rhetoric notwithstanding, institutional autonomy has been compressed significantly in the decade since its introduction (Donina, Meoli and Paleari 2015).

When linked to performance, as in US studies, the qualitative approach has obtained contradictory empirical evidence with respect to the relationship between centralized/decentralized governance in state public systems. For example, Richardson and Martinez (2009), in their study of five states, have shown how those with a centralized governance system performed better (with respect to access and graduation rates) than those with a decentralized design. Thus they reached exactly the opposite conclusion of that offered in a previous study (based on a multivariate analysis of all the PhD-granting universities), which concluded that decentralized state governance performs better (Knott and Payne 2004).

Many studies agree on the strategic role played by allocation rule changes in reforming HESs. They also emphasize that similar rationales have been adopted to justify very different funding systems (Frolich, Schmidt, and Rosa 2010). Performance funding has been designed differently across countries, as have performance contracts (Cheps 2015). Governments have adopted very strict contracts in Denmark (with clear performance goals to be reached by institutions), and the result seems very positive in terms of access and graduation; in Italy, as well other countries, these contracts are softer and yielded mixed results.

Another stream of work has focused on quality assurance systems, again reporting how they have spread worldwide, but with very different effects than expected (Dill and Berken 2010; Rosa and Amaral 2014). As such, related performance – for instance, in terms of wider access or graduation rates as recorded in international statistics – is quite different, and it always makes sense as a national effect.

3.3. Summing up

Mainstream assessments of the explanatory capacity of HES governance thus seem to be affected by different syndromes that weaken the credibility of results – at least as a strong basis for policy prescriptions at the case level.

On the side of net-effect analyses, even when the generative process is explicitly cited as the rationale for the specification, governance measures always raise issues of endogeneity and heteroskedasticity that instrumentation and fixed effects both help to address from a statistical viewpoint; yet, these solutions also overburden the model with assumptions that may weaken the portability of results. These analyses also operationalize jointness in causation as composite indexes or discrete interactions among a limited number of factors, each with a separate net effect on the outcome. The result may be a mis-specified model, which again casts doubt on the soundness of the resulting conclusions despite their statistical face validity. Moreover, these assessments may conceal the existence of outliers – hence, of signals that alternative explanations exist.

Outliers instead are the standard assumption of analyses intended to understand the effect of governance. They tend to shift focus from outcomes to individual governance reforms and to portray them as the result of some ongoing process that is highly conditioned by the political and cultural context. Cases thus provide rich empirical evidence of process specificities but are often contradictory and far from conclusive. Thus, they mainly highlight the considerable variety of real solutions that the governance puzzle may be given, but hardly provide trustworthy causation results.

Interestingly, almost no analyses seem to have addressed the puzzle of HES governance performance from a pure difference-oriented perspective, despite evidence that it can provide especially suitable explanations to system performance (Damonte 2014). In the next section, we will show how that gap can be filled.

4. EXPLAINING HES PERFORMANCE:

A MODEL FIT FOR DIFFERENCE-ORIENTED STRATEGIES

A mechanistic model for difference-oriented strategies focuses on modular governance elements that make a difference among good and bad performers by proving relevant to an unobservable generative process. Thus, the first step is to define the generative mechanism.

4.1. The generative process underlying system performance

We maintain that HES performance depends on the behavior of ‘producers’ (e.g., individual scholars and personnel organized in departments and universities) and ‘consumers’ (e.g., students, their families, and economic actors) who are assumed to interact in five crucial ‘functional areas’ that characterize HES, namely: (a) student access; (b) teaching; (c) research; (d) academic recruitment and career development; and (e) local development (third mission).

We also assume that HES actors’ behavior is far from free. On the one hand, individual resources and values limit behavior; on the other, behavior is shaped by the restraints, opportunities and incentives that policies establish to prevent harmful courses of action or to promote worthy ones. Our model primarily focuses on the conditioning side. This choice is justified by the consideration that actors’ resources and values are neither developed in an institutional void nor fixed. Instead, they are endogenous to policy interventions and subject to reshaping (Ingram and Schneider 1990).

Indeed, policy interventions can restructure the priority that individuals give to particular preferences by weighting some alternatives down with sanctions while facilitating others thanks to incentives. We recognize that the mechanics of preference reshaping can take different forms: abidance of authority, calculation of the costs and benefits of alternatives, or adherence to the underlying policy value. However, we also maintain that these forms are all equally functional to the expected result, that is, promoting the individual adoption of some consumption or production behavior that individuals would not otherwise adopt and that, at the aggregate level, define the performance of the overall system. What really can make a difference, therefore, is the nature and intensity of constraints that trigger and maintain performance. This allows us to regard the generative process of preference formation and

transformation into action as unobserved, and entailed by the expectation that it always occurs under effective inducements.

Preference shaping is therefore our invariant generative process; effective inducements are the 'efficient' modular explanations for system performance. However, what actually constitutes an effective inducement?

4.2. Efficient causes of system performance

More or less tacitly, all streams of the literature recognize that the efficient part of any governmental intervention consists of policy instruments. Scholars who have been engaged in their conceptualization have defined policy tools as, for instance, "an identifiable method through which collective action is structured to address a public problem" (Salamon 2002, 19); "a set of techniques by which governmental authorities wield their power in attempting to ensure support and effect or prevent social change" (Vedung 1998, 21); or the means for a government "to deliberately affect the nature, types, quantities and distribution of the goods and services provided in a society" (Howlett 2000, 415). Beneath their superficial differences, these definitions all entail the idea that instruments refer to the capacity that government has to 'get things done' even against individual preferences.

This idea is founded on a long tradition of mid-range theories concerning compliance. A seminal contribution comes from Etzioni's typology, according to which people can be made to comply by three strategies: a *coercive* one, exerted through physical obligations and restraints in the form of rules, directives and mandates backed by the threat of sanctions; a *remunerative* one, based on the allocation of economic benefits that reward special behavior or make it simply more convenient; and a *persuasive/normative* one, enacted through the manipulation of information and the allocation of symbols of prestige (Etzioni 1961). Brigham & Brown (1980) approached the issue as a problem of the success and failure of implementation and reduced the strategies to two basic alternatives: *penalties* (e.g., sanctions and taxes) and *incentives* (e.g., grants, exemptions and rewards). A concern for the degree of coerciveness embodied in each policy instrument subsequently led to an array of further typologies, which often arranged tools along a continuum from the more permissive to the more restraining pole (Hood 1983, Phidd and Doern 1983, Ingram and Schneider 1990, Howlett & Ramesh 1993).

Vedung (1998) further developed such theories and regarded the *nature* of instruments and their *coercive* capacity as separate, orthogonal dimensions. He thus sorted instruments by the basic motivational switch involved, whether economic ‘carrots’, such as *taxation* or *expenditure*²; *regulatory* ‘sticks’, which can be more or less soft depending on the type of control and sanction that back them; or *informational* ‘sermons’, which may range from neutral administrative communications on modes of accessing a service through reputational certification. Additionally, he identified different degrees of coercion, conceived as how free individuals are left to choose different alternatives than the ‘option’ that the switch turned on. Sermons correspond to the government’s minor interventions and the lowest restraint on people’s preferences when they take the shape of neutral administrative information. However, when they allocate prestige, they become more coercive and ‘structure’ options instead. Option structuration is also the typical functioning of economic carrots when used as rewards or inducements. Though, carrots can also be taken away from those who prefer the wrong behavior. Such deprivation equates to ‘biasing’ options, as people can still choose a different behavior if they can afford it. Option biasing is also operated by soft regulation as, once again, non-compliance is followed by partial restrictions only. When the stick becomes the command-and-control sort, however, it becomes a much tougher device that ‘forces’ the audience to adopt the single legitimate behavior, backed by severe sanctions.

Vedung’s typology therefore highlights a key point for our reasoning: causation is not merely a matter of the nature of an instrument (whatever it is defined as), or of its degree of coerciveness (the strict or lax conditioning of alternatives), because the two dimensions do not overlap. Rather, causation rests on a type of switch being given a special power to foster an option. This combination becomes observable in the form of ‘delivery vehicles’ (Salamon 2002), that is, as the *special shape that gives a switch the potential to foster an option once enforced*. Adopting actual vehicles as observable causes – instead of more general families or

² In our perspective, taxation can be thought of as an autonomous type of substantial instrument. We are perfectly aware that in other typologies, taxation is the chief component of larger types of 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 exemption, tax expenditure). We believe that expenditure and taxation have a different political economy and offer a different way to induce or restrain institutional and individual behavior in higher education. This means that expenditure and taxation design and induce different sets of potential preferences on the part of the recipients because they incentivize/constrain their behavior in significantly different ways.

types of tools – has the advantage of allowing a neater and more fine-grained operationalization. When vehicles are considered, the attention can narrow – from carrots to actual *taxation* vehicles, either as fee exemption or levying, and to *expenditure* vehicles, such as earmarked funds, lump-sum transfers, loans, and the like; from *regulatory sticks* to real *rules* laid down in laws, decrees, national or local regulation; and from *sermons* to the tangible *modes of circulation* of performance information – and grasp how each substantive tool actually works.

Vedung's typology, like many others, only focuses on 'substantive instruments', that is, on vehicles that directly affect the strategies of the recipients when implemented in the field. Their shape alone, however, is not enough to assume unimpeachable effectiveness. Effectiveness also depends on how credible taxes, spending, rules and information are deemed by recipients, that is, how certainly incentives (penalties) will follow when the fostered option has (not) been embraced. Credibility therefore is not a property of any vehicle as such. Instead, it depends on a further set of instruments that, following Howlett (2000, 2010; Salamon 2002), can be labeled as 'procedural'.

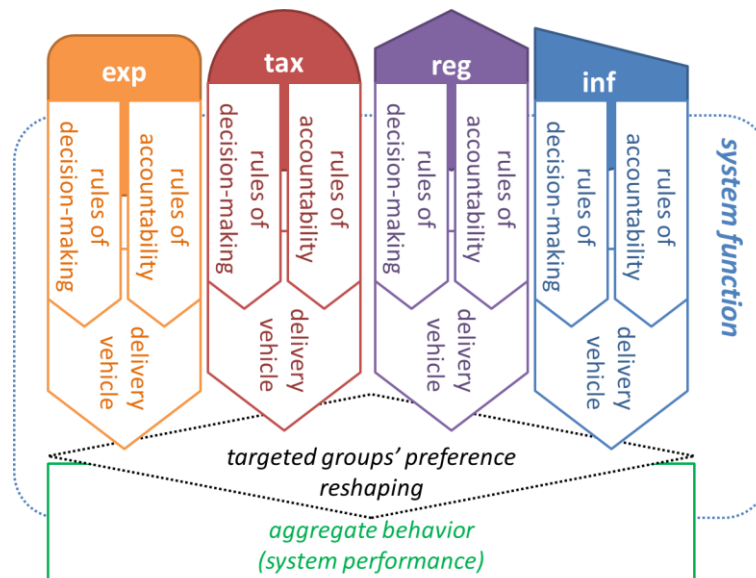
These special tools are rules, although they only impact recipients indirectly. They nevertheless actively contribute to the power of substantive instruments, as they define:

- a) who is responsible for vehicular settings ('decision-making') and
- b) to whom the policy actors have to be accountable and how they are held accountable for the uses they make of that vehicle ('accountability').

Indeed, depending on how the decision-making power is distributed to political, administrative or professional bodies across the system, vehicles may lose some of their credibility and, therefore, effectiveness. More important, weak or mismatching accountability provisions may lead the recipients to misuse or disregard some incentives or restraints (Hood 1986); when misuse is common, it eventually affects system performance. So, procedural tools further specify each vehicle's effectiveness.

In other words, our explanatory model accounts for the notion that each substantial delivery vehicle has its own procedural tools that jointly steer the behavior of targeted players within a functional area – as in figure 2 below.

Figure 2. Explanatory model of system performance by instruments



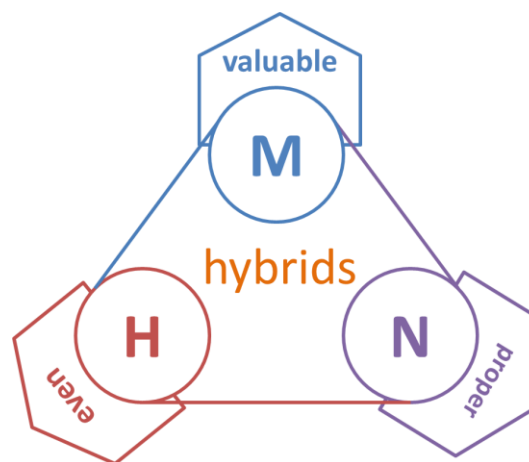
Procedural and substantial instruments therefore are justified efficient causes because they meet all the criteria of a sound explanatory model from a diversity-oriented perspective: they can shape an unobservable, invariant generative process by promoting some particular behavior through the reengineering of actors' preferences; also, they are modular, as their manipulation is expected to make the overall generative mechanism change in intensity, not in nature.

4.3. From single policy tools to governance regimes

According to our definition of good explanations, the power to initiate and maintain a generative process does not rest on individual vehicles or procedural instruments as such. Consistent with a diversity-oriented epistemology, causation is a configurational effect; thus, it springs from actual combinations of instruments' properties. This makes even more sense if we conceive of instrument mixes not as a mere combinatorial exercise but as 'regimes', the set of actual restraints that shape people's expectations, and strategies, and hence modes of coordination, while also embodying special values (Howlett 2012; Burch and Wood 1990; Krasner 1982). Thus, real governance regimes can be conceived as the actual *set of substantial and procedural instruments in use*, that is, *as the special mix of taxation, expenditure, regulation and information vehicles*, as well as *of those decision-making and accountability rules that govern them*.

Understanding instrument mixes in terms of governance regimes helps us to close the analytical space. Combinations of procedural and substantial instruments revolve around three well-known pure types of governance regimes, namely, hierarchy (H), network (N), and market (M) (Thompson *et al.* 1991). Thus we can expect that observable regimes either overlap with one of the three pure types, or form a hybrid of them (figure 3).

Figure 3. The analytical space defined by pure governance regimes



Keys: *M* = market; *H* = hierarchy; *N* = network;
circles = pure types; arrows = key property of intended results.

Governance regimes also allow expectations to be formed about performance. The three pure types constitute discrete alternatives because each empowers a different ‘principal’ – the national political majority (hierarchy); some functional, territorial, or professional community of reference (network); and individual ‘clients’ (market) – and serves distinct goals – respectively: evenness (through uniformity); appropriateness (through common identification and mutual obligations); and worth (through competition).

In higher education, pure hierarchic governance regimes can be expected to deliver more equitable results both from a territorial and social viewpoint, though irrespective of local specificities and at the price of excellence. Pure network governance regimes prize local and societal specificities; therefore, they should divide results along territorial or functional boundaries, sacrificing evenness without necessarily delivering excellence. Pure market competition is expected to deliver the most uneven results (above all, in terms of social

equity) while also disregarding the protection of local specificities, while allowing for cases of outstanding performance.

As real governance regimes are not monoliths but tool mixes, boundaries set by the three pure governance regimes also provide a clear guideline for the selection of relevant tool properties that can make an actual difference among regimes and, hence, performance. For instance, for substantial instruments in higher education, it can be important whether the vehicle:

- of state regulation is hard or soft, the former property being associated with hierarchy and the latter with the market and networks;
- of taxation (tuition fee) is income-based, implying a hierarchic rationale, targeted (network) or flat (market);
- of public expenditure is earmarked, implying a hierarchic rationale, or allocated through lump sum (network) or performance-driven means (market); and
- of information is top-down (hierarchic), horizontal and symmetric (network), or hidden in the tuition fees and thus asymmetric (market).

Equally, for procedural tools it becomes relevant that:

- decision-making is conducted at a proximate (institutional/community-based) vs. remote (state-driven) level to the intended beneficiaries; is concentrated vs. diffused; and is mainly shaped by political, administrative or academic bodies;
- accountability, again is concentrated vs. diffused; mainly due to political, administrative, or academic bodies; based on effectiveness vs. correctness.

In such a way, it becomes possible to operationalize the key properties of pure types – but at the level of observation lower than that where performance occurs, as mechanistic hypotheses require. Additionally, in a difference-oriented strategy of assessment, such disaggregation allows the researcher to consider the combination of each property with any other, therefore accommodating any possible hybrid of the three pure types – of which actual cases could become instances. In so doing, our framework can also test Krasner’s long-honored hypothesis that “if the principles, norms, rules, and decision-making procedures of a regime become less coherent, [...] then a regime has weakened” (1982, 189). Indeed, governance regimes so conceptualized can account for any national exceptionalism simply by regarding it as a measurable and comparable hybrid instrument mix in use.

5. CONCLUSIONS

The effectiveness of governance design in higher education has often been probed following the assumption that some special tool (e.g., performance funding, accountability, institutional autonomy, evaluation) could make the difference. There is no clear empirical evidence that this has been the case. A perspective focusing on net effects often leaves doubts on the actual effectiveness of selected tools in single cases. Variable-oriented analyses assesses whether a variable of interest has the power to influence the probability that the outcome improves as expected – on average, at the population level, and, often, regardless of interactions. On the basis of such results, prescriptions are drawn with ambitions of general applicability: however, the assessed predictive power merely gauges its average potential. As shown above, institutional autonomy can be more effective in some contexts than in others; the same amount of public funding can have different effects according to the other features of systemic governance, and so on. Actual causation is a conditional occurrence based on the working of some special mechanism after a conjunction of efficient causes. It therefore is in the interaction with ‘contexts’ that a predictor triggers or defuses causation, and actual policy performance follows. Yet, a perspective focused on contexts alone can dissolve the impact of policy tools in the complexity of ongoing processes and reciprocal influence among variables.

To solve this puzzle, we have proposed a specific framework based on a mechanistic hypothesis, fit for a diversity-oriented strategy, and focused on the performance of real HES governance regimes. According to this framework, a proper explanation of variations in performance across cases accounts for those combinations of factors that ‘made a difference’ between cases of policy success and failure. Here, differences among governance regimes in use are appreciated as differences in procedural and substantive tool mixes, which entail differences in the constraints and incentives affecting players’ preferences, therefore in their behavior, and ultimately in HES performance.

We consider that the same national governance regime does not necessarily produce homogeneous performance in the five functional areas of higher education. Thus, our theoretical framework accurately narrows on the performance of each HESs’ specific function. Also, our framework does not assume that any pure type of governance can perform better than any other, or that pure types outperform hybrids. Instead, we suppose that all substantive tool mixes can deliver results – although with different trade-offs among desired

outcomes. On the whole, we would rather expect that the difference between good and bad performers mainly rest on differences in the suitability of procedural instruments of decision-making and accountability.

These are, however, merely expectations on which empirical evidence alone can adjudicate.

REFERENCES

- 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.
- Bennett, A., & Checkel, J. (2014). *Process tracing: From metaphor to analytic tool*. Cambridge: Cambridge University Press.
- Bhaskar, R. A. (2008). *A realist theory of science*. London: Routledge.
- Brady, H., & Collier, D. (2010). *Rethinking Social Inquiry*. New York: Rowman and Littlefield.
- Braga, M., Checchi, D., & Meschi, E. (2013). Educational policies in a long-run perspective. *Economic Policy*, 28, 45–100.
- Brigham, J., & Brown, D. W. (1980). *Policy implementation: Penalties or incentives?* Beverly Hills: Sage Publications.
- Boudon, R. (1981). *The logic of social action*. London: Routledge & Kegan.
- Brunner, J. J., & Villalobos, C. (2014). *Políticas de Educación Superior en Iberoamérica 2009-2013*. Santiago: Ediciones Universidad Diego Portales
- Burch, M., & Wood, B. (1989). *Public policy in Britain*, 2nd ed. Oxford: Blackwell.
- Capano, G. (2013). Change in the governance of education policy in Australia and Canada. *Rivista Italiana di Politiche Pubbliche*, 12(3), 425–456.
- Capano, G. (2015). Federal strategies for changing governance of higher education: Australia, Canada, and Germany Compared. In G. Capano, M. Howlett & M. Ramesh (Eds.), *Varieties of Governance* (pp. 103–130). London: Palgrave Macmillan.
- Capano, G., & Piattoni, S. (2011). From Bologna to Lisbon: The political uses of the Lisbon ‘script’ in European higher education policy. *Journal of European Public Policy*, 4, 584–606.
- Capano, G., & Regini, M. (2014). Governance reforms and organizational dilemmas in European universities. *Comparative Education Review*, 56(1), 73–103.
- Capano, G., & Turri, M. (2015). *Same governance template but different policy autonomy. Types of evaluation agencies in higher education*. Mimeo.
- Clark, B. (1983). *The higher education system. Academic organization in cross national perspective*. Berkeley: University of California Press.
- Cartwright, N. (1997). ‘What is a causal structure?’. In V.R. Mckim & S. P. Turner (Eds.), *Causality in crisis?* (pp. 342–358). Notre Dame: Notre Dame University Press.
- Cartwright, N. (2007). *Hunting causes and using them: Approaches in philosophy and economics*. Cambridge, U.K.: Cambridge University Press.

- CHEPS (2015). *Performance-based funding and performance agreements in fourteen higher education systems*. Twente: Cheps.
- Coleman, J. S. (1990). *Foundations of social theory*. Cambridge: Harvard University Press.
- Craver, C. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Toronto: University of Toronto Press.
- Damonte, A. (2014). Policy tools for green growth in the EU15: A Qualitative comparative analysis. *Environmental Politics*, 23(1), 18–40.
- Dill D. D., & Beerkens, M. (2010). *Public policies for academic quality: Analyses of innovative policy instruments*. Springer: Dordrecht.
- Dobbins, M & Knill, C. (2014). *Higher education governance and policy change in Western Europe*. London: Palgrave.
- Dougherty, K.J. et al. (2014). Performance funding for higher education: forms, origins, impacts, and futures. *The Annals of The American Academy of Political and Social Science*, 655, 163-184.
- Donina D., Meoli M., & Paleari S. (2015). Higher education reform in Italy: Tightening regulation instead of steering at a distance, higher education policy. *Higher Education Policy*, 28(2), 215–234.
- El-Khawas, E. (2005). The push for accountability: Policy influences and actors in American higher education. In A. Gornitzka, M. Kogan & A. Amaral (Eds.), *Reform and change in higher education* (pp. 287–303). Dordrecht: Springer.
- Elster, J. (1989). *Nuts and bolts for the social sciences*. Oxford: Oxford University Press.
- Enders, J., De Boer, H., & Weyer, E. (2013). Regulatory autonomy and performance: The reform of higher education re-visited. *Higher Education*, 65(1), 5–23.
- Etzioni, A. (1961). *Comparative analysis of complex organizations*. New York: The Free Press.
- European Commission (2006). Delivering on the modernization Agenda for Universities: Education, Research Innovation (COMM 208), Brussels, 10 May.
- European Commission (2014). *Education and Training Monitor 2014*. Brussels: European Commission.
- European University Association (EUA) (2011). *University autonomy in Europe II: The scorecard*. Brussels: EUA.
- European University Association (EUA) (2014). *European Funding Observatory 2014*. Brussels: EUA.
- Fagan, M. B. (2012). The joint account of mechanistic explanation. *Philosophy of Science* 79(4), 448–472.
- Frolich, N., Schmidt, E., & 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.
- George, A., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press.
- Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, 44(1), 49–71.
- Gornitzka, A., Kogan, M., & Amaral A. (2005). *Reform and change in higher education*. Dordrecht: Springer.
- Hall, P. A. (2006). Systematic process analysis: When and how to use it. *European Management Review*, 3, 24–31.
- Hedström, P., & Ylikoski, P. (2010). Causal mechanism in the social sciences, *Annual Review of Sociology*, 36, 49–67.
- Hood, C. (1983). *The tools of government*. Chatham: Chatham House Publishers.

- Hood, C. (1986). *Administrative analysis: An introduction to rules, enforcement, and organizations*. Brighton: Wheatsheaf.
- Howlett, M. (2000). Managing the “hollow state”: Procedural policy instruments and modern governance. *Canadian Public Administration*, 43(4), 412–431.
- Howlett, M. (2011). *Designing public policies: Principles and instruments*. London: Routledge.
- Howlett, M. and Ramesh, M. (1993), Patterns of policy instrument choice: Policy styles, Policy learning and the Privatization Experience. *Review of Policy Research*, 12(1-2), 3–24.
- Huisman, J. (2009). *International perspectives on the governance of higher education*. London: Routledge.
- Johnstone, D. B. (2006). *Financing higher education: Cost-sharing in international perspective*. Boston: Boston College Center for International Higher Education, and Rotterdam: Sense Publishers.
- Kickert, W. (1997). Public governance in the Netherlands: An alternative to anglo-american managerialism. *Public Administration*, 75(4), 731–752.
- Krasner, S. D. (1982). Structural causes and regime consequences: Regimes as intervening variables. *International Organization*, 36(02), 185–205.
- Lane, J. E. & Kivisto, J. A. (2008). Interests, information, and incentives in higher education: A review of principal-agent theory in higher education governance. In C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 141–180). New York: Springer.
- Lazzaretti, L., & Tavoletti, E. (2006). Governance shifts in higher education: A cross national comparison. *European Educational Research Journal*, 5(1), 18–36.
- Lewis, D. (2001). Truth making and difference making. *Noûs*, 35(4), 602–615.
- Maassen, P., & Olsen, J. (2007). *University dynamics and European integration*. Dordrecht: Springer.
- Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly*, 2(4), 245–264.
- Mahoney, J. (2008). Toward a unified theory of causality. *Comparative Political Studies*, 41(4-5), 412–437.
- Mahoney, J. (2012). The logic of process tracing tests in the social sciences. *Sociological Methods & Research*, 41(4), 570–597.
- Mahoney, J., & Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis*, 14(3), 227–249.
- McLendon, M. K., & Hearn, J. C. (2009). Viewing recent U.S. governance reform whole: “Decentralization” in a distinctive context. In J. Huisman (Ed.), *International perspectives on the governance of higher education* (pp. 161–181). New York: Routledge.
- Mok, K. H. (2010). *The search for new governance of higher education in Asia*. New York: Palgrave Macmillan Press.
- Neave, Guy R., & van Vught, F. (1991). *Prometheus bound: The changing relationship between government and higher education in Western Europe*. Oxford: Pergamon.
- OECD (2008). *Tertiary education for the knowledge society: VOLUME 1 : Special features: Governance, funding, quality*. Paris: Oecd
- OECD (2014). *Education at a glance*. Paris: Oecd.
- Orr, D., Wespel, J., & Usher, A. (2014). Do changes in cost-sharing have an impact on the behaviour of students and higher education institutions? Evidence from nine countries. Luxembourg: Publications Office of the European Union.
- Paradeise, C. et al. (2009) *University governance. Western European comparative perspectives*. Dordrecht: Springer.
- Pawson, R. (1989). *A measure for measures: A manifesto for empirical sociology*. London: Routledge.

- Pawson, R. (2008). Causality for beginners. *NCRM Research Methods Festival*, eprints.ncrm.ac.uk/245/1/Causality_for_Beginners_Dec_07.doc
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Francisco: Morgan Kaufmann.
- Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge: Cambridge University Press.
- Phidd, R., & Doern, G. B. (1983). *Canadian public policy: Ideas, structure, process*. Toronto: Methuen.
- Pick, D. (2006). The Re-Framing of Australian Higher Education. *Higher Education Quarterly*, 60(2), 229–241.
- 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. (1997). Turning the tables. *Comparative Social Research*, 16, 27–42.
- Ragin, C. C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.
- Ragin, C. C. (2008). *Redesigning social inquiry. Fuzzy sets and beyond*. Chicago: University of Chicago Press.
- Rosa, G., & Amaral, G. (2014). *Quality assurance in higher education*. Basingstoke: Palgrave Macmillan.
- Richardson, R. C., & Martinez, M. C. (2009). *Policy and performance in American higher education: An examination of cases across state systems*. Baltimore: The John Hopkins University Press.
- Rutherford, A., & 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, (185–209).
- Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- Salamon, L. M. (2002). *The tools of government: A guide to the new governance*. Oxford: Oxford University Press.
- Schneider, A., & Ingram, H. (1990). Behavioral assumptions of policy tools. *The Journal of Politics*, 52(02), 510–529.
- Schneider, C. Q., & Wagemann, C. (2012). *Set-theoretic methods for the social sciences*. Cambridge: Cambridge University Press.
- Schuetze, H. G., Bruneau, W. & Grosjean, G. (2012) *University governance and reform – Policy, fads, and experience in international perspective*. New York: Palgrave Macmillan.
- Shadish, W. R., Cook, T. D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton.
- Shanahan, T., & Jones, G. A. (2009). Shifting roles and approaches: Government coordination of post-secondary education in Canada, 1995-2006. In M. Tight, K. H. Mok, J. Huisman, & C C. Morphew (Eds.), *The Routledge International Handbook of Higher Education* (pp. 315–325). New York: Routledge.
- Shattock, M. L. (2014). *International trends in university governance*. London: Routledge.
- Tandberg, D., & Hillman, N. (2014). State higher education performance funding: Data, outcomes, and policy implications. *Journal of Education Finance*, 39(3), 222–243.
- Thompson, G., Frances, J., Levacic, R., & Mitchell, J. (1991). *Markets, hierarchies and networks*. London: Sage.
- Trakman, L. (2008). Modelling university governance. *Higher Education Quarterly*, 62 (1–12), 63–83.

- Usher, A. (2005). *A little knowledge is. A dangerous thing: How perceptions of costs and benefits. Affect access to education*. Toronto: Education Policy Institute.
- Van Vught, F. (1989). *Governmental strategies and innovation in higher education*, London: Jessica Kingsley.
- Vedung, E. (1998). Policy instruments: Typologies and theories. In M. L. Bemelmans-Videc, R. C. Rist, & E. Vedung (Eds.) *Carrots, sticks, and sermons: Policy instruments and their evaluation*, (pp. 21-58). Transaction.
- Viaene, J. M. A. & Zilcha, I. (2013). Public funding of higher education. *Journal of Public Economics*, 108, 78–89.
- Volkwein, J., & Tandberg, D. (2008). Measuring up: Examining the connections among state structural characteristics, regulatory practices, and performance. *Research in Higher Education*, 49(2), 180–197.
- Winter-Ebmer, R., & Wirz, A. (2002). *Public funding and enrolment into higher education in Europe*. IZA Discussion Papers 503, Institute for the Study of Labor (IZA).
- Woodward, J. (1997), Explanation, invariance and intervention, *Philosophy of Science*, 64, 26–41.
- Woodward, J. (2000), Explanation and invariance in the special sciences, *The British Journal for the Philosophy of Science*, 51, 197–254.
- World Bank (2008). *Global trends in university governance*. Washington: World Bank.
- Wright, S., & Orberg, J. (2008). Autonomy and control: Danish university reform in the context of modern government. *Learning and Teaching*, 1(1), 27–57.