# Università degli Studi di Napoli "L'Orientale"



Dipartimento di Scienze Umane e Sociali

## Dottorato di Ricerca in Studi internazionali

## XXXIII CICLO

Settore Scientifico Disciplinare: SPS/04

## Evaluated Through Performance or Performing Evaluation?

An historical investigation of the role of vice-chancellor committees in shaping the genesis of managerial governance in English and Italian universities

Candidata: Laura Giovinazzi

Tutor: Prof. Rosario Sommella

Co-Tutor: Prof. Domenico Rizzo

Coordinatore: Ch.mo Prof. Raffaele Nocera

ANNO ACCADEMICO 2019-2020

# Abstract

The New Public Management literature represents an intersection of different disciplinary traditions and importantly explores how management ideas and practices play a part in governing public institutions through evaluation criteria. However, the debate on higher education focuses on State-level developments, whilst being too general when analysing university contexts. This tendency hinders a proper understanding of how management and evaluation practices first emerged and developed within universities. To remedy this gap, this study advances an historical investigation of the first uses of such practices within the English and the Italian university sectors.

I argue that whilst scholars acknowledge that vice-chancellors have acquired managerial roles, they tend to limit their conceptualisation of these agents to mere subjects of governance as NPM principles emerged. As a result, they underplay the role of groups who consistently developed and launched evaluation practices within the university context. From this critique, the thesis examines the rise of managerial governance in universities by historicising vice-chancellor committees' use of evaluation practices. In so doing, the study critically analyses committees' actions in relation to States' policies. It also draws out the gradual and bottom-up construction of an evaluation-managerial nexus.

The study's key historical argument is that performance assessment indicators in English and Italian universities did not emerge from State policies geared at creating competition between universities but entered university walls due to the active role of vice-chancellor committees. Far from being mere marketizing instruments imbued with a competitive logic, performance assessment indicators are extremely contested technologies which reflect constantly changing power struggles between university groups and between university and non-university groups. The educational evaluation indicators analysed here were first constructed by research networks in the post-war international context, in the midst of a "planning movement" tracing back to the 1960s at the OECD. As vice-chancellors borrowed these practices, they frequently re-articulated them for acquiring greater leverage of power within universities, developing a managerial connotation of evaluation. Explaining the rise of managerial governance through evaluation in terms of universities' forced adaptation to the pressures of the New Public Management paradigm, I argue, generalizes the complex and still largely unexplored historical lineages of these practices.

# Acknowledgments

# List of Abbreviations

AAU – Academic Audit Unit

AHRC – Arts and Humanities Research Council

BBSRC – Biotechnology and Biological Sciences Research Council

CEDE – Centre Européen pour le Développement de Entrepreneuriat

CENSIS – Centre for Social Investment Studies ("Centro Studi Investimenti Sociali")

CERI – Centre for Educational Research and Innovation

CIPP model – Context, Input, Process and Product

CIVR – Research evaluation committee ("Comitato di Indirizzo per la Valutazione della Ricerca")

CPRS – Central Policy Review Staff

CRE – Conference of European Rectors

CRUI – Conference of Italian University Rectors ("Conferenza dei Rettori delle Università Italiane")

CSD – Civil Service Department

CSTP – Committee for Scientific and Technical Personnel

CNR – National Research Council (Consiglio Nazionale delle Ricerche)

CUA – Conference of University Administrators

CVCP – Committee of Vice Chancellors and Principals

DfEE – Department for Education and Employment

DES – Department of Education and Science

DQA – Division for Quality Audit Policy

EIP – Educational Investment and Planning

EPA – European Productivity Agency

EPSRC – Engineering and Physical Sciences Research Council

EQUIP project – Enhancing Quality through Innovative Policy & Practice (carried out by the European University Association)

ESRC – Economic and Social Research Council

EUA – European Universities Association

FFO – Fondo di Finanziamento Ordinario

FMI – Financial Management Initiative

FMU – Financial Management Unit

HEFCE – Higher Education Funding Council for England

HEI – Higher Education Institutions

HEQC – Higher Education Quality Council

IIEP – International Institute for Educational Planning

IEUs – Internal Evaluation Units

IMHE – Programme on Institutional Management in Higher Education

JPIWG – Joint Performance Indicators Working Group

KEF – Knowledge exchange framework

MIMIR – Modernisation of Institutional Management of Innovation and Research in South Neighbouring countries project (carried out by the European University Association)

MIS – Management Information Systems

MRC – Medical Research Council

MRP – Mediterranean Regional Project

MUR – Ministero dell'Università e della Ricerca

NERC – Natural Environment Research Council

OSTP – Office for Scientific and Technical Personnel

PA – Public Administration

PAC – Public Accounts Committee

PAR – Program Analysis and Review

PARC – Program Analysis and Review Committee

PESC – Public Expenditure Survey Committee

PI – Performance Indicators

PPBS – Planning Programming Budgeting System

PSRU – Public Sector Research Unit

QAA – Quality Assurance Agency for Higher Education

QMS – Quality Management Systems

RAE – Research Assessment Exercise

RAND Corporation –Research ANd Development Corporation

RE – Research England

RSE – Research Selectivity Exercise

SECIN – Internal Evaluation Units ("Sezioni di Controllo Interno")

SRHE – Society for Research in Higher Education

STEM subjects – Science, Technology, Engineering and Mathematics

STFC – Science and Technology Facilities Council

SUP – Sustaining University Program

SVIMEZ - Association for the Industrial Development in Southern Italy ("Associazione per lo sviluppo dell'industria del Mezzogiorno")

TCPS – Technical Committee for Public Spending ("Commissione Tecnica per la Spesa Pubblica")

UFC – Universities Funding Council

UGC – University Grants Committee

UUK – Universities UK

UKRI – United Kingdom Research and Innovation

VNSU – Association of Universities in the Netherlands ("Vereniging van Universiteiten")

VPS – Evaluation of Scientific Production ("Valutazione della Produzione Scientifica")

# List of Tables

# List of Figures

# Appendix List

A. The 'University constituencies'
B. Presidents, Vice-Presidents, and General Secretaries of the CRUI 1963 - 2002
C. Life cycle of an educational activity
D. Members of the 'Steering Committee for Efficiency Studies in Universities)
E. Membership of the CVCP/UGC 'Performance Indicators Technical Committee' in 1986
F. Membership of the CVCP/UGC 'Performance Indicators Steering Committee' in 1987
G. University Management Statistics and Performance Indicators
H. List of the 37 cost-centres in British universities
I. Membership of the 'Research sub-group of the Joint Performance Indicators Working Group' in 1993
J. The Composition of the Fund for Ordinary Funding in 1998
K. Members of the CRUI's 1991 evaluation committee
L. List of the 17 Macro-scientific disciplinary sectors following Ministerial Decree n.152 23/06/1997
M. List of the 14 Scientific disciplinary Areas following Ministerial Decree 4/10/2000
N. Members of the CRUI's research quality indicators working group in 1999
O. Research quality indicators devised by CRUI in 1999
P. Staff weightings of CRUI's research quality indicators
Q. Comparison between the CRUI's weightings for industrial engineering and economics

# Table of contents

# Introduction

Universities have lived a profound managerial shift over the last 50 years, characterised by two key developments: the widespread use of evaluation criteria and the marginalisation of academic powers. As the funding and organization of activities have become subjected to notions of efficiency, economy and effectiveness, the principle of collegial governance has been increasingly side-lined by that of managerial governance. As management notions and practices continue to expand, they deeply redefine the university's existentially social demand function towards an economic one.

This complex period of transformation is referred to as one of the typical examples of the New Public Management (NPM) paradigm, in which evaluation is a key instrument allowing States to steer public sector activities into productivity (Mellon, 1993; Hood, 1991; Ferlie et al., 1996). The benefits brought about by NPM governance include, amongst others, significant savings in public resources, now channeled conditionally on the basis of institutional efficiency rather than being distributed incrementally in the public sector. Furthermore, the use of evaluation is allowing States to retreat from previous burdens associated with the full control over public sectors. This is because resource management and distribution powers have been devolved to the institutional level and/or to private-sector third parties whilst being monitored at a State level through frequent performance assessment. As Neave put it, thanks to evaluation States can '[…] take refuge in the commanding heights of strategic management' (1988, p.12) and thus exert political influence by setting and monitoring objectives without being involved in administering daily activities for their achievement. The concept of "steering at a distance" thus appears well-fitting for understanding managerial developments occurring in the university sector, where the funding of research and teaching activities is subjected to the efficiency-oriented evaluation criteria developed by quasi-governmental university evaluation agencies.

This perspective helps to understand how States have used evaluation to redefine their powers, providing important insights also for critically assessing the ongoing dismantlement of the welfare State model. However, it marginalises in-depth analyses of institutional changes occurring within NPM. Indeed, capturing how the new "regulatory State" (Palfreyman & Tapper, 2014) engages in  steering at a distance tells us little about how governance has turned managerial at the university level. Whilst new modalities of exerting State power are bound to influence institutional development, they can hardly be said to fully shape it. This is even more the case for universities, who hold statutory autonomy and thus have an important degree of leverage over internal

governance.[1] A fully rounded understanding of the managerial turn in universities must thus keep account of how both State power and institutional power have transformed since the advent of NPM-oriented evaluation.

Studies of NPM in universities agree on the argument that institutional powers changed in line with State powers and turned managerial following NPM-inspired reforms to higher education (Palfreyman & Tapper, 2014). This shared perspective rotates around the conviction that efficiency-oriented practices carried an unstoppable transformative power given their ability to respond to the growing costs of administering public sectors following their expansion between the 1960s and the 1970s. Management's ideological appeal and its rise as a new dominant "cultural archetype" (Deem et al., 2007) is often identified as that key force driving institutions' alignment with State reforms. The suggestion emerging from these readings is thus that universities – just as other public institutions – were inevitably subjected to the strength of management practices which were becoming increasingly appealing to States. The result is that scholars frame the rise of university chief executives using evaluation to govern universities **as evidence of an institutional adaptation to State pressures** (Deem et al., 2007; Palfreyman & Tapper, 2014;McGettigan,2013) **and to NPM more broadly.**

This top-down framing of universities' managerial transformation, however, clashes greatly with context-specific studies on the role that university networks play in shaping higher education governance (Cave et al.,1997; Rizzi & Silvestri, 2002;Vaira, 2011; Kogan & Hanney, 2000; Paradeise et al., 2009;). Kogan and Hanney, for example, suggest that the complex developments occurring in higher education hardly fit into a '[…] coherent philosophy of government' (2000, p.238). In showing how university agents importantly shape institutional alignment or resistance to NPM policies (Paradeise et al., 2009), studies have problematized NPM scholarship's inability to account for university networks' influence on the sector's managerial development. One of such networks, that of VC committees, stands out from the university world. Internationally, the Conference of European rectors has been a key group in developing higher educational policymaking. It has forged crucial features of the current European university context, such as the ERASMUS program, a landmark in European Higher Education policy. Following its development into the European Universities Association (EUA), further, the Conference has been consulting universities on evaluation criteria, setting out strategic management guidelines (see for example its "Institutional evaluation program"[2] ) which are used and shared by adhering universities. In specific national contexts, studies have shown how vice-chancellors actively participated in the construction of evaluation criteria in England (Cave et al., 1997) and Italy (Rizzi & Silvestri, 2002) when

---

[1] Italian universities' autonomy is granted to them by the constitution, a provision passed following the dramatic effects of the fascist regime (see Art.33 of the Italian Constitution).
[2] https://www.iep-qaa.org/downloads/publications/20year_contribution_to_institutional_change.pdf

efficiency principles began to diffuse. In appropriating notions of evaluation and management, VC committees have long been developing specific articulations and practices around efficiency which were often readily used in universities and yet remain largely unexplored by mainstream studies.

The co-existence of multiple groups developing specific articulations of university efficiency within the NPM paradigm, I argue, compels scholars to analyse the rise of evaluation as a *political process*. This means analysing how different groups of State and non-State agents have advanced their own understanding of efficiency as well as how they interacted, rather than assuming that university agents merely aligned to State demands. This can bring to the foreground new insights for unveiling how and why university evaluation is developing into an instrument of managerial governance *at both the State and institutional levels*. The concern of this thesis, however, is that NPM studies on higher education do not permit to make this step for they are exceptionally a-political compared to research on how NPM is playing out in other sectors. In so doing, they reproduce the view that the university context is inherently subjugated to State policy and must be studied as such. If studies on the managerial turn in the public administration have pointed to the existence and persistence of political struggles between civil servants and politicians (Lewis, 2011; Rebora et al., 2016), those on higher education define it as a process of forced adaptation to external pressures (Deem et al., 2007). As this research discusses, the tendency to conceptualise the public administration as a more political realm than other public institutions rests on a structuralist understanding of power, whereby political struggles are conceptualised as those contestations occurring within political institutions, rather than as contestations between groups of agents. As I argue, this translates into the assumption that institutional agents' actions are inherently *re-active* to those of political agents, with the result that a historical reconstruction of their role in shaping broad developments such as NPM is side-lined.

The key focus of this thesis for re-politicising the rise of managerial governance in universities is the role played by Vice-Chancellors in (VCs from now on) in shaping key NPM principles such as performance evaluation, institutional management, and accountability. This analysis is carried out for two specific university sectors, namely the English and the Italian; and their respective Vice-Chancellor committees, the Committee of Vice-Chancellors and Principals (the CVCP, currently Universities UK, UUK) and the Committee of Rectors of Italian Universities (Conferenza dei Rettori delle Università Italiane, CRUI). The choice of these cases relates to the debate (Clark, 1983, Capano,2008) on the relationship between different configurations of State power and types of university governance, which importantly inform studies on NPM in universities. The two sectors are conventionally analysed through the categories of Clark's triangle, pioneered by Bruton Clark in 1983. An agency-centred (Knafo, 2010) comparative analysis of two opposites in Clark's triangle – the Italian being an example of the "continental/academic oligarchy"

model and the British of the "collegial /market" model – will allow to problematize how this approach feeds into structuralist accounts of managerial governance in universities. Surely, this approach has an appeal, for there is no question that the universities' surrounding political-economic context influences their development. Clark's model allows to outline key differences between countries' university sectors, providing a starting point for reflecting on how the broader institutional setup within which universities operate may predispose certain types of relationships between universities and their "external environment" (Clark, 1983), such as their relationship with the State. Whilst being more specific to the university world than NPM scholarship, however, Clark's model also falls in NPM scholars' trap of analysing university governance through non-university categories and agents. Changes in institutional governance are treated as the residual effect of changes to superior powers such as the State and the market, often interacting in a trade-off dynamic. The role that university agents concretely play in defining institutional governance within a given "external environment" and within given State-market configurations is marginalised to a secondary concern for understanding how universities transform. Eventually, State or market constraints are seen as fully shaping institutional governance. This research hence takes two contrasting cases in Clark's model to explore how vice-chancellors' actions in light of State policies have shaped two very distinct forms of institutional governance, recalibrating assumptions on the structural determinants of these differences. This allows to study the two sectors' managerial turn dynamically and contextually,  exploring vice-chancellors' changing actions in relation to changes in both the university and the political-economic context: studying bottom-up developments in relation to top-down ones.

This change of perspective will allow to appreciate the key contribution of this research.  As I argue, managerial practices around evaluation entered the two university sectors through a complex and context specific interplay of State and vice-chancellor powers. These notions historically emerged much earlier than neoliberal governance. Far from having been transmitted by NPM-inspired State policies, educational evaluation practices were first theorised and piloted in the 1960s by research networks working in emerging post-war international organisations, particularly the OECD. Although scarcely analysed in relation to the rise managerial governance in universities, this historical lineage left a significant legacy amongst participant university administrators and vice-chancellors, who confronted States' NPM policies with much greater knowledge and active participation than is often assumed and documented in scholarly literature. As Vice-Chancellors borrowed practices and notions developed there, they re-articulated their use for acquiring greater leverage of power within universities, further developing a managerial connotation of evaluation. Explaining the managerial shift in universities as a case of forced adaptation to State pressures generalises the complex historical lineages of university evaluation and what they tell us about the

rise of managerial governance. To expand this argument and present its content, I therefore propose to significantly revisit existing analyses of universities' managerial shift along different historical and theoretical axis to those of NPM scholars. Historical enquiries often inevitably stem from the theoretical assumptions guiding them, with the result that history conventionally becomes a validation of specific theoretical approaches. Through a historicist approach, however, this research overturns this tendency; seeking to let the historical developments draw the perimeters of the theoretical categories that I mobilize to account for the rise of managerial governance within universities. From this perspective, my critique of existing theoretical categories used to explain this phenomenon relies crucially on the contradictions which emerge when historical developments suddenly appear counterintuitive to theoretical predictions. As I argue, this reflects the gaps of deductive approaches, which this research tries to rebalance through historicism. For this reason, both of the new axis proposed in this thesis for revisiting the rise of managerial governance in universities are theoretical critiques informed by historical research.

## Studying managerial governance in its own terms

By studying managerial governance in its own terms, we can understand how management has turned into a decision-making exercise instead of assuming that this generally occurred in the 1980s when management practices were appropriated by politics. At a theoretical level, I argue that analyses on the development and consequent rise of managerial powers are distorted by structuralist accounts of power. The view that managers' agency in institutions is an effect of State agency is thus a cause and a reinforcement of scarce historical research into the diverse managerial agents that make use of management practices within the NPM paradigm. Borrowing from existing critical Political economy studies (Knafo et al., 2019, Knafo, 2020) on managerialism, I suggest that this trend is strictly related to the scholarly conviction that managerialism reflects the progression of neoliberal governance, with the result that there is scarce historical attention on the concrete ways and groups through which management practices have been developed into instruments of decisional power by managerial groups themselves. By prioritising the use of neoliberal categories for explaining a managerial phenomenon, the emergence of NPM and its ongoing development is readily and consistently reconducted to a question of how States act in order to reinforce neoliberal governance and market principles (Knafo et al., 2019). As has been critically argued (Bacevic, 2019) the way that this tendency plays out in academic critiques of the neoliberal university is actually weakening scholars' attempt to use academia as an instrument of resistance. They seem to be leaving a critical examination of the grounds of their working conditions too unexplored. Convinced that knowing neoliberalism is sufficient to change it (Bacevic, 2019), academics continue to attack neoliberal logics for the dismantlement of the sector whilst witnessing on their

own skin the outbreak of an increasingly extending managerial arm. In the UK, there have recently been significant academic strikes and mobilisations across the country, yet the VC Committee (now UUK) has managed to actually reinforce its position.[3] In the case of evaluation, scholars emphasize its neoliberal logic without debunking how the managerial governance of quality criteria may be reinforcing itself precisely thanks to the fact that these are seen the consequence of overarching neoliberalism and not as the result of the active work of managerial groups. By speaking out against neoliberalism's systemic power, in other words, academic critique does managers a favour. The result, I argue, is that there is still a marginal analysis of universities' managerial agents and practices *in their own terms*, which would allow to trace how these acquired new powers over institutions through management. On one hand, the rise of the "quality movement" (Dahler-Larsen,2019) has fuelled discussions on technocratic élites  and those processes through which States centralize power through biopolitical discipline (Pinto,2012). On the other, the proliferation of quality benchmarks is seen as the emulation of market dynamics (McGettigan, 2013), functional to neoliberal States' retreat. In both cases, scholars' object of analysis is the neoliberal State, and not the managerial cadres in university sectors. Conceptualising the sector's dismantlement as a phenomenon caused by exogenous pressure only confirms that academics are too concerned with 'knowing neoliberalism', and this deepens existing challenges of unpacking how neoliberal policies (such as the sector's privatisation) are developing not in isolation, but instead simultaneously to the diffusion and empowerment of managerial groups *within* universities. Considering that neoliberalism is only the tip of universities' managerial iceberg can help us to unpack how planning, evaluation and conditional funding find legitimation at the higher hierarchies and are instead often outright rejected by the lower-level and precarious workers in universities, including subjects within academic governing bodies. understanding the neoliberal reasons for using evaluation to govern the sector, in other words, still leaves the managerial dynamics unexplored, instead it is important that they be studied in their specific developments. This is key, for if on one hand the use of managerial practices is strengthening neoliberal governance, ***it is also empowering managerial control over organizations' procedures, goals, and constraints[4]***.

The first key developments in managerial practices  used for decision-making, in fact, developed from a military approach called "systems analysis" emerging in the post-war period. Different to scientific management practices, performance assessment ones were focused on

---

[3] As the University and college union (UCU) stated on the development of the strikes, '[…]branches need to make clear to employers that the cuts to defined benefits proposed by UUK are absolutely unacceptable and that UCU members are ready to campaign and take industrial action to force employers to withdraw them' (UCU, 19 April 2021) See more at https://www.ucu.org.uk/media/11469/USS-branch-briefing-Apr-21/pdf/USS_branch_briefing_19Apr21.pdf

[4] This applies also to the exponential growth in the number and diversity of academic consultancy agencies. Consultancy agencies exist for the whole spectrum of university activities, including  research management, research evaluation, quality assessment.

strategic decision making, a concept initially used in military strategy in order to take the best operative decision amongst multiple hypotheses (Knafo et al., 2019). The historical development of systems analysis challenges widespread assumptions on managerialism as the product of new public management in the 1980s, mostly because it shows that criteria for quantifying and achieving efficiency, performance and control **were being used much earlier**. A key example is the Eisenhower administration's use of the Planning Programming Budgeting System (PPBS) in education departments to steer the sector into higher scientific productivity, using concepts of objective-setting, performance assessment and efficiency to take an important policy decision, namely the National Defense and Education Act (passed in 1958). Whilst in the US the PPBS framework was used to tailor policy decisions during the late 1950s, its diffusion into international networks led to the development of new methods and concepts around educational governance in output-based terms (Bürgi, 2016a, 2017b). Exploring how educational evaluation indicators came about, and how this movement intersected -and still intersects- with broader developments in managerial theory is crucially relevant for politicising the debate around the use evaluation as an instrument of governance. Considering the broad lineage of management practices around performance criteria opens research paths in the field of educational evaluation criteria specifically.

### Studying the international lineage of educational evaluation practices

Due to the convinced assumption that States appropriated efficiency-oriented evaluation criteria from the corporate sector in the 1980s, scholars suggest that educational evaluation criteria were managerial from the onset, in other words the word 'imported' is often used to describe the transfer of criteria from the private to the public sector, suggesting that criteria had already been devised by other groups. Many see the beginning of the use of evaluation as a process where university agents gradual adapted to pre-determined criteria, frameworks, and principles, as if these entered the university walls in the form of a package ready to be put to use. However, this does not explain two key aspects of university evaluation criteria which depend on their international lineage. Firstly, the fact that, as this thesis shows, **university-tailored evaluation criteria emerged previously to the establishment of national evaluation agencies**. Secondly, that **there are many different types of evaluation criteria,** often depending on different university contexts. The first university criteria used in Italy and England, for example, borrowed from the frameworks of the CIPP and the PPBS models. To explain these points, this research critically addresses the assumption that university evaluation criteria were inherently managerial from their genesis. I suggest that *t*he management-educational evaluation nexus was a historical development stemming from innovations in military strategy design methods at the U.S. department of defense** (Bürgi, 2016a; Hirsch, 1965). As such, I seek to historically trace how, why and through which groups the military lineage

influenced the redefinition of university evaluation from peer-review - an academically controlled and highly informal system rather autonomous from political-economic interference - to a system of managerial control - dominated by conditionality logics, metrics and guidelines created by a galaxy of experts. I trace the transmission of the military lineage to university contexts in order to understand which agents constructed the relationship between management and evaluation in time.

As the research discusses, both the program budgeting framework and the CIPP model found fertile grounds in international networks of educational planners and researchers active at the OECD and UNESCO between the 1950s and the 1970s. This emerging class advocated for a technical and rationalising approach to educational governance, for example by framing changes such as student enrolment increases in terms of efficiency gains and cost control (Coombs, 1968). Populated by university researchers, administrators and vice-chancellors, international research centres such as the Centre for Educational Research and Innovation (CERI) and the International Institute for Educational Planning (IIEP) developed new educational indicators and institutional management models which were tested through 'pilots' in member States' respective education sectors.. Interest groups were thus able to use international research to test their views and studies on educational productivity. This set-off a process of constant experimentation – especially between the 1960s and the 1980s in Western Europe – where schools and universities were laboratories for educationalists, economists, and planners, all keen to influence how the post-war educational expansion would play out (Bürgi,2016a,2016b). Acting as a 'platform' (Centeno, 2019), international organizations were key *loci* where diverse interests met and could potentially find legitimation by interacting. These networks did not always depend on political support or specific policy reforms for a time being, because research itself was enough to justify why new experiments in educational evaluation were being discussed and developed. As researchers debated different educational contexts, for example, they formed new projects and collaborations to develop their research and enhance comparative analyses, a process which crucially explains the proliferation of educational indicators internationally. Diverse research projects at the OECD and UNESCO reflected in this way diverse networks of agents working on educational development.

By tracing the work of a military-academic network of system analysts, this research reconstructs how PPBS was framed as a management model for university administrators and leaders. Similarly, researchers concerned with educational attainment indicators fuelled the construction of the CIPP evaluation frameworks used mostly in schools bust also in university context such as Italy's. Importantly, it was through these emerging classes that the idea of educational evaluation and management principles interwove and led to diverse experimentations. By addressing the institutionalisation of these international networks, this research thus also

addresses the institutionalisation of methodological innovations which built an output-based approach to educational governance (Bürgi, 2016a).

## Research Methods and Data Sources

This thesis thus recasts the question of the rise of managerial governance in the Italian and British university sectors and VCs' role in shaping it by taking the abovementioned historical lineages as key starting points. Specifically, it looks at the moments of emergence of university evaluation in the two contexts in order to historically trace the interaction between a rising class of university managers and State reforms to higher education.

As such, this research firstly studies scholars' insights on the significance of Cold War context for the development of a systems analysis approach to educational policy and evaluation, tracing the networks of educational experts which emerged. The key intent of this international-level analysis is to explore the presence of Italian or British VCs during this key period, seeking to understand their involvement in the OECD's educational planning movement. This permits to analyse whether VCs acted as conduits between international research on evaluation and national-context developments in higher education policy. I analyse secondary sources as well as OECD reports[5] to contextualise novelties in educational governance and to identify PPBS-inspired models for university evaluation.

Secondly, the research looks at how the two VC committees – the CVCP and the CRUI – articulated their changing relationship with policymakers, typically brought about by State reforms. In this national-level analysis, the research asks how VC committees may have influenced State policies geared at introducing evaluation, or how they may have re-articulated their scopes. The research analyses committees' research, reports, and statements to understand how VCs acted in relation to their changing "external context", re-balancing State-biased readings on the rise of managerial governance. This allows to uncover how national developments played out from the bottom-up. Specifically, this thesis compares evaluation methods advanced by States and those advanced by the VC committees, seeking to understand what their differences tell us about how managerial concepts entered the university walls.

I use both primary and secondary sources to carry out this research. Secondary sources include scholarly studies on two sectors' structural and socio-political developments, debates around the rise of evaluation, and university reports. I particularly focus on research published during the context of NPM reforms to grasp the specificity of NPM and explore how scholars of the time interpreted the rise of an "Evaluative State". These secondary sources are analysed in relation

---

[5] These have been consulted online as they were easily accessible on the ERIC platform and the OECD online archive platform.

to primary sources, used mostly for investigating VC committees' perspectives and actions. Specifically, I have consulted the CVCP's archive at the Modern Records Centre at the University of Warwick and the CRUI's at the Foundation's headquarters in Rome. Documents analysed include reports, working papers, minutes of meetings, conference proceedings and correspondences. The CRUI's archiving gaps for key periods in Italian higher education (i.e., 1969 – 1993) led me to get in direct contact with the committee's past members for a fully-rounded and well-informed analysis. I thus had a few non-structured interviews with members of the CRUI to get more background information for the Italian case. Overall, this research has benefited from the combination of direct consultation of archive material and secondary sources, given the significant chasm – except for a few cases (Cave et al., 1995, 1997; Shattock, 2012; Corradi, 1998) – in historical reconstructions of the committees' history and work on evaluation specifically.

**<u>Structure and Argument of the Thesis</u>**

This thesis has six chapters. The first chapter reviews how two approaches analyse the rise of the new public management governance paradigm in higher education in the 1980s. The first approach considered is that of New Public Management scholars (Hood, 1991; Hood & Jackson,1991; Osborne & Gaebler, 1992,), the second is the neo-institutionalist approach (Deem et al., 2007; Deem & Brehony, 2005; DiMaggio & Powell, 1983). Both these accounts recognize a managerial shift in the governance of higher education sectors but underplay the agency of higher education actors in shaping it because they assume that their actions were secondary to those of political agents and of neoliberal governance more broadly. Specifically, the rise of managerial evaluation practices is treated as a State-led process which was developed to sustain marketization (Tapper & Palfreyman, 2014; McGettigan, 2013) or universities' adaptation to its political agenda (Deem et al., 2007). Together, the two approaches conceptualise the use of radically novel evaluation techniques as State instruments used to subject higher education to new rules, with the result that the question of the emergence of managerial governance is seen as requiring little or no bottom-up analysis. As I argue, this overarching understanding de-politicises the issue as a case of neoliberalism's systemic progression. More concerned with fitting NPM's emphasis on evaluation into neoliberal categories, scholars have produced what I define a "managerial class paradox": managerialism appears as a governance paradigm which does not necessitate governors. Whilst emphasizing NPM's profound novelty and transformative power, scholars do not problematize the specific powers which gradually constructed consensus around a new form of governance at different levels and their respective strategies in this process. This is evident in the case of higher education, where scholars leave the actions and interests of university agents close to fully unexplored when exploring the effective implementation of NPM policies. To address this gap, the chapter discusses how the theoretical tools of an agency-centred approach (Knafo, 2010) permit to historicize the actions of groups which constructed key grounds of a managerial governance approach by developing and diffusing evaluation methods and principles in the university sector.

Building on this critique, chapter two opens the historical research of this thesis. It problematizes the view of managerial governance as a systemic progression of neoliberalism by studying management practices of performance assessment and efficiency in their own terms. These notions have their roots in the development of a systems analysis approach in the US, which emerged and evolved during the early years of the Cold War. This approach lay the grounds of a science of decision-making borrowed from military research, and it influenced how organizational evaluation was conceptualised by groups of evaluation researchers and military intellectuals. The first applied systems thinking to school evaluation by devising an information-oriented model,

CIPP. The second group came from the RAND corporation and constructed a resource allocation and program assessment model, PPBS, to consult federal agencies as well as schools on how to improve the efficiency between investments and outputs.

As I show in chapter three, during the 1960s and 1970s the systems analysis lineage ramified to research on university governance at the OECD through key research networks, adding new innovative inputs to the organization's key role in framing educational governance as an exercise of manpower planning oriented at output-maximisation (Bürgi, 2016a; Ydesen, 2019). Concerned with tracing the relationship between concepts, agents, and the use of new practices of educational governance, I then trace the interactions between OECD research networks and Italian and English research groups participating in them. As I show, English administrators were particularly active members of emerging projects which sought to implement PPBS principles to better govern universities and contributed in this way to the rise of the notion of "institutional management" in universities. Following initial experiments with educational planning along the lines of the "economics of education" approach, Italian researchers actively worked on the CIPP model to think about educational evaluation. Vice-chancellors, too, inherited the legacy of the OECD's work on institutional management thanks to a series of seminars and networks. These developments help to explain how international research on educational evaluation and governance was borrowed and used to develop specific evaluation frameworks in the two case studies, Italy and England.

Chapters four and five turn to analysing national-level dynamics which can account for the rise of efficiency and evaluation principles and methods in the two university contexts. Chapter four looks at the historical development of such principles in England. Firstly, the chapter observes how management was approached and used by different governments to reform the public administration and explores the practical influence of these political attempts on the practices of policymaking. Secondly, the chapter analyses the content of key NPM-inspired university policies to historically address the mainstream argument that evaluation practices entered the higher education sector through the policies of a newly born Evaluative State. As I suggest, while policies reflected new and key funding constraints for universities, they do not exhaustively account for the wide array of developments of management models and evaluation criteria occurring *within* universities. The third and final part of the chapter therefore problematizes the notion of adaptation by discussing the narrative and practical innovations of the British VC committee, which used institutional management practices to construct university evaluation ones. Following a discussion of the rising managerial class in English universities and its role in evaluation, the chapter considers the challenges and opportunities faced by the committee in attempts to institutionalize research quality indicators specifically.

The fifth chapter on the Italian case follows the same structure as chapter four, exploring the specifically national drivers behind the construction of efficiency indicators in the public administration (PA) and in university policies between the 1980s and the mid-1990s. Specifically, it discusses the key objectives of the university autonomy reforms, which entered the sector with strings attached. The chapter continues by discussing how autonomy reforms were perceived in the university sector, contextualising the Italian VCs committee's actions in particular. This opens a key reconstruction of concrete attempts to turn political pressures into new opportunities such as the construction of bottom-up definitions of university autonomy and university evaluation. The chapter continues by exploring the successes and failures of the committee's advocacy of new evaluation indicators and principles. Both chapters four and five discuss the turn to management and evaluation in general terms as well as looking specifically at the development of research quality indicators. This allows to break down the similarities and differences in how a managerial articulation of internal/institutional evaluation methods and research quality ones developed within universities. Secondly, both chapters historicise the actions of the institutional representatives in the two sectors. The reasons for this methodological choice stem from the study's critiques of scholars who conceptualise institutional agents as mere subjects of NPM policies. As I argue, a historicist agency-centred approach to the study of the rise of NPM principles and methods in university sectors allows to reconstruct the role which non-State networks played in shaping the rise of managerial governance through evaluation. These insights, it is argued, can fuel further research on the interaction between university groups and political ones and help to conceptualise the rise and development of managerial governance as a contested political process between them, rather than as an inevitable progression of neoliberal governance. Both chapters four and five argue that, in different ways both British and Italian Vice-chancellors were very innovative as they researched, constructed and partially institutionalised internal efficiency criteria and research quality evaluation criteria.

Finally, chapter six draws out a comparative analysis of the rise of evaluation criteria in the English and Italian university sectors. Specifically, it focuses on a comparison of the actions of institutional representative committees in terms of their engagement with political networks, their intellectual terms of reference and methods, and their interests for implementing evaluation in universities and in the sector.

To conclude, the thesis discusses the implications of historicising the rise of a managerial class within universities for how the commodification of research and instrumentalisation of the university sector can be conceptualised. The conclusion problematizes the notion of a managerial turn in the sector as a State-led process ignited by politicians'

appropriation of NPM principles, suggesting that university agents played a significantly influential role in shaping such principles in their respective contexts –and particularly in the early stages of the introduction of evaluation – and that this insight can free academic debates from the rigidity and reification which generic studies on NPM in higher education create. Specifically, this study advances a more complex reading of NPM which rests on the acknowledgment that the very definition of universities' quality was and is a highly contested process which precisely for this reason needs to be analysed in context specific and agency-centred terms. The historical insights of this study assist this broader research agenda by rebalancing scholarly assumptions on the neoliberal roots of managerialism itself, showing that structural developments of neoliberalism cannot exhaustively account for managerial governance in the university sector.

Chapter I

New Public Management and Neoliberal governance

As discussed in the introduction, re-politicising the rise of evaluation is strictly tied to a greater appreciation of managerial governance in its own terms. In this chapter, I discuss why this is so relevant in existing scholarly debates on the NPM paradigm. Evaluation is commonly accepted as a managerial instrument through which power is veiled in "objective" criteria transmitting a normativity which others must share and participate in. The diffusion of evaluation practices for regulating the activities of both private and public institutions, further, is running parallel to the growth of managerial cadres, signalling that evaluation is a managerial-friendly instrument. Critical evaluation scholars (Borrelli & Giannone, 2020; Zarka, 2020; Dahler-Larsen, 2020) argue that evaluation can be controlled to personify given forms of knowledge and their advocators, used to exert power by discriminating and de-validating other "truths" (Zarka,2020). Simultaneously to this broader relationship between evaluation and the exercise of power rooted in the validation of given normativity against others, evaluation is also developing into an instrument for organising labour by defining its conditions of access to and participation in the labour market (Nicoli,2010,2011; Pinto, 2012,2019; Nicoli & Paltrinieri, 2014), and by defining benchmarks and processes necessary for the distribution of resources in both enterprises and public institutions. A second significant dimension of evaluation is thus the way it is used to organize and control human and material capital. Differently to scientific management, evaluation criteria are more than ex-post assessments of production or benchmarks against which it can be considered (un)productive. Rather, they are devices which empower decision-makers by allowing them to define workers' boundaries of action and constantly monitor workers' activities. Indeed, it is often by changing evaluation criteria such as the scales of research quality rankings that decision-makers can re-affirm their control over workers, who are driven to adhere given the punitive dimension of evaluation (most evident in performance-based-funding (PBF), see Zacharewicz et al., 2019). Thus, evaluation is importantly a device at managers' disposal when used in these ways and can be used to set-off new organizational behaviours which directly influence workers.

These examples of the managerial dimension of evaluation point to the importance of investigating the novelties which its use in public sector governance is bringing about. However, this is a challenging task because managerial governance within the broader New public management process is predominantly seen as a form of neoliberal governance, rather than as a paradigmatic shift from it. Scholars agree on the idea that managerialism is a new governance paradigm because of the new techniques used by neoliberal policymakers for pursing neoliberal principles (Mascarenhas, 1993; Pollitt, 1990; March & Olsen, 1989), and less so because of the new

ways in which it reshapes governance, i.e., those who govern and those who are governed. In fact, many literatures define managerialism as the mere application of private-sector concepts – such as those of efficiency, incentives, and performance – within State apparatuses and public institutions, reducing the significant transformations brought about by management practices to a case of using different practices for the pursuit of neoliberalism's conventional objectives: a shrinking State, the creation of market dynamics and the use of incentives for incrementing competitiveness.

This thesis dialogues with such literatures, inviting them to engage in historicist studies on managerialism. In an interdisciplinary vein, I address how different traditions obscure the specific histories of how managerial practices have become so central for governing diverse contexts. To be sure, the main intent of this study is to uncover the ways in which managerial governance has come about through the use of evaluation in Italian and English universities, seeking to provide points of reflection for addressing, more broadly, the advent of managerial governance in public universities. In this intent, the research inevitably speaks to and with studies on neoliberalism, because it is in the context of neoliberal governance that the dismantlement of public universities and the implementation of new public management principles is playing out. Far from seeking to re-define neoliberalism or negate that managerial transformations are occurring within a neoliberal context; I problematize the view that this be treated as sufficient reason for generalising managerial practices as having emerged for, and at the disposal of, neoliberal States. Borrowing from existing studies (Eagleton-Pierce & Knafo, 2020), I interpret the fact that managerial classes have been rising and gaining leverage in extremely diverse contexts – such as business schools, public administrations, public institutions of different types, regional/local administrations and NGOs, amongst others – as evidence that the phenomenon of managerialism is complex and hardly generalisable. Indeed, these multiple manifestations of the use of managerial practices hint that managerialism is influenced by the way that different institutions make use of, or develop, managerial instruments of which evaluation is just one. I follow the suggestion that such diversity brings to a more attentive and open-ended definition of managerialism as '[…] as a set of ideas, cultural frames and technologies, that, although operating in a capitalist environment, have a certain autonomy' (Eagleton-Pierce & Knafo, 2020, p.772). ***By rejecting the assumption that managerialism manifests itself equally in different contexts***, there is more scope for studying it in the university context by asking which developments might be specific to the university sectors compared to managerial developments elsewhere, such as in the public administration. Investigating the emergence of university managers by historicising developments in the university context opens up a methodological frame focused on specificity, and this means acknowledging that the rise of managerial classes can develop very differently depending on the context. Instead, NPM scholarship does not provide sufficient methodological instruments for thinking about managerialism outside the State walls and

historicising managerial classes *in* the public sector, for it is heavily focused on exploring the managerial governance *of* the public sector. Scholars study managerial practices to explain neoliberal governance and/or the neoliberal State. This structuralist reading has a strong top-down bias when it comes to explaining how and why agents in public-sector institutions first used and continue to develop managerial practices, the answer given is indeed that there was no alternative but align to the developments occurring inside State walls.

I this chapter, I problematize the argument that evaluation is a managerial instrument at the disposal of States. Firstly, I show how political economy and new public management literatures convey this view when they suggest that managerialism is merely functional to neoliberal governance. By underestimating the significant differences between managerial and neoliberal thought, literatures agree on the view that NPM reforms in public administrations reflect a convergence between managerialism and neoliberalism. As I argue, this alleged alliance between political and managerial classes reflects a *managerial class paradox*, managerialism is discussed but the managers enacting it are scarcely historicised. To explain the managerial class paradox, I refer to Ralph Miliband's concept of "structural abstractionism" (1973). Secondly, I present neo-institutionalist approaches which use the category of isomorphism to account for institutional change, discussing both broad literatures (Powell & DiMaggio, 1983) and those which discuss the university context specifically (Deem et al., 2997; Deem & Brehony, 2005). As I point out, the neo-institutionalist approach is highly concerned with unpacking institutions' organizational change; yet institutional agents (i.e. university agents in this case) are considered mere subjects of top-down governance or changes. The implication is that institutional agency is conceptualised as being *adaptive* to systemic developments. From an agency-centred approach, this is problematic because agents' concrete actions are simply treated as the result of something else, namely institutional logics. As a consequence, the growth and empowerment of a managerial class is detached from the agents themselves, leading to risky reifications. In other words, this chapter shows how scholars from different traditions conceptualise managerial governance either as emerging from neoliberal necessities or as existing only thanks to neoliberal structures and/or agents. I conceptualise this State-bias as a structuralist understanding of power which is closely interwoven with weak historical enquiry. As I argue, this has repercussions on how evaluation is studied: those literatures with generic readings of the rise and development of managerialism in fact treat evaluation as a structural transformation where specific groups' actions make littledifference. Finally, I discuss how historicization is central for theorizing, as recent Political Marxist literature (Knafo, 2010; Knafo & Teschke, 2017) shows. Using Political Marxism as my methodological reference, I discuss how the agency-centred framework can free conventional conceptualisations of agents from the theoretical constraints which scholars frequently impose on them when placing too much analytical

attention to the structures within which they act. Building on this, ***I argue that historicising managerialism through its agents*** is a crucial step for re-politicising the rise of evaluation and for theorizing on how its use for governing institutions has altered power relations in favour of managers. Attentive and agency-centred historicization allows to connect new evaluation and managerial practices to the specific groups of agents devising and using them, or to groups contesting and re-defining them. Specifically, this means exploring how the use evaluation instruments can have advantages or disadvantages depending on the groups being explored; rather than simply assuming that university evaluation has, generically, weakened the university sector. For if on one hand there is increasing evidence of the relationship between the rise of evaluation and the erosion of universities' public character, on the other there is also increasing concern about the rise of power – and salaries – of managerial groups within universities (vice-chancellors, administrators, finance officers) and in the sector more broadly (evaluation and funding agencies).

## 1.1. The managerial class paradox in new public management scholarship

The rise of managerial powers within capitalism has long been documented, often to point out the changing nature of capitalist power as having gradually shifted from the ownership of the means of production to their control (Berle & Means, 1932; Burnham, 1941). When notions of "managerial capitalism" were first coined (Marris, 1964) scholars emphasized the radical implications of this development. Arguments were in line with previous claims that managerialism was detaching itself from its conventional form of the immediate post-war period (Burnham, 1941) and that it appeared less concerned with directing and coordinating production than with constructing a proper theory of the firm based on managers' decision-making powers. As such, managers were rising to power in an unexpected way and through innovative means. In his book "the ungovernable society" ("la société ingouvernable"), Grégoire Chamayou (2018) warns readers not to under-estimate the rising political powers of the classes of managers emerging in the U.S. corporations of the 1950s and 1960s. As he suggests, managers ought to be considered as political agents because they began developing new ideas and practices for creating '[…] an analogy of political governance within the private management of economic affairs' (p.49). For example, managers tried to appease factory workers' uprisings of the 1970s by implementing notions of "participative management" (see Marrow, 1975 for an example of this), as was the case in the Topeka factory of General Foods (Chamayou, 2018). They also defended themselves from the criticisms of civil society by taking a strategic turn and engaging with "corporate social responsibility" (CSR) practices. Through CSR, managers created an ethical veil which allowed to rhetorically claim that business was following ethical conducts whilst, in practice, they disengaged with the practical and unprofitable consequences that following such ethical codes would bring. In

this way, managers asserted that they were keeping new concerns in consideration, bringing moderate critics to their side and thus politicised their business roles (Sanders, 1998) in a way that was advantageous for them. Interestingly, Chamayou (2018) points out that these managerial strategies were completely at odds with how neoclassical economists framed the firm; namely as a locus with horizontal, a-political hierarchies driven by economic interests and rationality (see Alchain & Demsetz, 1972). Whilst neoliberal theories of the firm negated power relationships, for example by defining labour contracts as mere transactions, managers were concerned with re-politicising the firm; and '[…] saw conflict everywhere' (Chamayou, 2018, p.137). For these managers, economic theory completely ignored the fact that business needed to develop a "political expertise" (Sanders, 1998). Chamayou's work is coherent with other studies that explore the different principles of neoclassicals' public choice economic theory and managerial theory, especially strategic management (Knafo et al,2019; Knafo, 2020). By discussing how neoclassical thinkers have tended to re-frame governance through market-dynamic axioms, scholars suggest that public choice theorists were actually hesitant to managerial techniques because they did not trust bureaucratic powers and their expansion following the post-war period. Indeed, a key neoliberal icon such as Margaret Thatcher reflects this, for once elected she set out to revert her predecessor's managerial reforms (Lewis, 2011). As the historical development of PPBS also reveals, managerial groups formulated organizational instruments geared at giving managers new decision-making powers at the top of organizations. Notions of strategy, planning and efficiency were thus developing simultaneously and in antithesis to neoclassical ones of transaction costs, information asymmetries and market dynamics within firms. Once analysed together, these studies (Chamayou, 2018; Knafo at al., 2019; Knafo, 2020) indicate that neoclassical economic theory did indeed play a part in the development of new theories around the managerial turn in the business context, yet in terms of opposition to managers' theories. Further, these authors venture to propose that managerial theorizations around decision-making predated the ascendancy of neoliberal politics. We should, thus, be careful **not to conflate these two approaches on governance and their historical manifestations as coherent and aligned to each other**. In what follows, I discuss how this conflation plays out in political economy and new public management scholarship.

New Public Management (NPM hereafter) stands as they primary paradigm used for conceptualising the development of neoliberal governance. For many scholars, it describes a new stage of neoliberalism at the turn of the 1980s. Engaged in the dismantlement of their welfare economies, Anglo-Saxon States first, and European ones after; are said to have appropriated management practices to establish quasi-markets in public sectors. The NPM paradigm thus arguably allowed States to marketize the public sector by handing governance down to public-sector institutions (hence the term "devolved governance") whilst monitoring – and when necessary,

controlling – their  development "at a distance" thanks to performance criteria. A puzzle, however, characterizes these readings of how NPM plays out: managerial practices (such as evaluation criteria) are seen as both centralizing and marketizing instruments at the disposal of States. The NPM framework is thus frequently mobilized to explain contrasting governance trends (Moini,2017), such as Thatcher's *roll-back* neoliberalism and Clinton's *roll-out* neoliberalism (Moini,2017; Peck & Tickell, 2002). The idea conveyed is that management represents a novel State instrument for enacting neoliberal governance, a view which I argue is symptomatic of gaps in the historicization of managerialism. Indeed, the puzzle as to of which kind of neoliberal strategy managerialism allows to pursue is  also evident in studies on public universities' transformations. On one hand, evaluation allows to construct a quasi-market in the supply and demand dynamics for higher education (especially in Anglo-Saxon countries) (McGettigan, 2013; Brown & Carasso, 2013; Ball, 2009). In this view, rankings signal quality to "consumers" and universities increasingly turn to private sources of income. Further, States have set off competitive behaviour between universities for funding, thus withholding funding from less deserving universities (Brown & Carasso, 2013; McGettigan, 2013; Tapper and Palfreyman, 2014, Slaughter & Rhoades, 2004). On the other hand, evaluation is treated as a centralizing device precisely because of its direct relationship to State funding, such that universities are increasingly vulnerable to political steering and to new forms of discipline (Pinto, 2012; see Power, 1997). From the academic literature, it is still unclear what novelties management practices are bringing to governance. Rather, these appear as generic and universal efficiency instruments which can be mobilized to implement different neoliberal agendas. This fuels the idea that managerial practices are, either way, functional to and in alignment with the pursuit of neoliberal governance, helping to pursue classical tenets of a limited State and encouraging market dynamics. In highlighting that neoliberal governance has increasingly endorsed a more "moderate" function through strong (rather than limited) States which ensure the conditions for the operation of the market, recent scholarship has further supported the idea of a tight convergence of aims between managerialism and neoliberalism. Indeed, neoliberal States constructing quasi-markets are – in contrast to Keynesianism and to classical liberalism– concerned with why intervene and what to intervene in (Giannone, 2019) rather than *whether to intervene in the first place*. Having moved away from classical theorists who saw market dynamics as a natural element of society itself, neoliberal governors since NPM use managerial instruments to ensure that markets work. They thus engage in *neoliberalisation* of society broadly*,* using technical-economic evaluation to re-mould institutions into competitive market-players (Giannone, 2019). Recent critical studies on the relationship between managerial practices and neoliberal governance allow to unpack why the argument that management centralizes State power may be coherent with neoliberal governance, showing that neoliberal governance is no longer intrinsically tied to classical economic

theory (Boltanski & Chiapello, 2005; Bruff, 2014; Davies, 2016; Duménil & Lévy, 2015,2018). At the same time, it is through the frequent re-definition of neoliberal governance since NPM that scholars can keep the phenomenon of managerialism into account without concretely historicising how its differences to neoliberal economic theory have arguably dissolved or converged. This perpetuates a generic understanding of managerialism, whereby if we accept that neoliberal governance is changing, then we can accept that new public management is both about centralization and marketization, about limiting and enhancing State power simultaneously. It emerges that managerialism does not seem to be shaped by specific cadres[6], and can therefore be framed as a progression of new neoliberal goals through different means. Scholarly literature is in this way actively blending the boundaries between managerial and neoliberal governance, detaching the still open-ended development of managerial governance practices from the historical context and agents that they emerged from.

As I suggest, in NPM scholarship this view is interwoven with the assumption that NPM was the result of a convergence of interests of State and managerial powers. Public Administrations are the first key locus where NPM reforms were enacted, reason why an array of NPM scholarship analyses administrative change (Hood, 1991; Hood & Jackson, 1991, Aucoin, 1990, Pollitt, 1990). When the reform movement began, scholars defined NPM as 'a new global paradigm' (Osborne & Gaebler, 1992), with an unprecedented pace and intensity which began to manifest itself first within the State machinery and then all the way down to the institutional level (Dorrell, 1993). The key perception was that States were being asked to act like the private sector, and that the reforms reflected the ginning of a process which would eliminate political complexities intrinsic to the expansion of expanding public administrations thanks to managerialism. As Aucoin (1990) pointed out during the early stages of NPM scholarship, however, there were significant differences between managerial and political perceptions of NPM administrative reform. ***The public administration reforms of the 1980s, for example, aimed to reduce the growing powers of public bureaucrats vis-à-vis that of politicians.*** Public choice theorists sought to elevate principles of representative democracy to re-enforce the role of politicians and advocate for bureaucrats' disempowerment. Reforms thus also geared at reducing managers' autonomy[7] within the PA itself,

---

[6] Duménil & Lévy, for example, concern themselves with conceptualising the rising wealth of managerial cadres within capitalism in "Neoliberal managerial capitalism: Another reading of the Piketty, Saez and Zucman data" (2015). Although they substantiate how the managerial rise in contemporary capitalism owes much to managers' appropriation of surplus value in the form of high wages, the scholars seem to suggest that capitalist logics are merely re-manifesting themselves through a new composition of the capitalist class whose build-up is scarcely historicized, with the result that one gets the general feeling that managerialism is a natural progression of capital accumulation logics within neoliberalism itself.

[7] Knafo elaborates on the specific power interplays in the use of managerial practices to explain its contrast to public choice positions: '[…] public choice theorists have long been wary of performance management because it requires generating data processed by a bureaucracy. Performance management produces information asymmetries that public bureaus and public decision-makers can manipulate to serve their own interest. On the one hand, lower levels of a

for in the 1970s they had gained increasing leverage within the bureaucracy, in large part through a technical approach to public expenditure. As Aucoin put it, '[…] elected politicians […] tend in any event to be suspicious of "budget maximizing bureaucrats"' (1990, p.116). For public choice theorists, thus, NPM reforms to the PA would have ideally re-centralized leadership in the hands of politicians by, for example, reducing the size of decision-making *loci* such as executive and ministerial committees (Aucoin, 1990). From the public choice perspective, managerial reform referred to the changing leadership of ministers who, in the attempt to re-gain political influence over bureaucracy, would act more decisively as executives of policy *as well as* of administration. On the contrary, new and complex managerial practices – such as budgeting and policy review – advocated decentralization and delegation of powers if efficiency was to function. In part, this is because efficiency practices first relied on the use of feedback indicators and reporting exercises which required that information be gathered at different levels and controlled and elaborated at the higher managerial level. Managers sought to de-bureaucratize administrative processes by downplaying the political dimension of decision-making to the advantage of more "rational" practices such as resource planning. As such, the rise of NPM in Anglo-Saxon PAs was characterized by this paradox whereby for politicians to gain leverage over civil service through administrative reform they nevertheless mobilized principles which could actually re-enforce a detachment between the technically specialized functions of bureaucrats-managers and the political roles of politicians. As Knafo has recently summarized, '[…] from a public choice perspective, speaking of performance solves nothing if it is determined by bureaucratic and managerial processes' (2020, p.7). As NPM reforms were introduced, contradictions emerged as to whether they would swing towards public choice theorists' advocacy for greater political control over the PA or towards managers' emphasis on decentralizing political powers for the sake of efficiency. As Aucoin pointed out, the reforms promoted '[…] movement in both directions simultaneously' (1990,p.129). Taking the British example once more, Margaret Thatcher criticized Edward Heath's managerial reform –known as the "reorganization of powers"– precisely on the basis that these were too distant from political reality (Lewis, 2011). This suggests that managerial principles and their unfolding within the PA were not at all smoothly welcomed by typical neoliberal politicians. Nevertheless, NPM reforms have largely been discussed as a quasi-immediate success of managerial practices within politics, as these importantly responded to State concerns on cutting costs. This view is frequently strengthened by referring to the systemic pressures that States themselves lived as they turned neoliberal at the turn of the 1980s. Often, reference is made to the

---

bureaucracy have to produce numbers which they know will impact upon them, thus creating strong incentives to dress up performance data so as to flatter those producing the services being evaluated, a problem that has crept up time and time again under NPM. On the other hand, it also makes it easier for those at the top, who collect and use this data, to manipulate these numbers so as to use them for political purposes (e.g., to lobby for further resources or exaggerate their accomplishments)(2020, pp. 7-8).

rise of "large States" between the post-war period and the 1970s which endorsed social-democratic principles in the welfare-State model. No longer able to sustain such model following the oil crisis of 1973 and the rise of right-wing free-market politics, States turned their back to the public sector and initiated deep reforms which later brought to new political-economic agendas. The arrival of efficiency-oriented practices in public administrations was thus itself a response, on behalf of States, to global pressures which trapped States into pursuing cost-cutting as the way out. It is thus from these neoliberal constraints that States appropriated private-sector logics and began to streamline the PA first and the rest of the public sector later (Pollitt, 1990; Larbi, 1999, Mascarenhas, 1993). As a "management consultocracy movement" (Hood & Jackson, 1991) set-off within State machineries, the "alliance" between politicians and managers developed and allowed to set out new methods of public governance. By hiring management consultancy firms which brought new operational guidelines in public administrations, States arguably devised the grounds of new public management and began to experiment unprecedented exercises such as civil staff activity reporting, performance measurement, and the use of cost-benefit analysis in policy formulation. In other words, developments in the public administration further reinforce the idea that NPM can be easily generalized as an attempt to have "[…] the public sector operate more like the private sector" (Best, 2014). Similarly, scholars have downplayed the significance of frequent "implementation gaps" (Ongaro & Valotti, 2008) of administrative reforms. To make an example, as of 2014 the Italian PA was still struggling to implement performance-assessment practices and set-off new managerial practices more broadly (Rebora, et al., 2016). Still, this has been downplayed as a case of contrasting "administrative paradigms" (Capano, 2003) and insufficient political will. The international proliferation of NPM research (Di Mascio & Natalini, 2018; Dill, 1998; Pollitt, 2013; Pollitt & Dan, 2011; Reale & Potì, 2009) attempts to account for the in-efficiencies of the public management governance model through a comparative outlook of NPM reforms in different States' public administrations. By exploring different countries' experiences, scholars benchmark them against Anglo-Saxon developments where NPM first set-off, suggesting that the scarce enactment of NPM is the result of diverse parliamentary and political structures. This is clearly an acknowledgment of the complexity of reforming the PA through managerial procedures that were previously unfamiliar to civil service staff. At the same time, these accounts answer the issue of a weak implementation of NPM reform by pointing to weak political abilities. This leaves the complexity of the techniques and the development of the managerial lineage within State walls historically unspecified. The assumption that these practices were central to improving political governance remains strong even if in practice they failed to do so and politicians themselves importantly reject them at first.

Scholars fuel what I define a managerial class paradox: public governance is said to have turned managerial, yet the managerial class making this possible is assumed rather than being concretely historicized. By situating most analytical attention to State governance, scholars do not thoroughly investigate conjunctural moments in which States formed new alliances with managerial groups. As a result, implementation gaps or neoliberal attacks against managers (Lewis, 2011) are marginalised, when instead they could significantly help to explain the very novelty of the governance turn that scholars seek to account for through NPM categories. Whilst the key role that consultancies in central government played in the construction of a new public-private governance model is often invoked, for example, specific groups and their methods are rarely brought to the foreground – with some exceptions – (Saint-Martin, 2000). The result is that it becomes difficult to grasp the *development* of newly emerging governance practices: how these were received, resisted, changed. Whilst it is not the scope of this thesis to question that a managerial-neoliberal alliance may have eventually formed, this research's focus on its emergence calls for greater historical specificity as a valid frame for grasping how managers' empowerment may have occurred. An example of the managerial class paradox at play is the scholarly tendency to associate the rise of NPM with the "third way" of the 1990s in the UK (Peck & Tickell, 2002; Pollitt, 2013), although management consultancies were present in central government as early as the mid-1960s (Saint-Martin, 2000). This distortion of the period in which managerial groups enter State walls and actuate important transformations exemplifies the abovementioned assumption that private sector practices set off in function to the rise of neoliberal governance, rather than being also the result of managerial groups' context-specific actions.

To overcome this reductionism and its effects, the managerial class paradox can be resolved by rejecting the notion of objective interests between two classes – such as capital accumulation – as the driver behind the emergence of new public management, turning instead to investigating interests in historical and context-specific detail. Precisely because the State is often faced with constraints or threats to its dominance, then new alliances geared at maintaining it ought to be analysed in their conjunctural character. This can allow to explain States' turn to diverse practices more specifically. At the same time, it can shed light on how new groups mobilize specific discourses, practices or resources in the attempt to become a central part of new political developments, such as the rise of neoliberal governance. By taking a historical perspective on managerial groups' actions, scholars can better explain turning points such as that of the 1980s. This could include, for example, starting with the question 'what groups and actions made managerial practices for State governance more successful in the 1990s than in the mid-1960s, when they were also being used in State walls?'. A question such as this will allow to grasp the core of an emerging alliance between politicians and managers because it seeks to account for changes in

the relationship between the two rather than taking it as a progressively positive alliance from its establishment onwards. As a corollary, this more specific framing allows to surpass a structuralist framing of power whereby the entrance of managers in State suffices to argue that they were dominant and that they acted in the interests of the neoliberal political class. Importantly, efforts must be made to understand what intrinsic features of a managerial class may have helped it to gradually acquire legitimation and power within State walls. This is quite in contrast to the suggestion that managerial groups successfully became dominant once they entered State walls, when their interests were activated by neoliberal class. Indeed, what counts is grasping how managerial interests were mobilised to appear attractive to State interests in the first place. This shifts the question of the legitimation of managerial interests to moments **prior to their entrance within the State.** This can help to explain how managers try to attract State agents, and therefore helps to account for the transition of managers from outside the State to inside the State,  as groups of experts working with politics to devise new forms of public governance. However, much is still to be done, because NPM scholarship (Hood, 1991; Hood & Jackson, 1991; Osborne & Gaebler, 2002) suggests that managerial groups were very easily welcomed by politicians, as if they made no effort to legitimise their practices. Authors believe that managers' advocacy of cost-efficiency, performance evaluation and rationalisation is sufficient for acquiring widespread political (and bipartisan) support. This reading de-politicises the political process through which public governance transformed following the rise of NPM principles: **managers' political rise is studied and narrated as an inevitable outcome of neoliberal establishment**, it is de-voided of the many challenges which managers may have faced in the process of legitimating their views and practices. Such struggles can only be grasped if we adopt a long-term perspective (genealogical in many ways) which identifies managerial groups which are dominant today and traces their steps to reconstruct groups' composition, backgrounds, and actions. Explaining managerial groups' reach to politics thus becomes an exercise of investigating these in relation to their context and their successes and failures at obtaining political attention. In a similar vein, Miliband provided Marxist theory with important methodological instruments for breaking down the terms of partnerships between State power and class power (1983). Whilst acknowledging that the relationship between capitalists and the State '[…] has on the whole been very good" (Miliband, 1983, p. 64), Miliband rejected the idea that State actions be explained in terms of constraints placed upon it. This is insightful because it helps to reflect more specifically why, besides the often-invoked expansion of the public administrations, States began to consider the use of managerial practices. Through contextual and relational analysis, therefore, both State and managerial groups' actions can be better articulated as a political process, avoiding the 'derealization' of dominant classes (Miliband, 1973) whose specificity is subjected to their broader intention of being dominant and thus pursuing, like

the State, broad capitalist interests. Rather than explaining the one or the other's empowerment in structuralist terms, these insights allow to study the rise of a managerial-neoliberal alliance in political terms as the result of a complex interplay between two groups who attempt to appeal to one another to achieve their specific interests.

## 1.2. The rise of managerial governance in universities as a case of institutional isomorphism

At an institutional level, scholars often analyse the degree of application of NPM through the concept of *institutional isomorphism,* coined by the Powell & DiMaggio in their notorious paper "The iron cage revisited: institutional isomorphism and collective rationality in organizational fields" (1983). Differently to other organizational theorists of the time (Hannan & Freeman, 1977; Child & Kieser, 1981) concerned with explaining organizational diversity, Powell and DiMaggio elaborated diverse hypotheses for explaining 'what makes organizations so similar' (1983). The authors observed organizational fields, exploring how the clustering and establishment of similar organizations drove yet others to assimilate the '[…] structure, culture and output' of such fields (p.147)[8]. Convinced that competitive dynamics and a *homo economicus* rationality were reductionist concepts for explaining institutional behaviour (Lanzalaco, 1995), the authors coined the term *institutional isomorphism* precisely to emphasize non-economic factors. They suggest that groups of organizations tend to go through a process of homogenization –*isomorphism*– as they become increasingly established. What most sparked their interest was not just that this process took place, but also that it does not seem oriented at achieving greater (productive) efficiency as much as at strengthening organizations' power and legitimacy. By making this claim on the process of isomorphism, the authors introduced new important conceptual points for reflecting on how organizations do not rely solely on their productive ability for affirming their position, but also on a complex interaction with each other (within a "field") as well as with organizations different to them. In many respects, Powell and DiMaggio attempted to open the black box of organizational behaviour. If on one hand they acknowledge that innovative organizational practices often emerge to improve productivity or performance, on the other they suggested how this eventually turned into a goal secondary to that of achieving legitimacy, recognition, and power. Using Bordieuan categories, they argue that as new practices spread and are shared in a given organizational field, they eventually acquire a symbolic meaning. Rather than changing for the purpose of production, thus, organizations may adopt practices new to them simply to obtain a position within a given field. Institutional isomorphism thus allows to explain why institutions adapt to behaviours or

---

[8] To be precise, the authors defined organizational fields as '[…]those organizations that, in the aggregate, constitute a recognized area of institutional life: key suppliers, resource and product consumers, regulatory agencies, and other organizations that produce similar services or products.' (p.148)

structures which don't typically belong to them and assimilate new behaviours surrounding them.

It is organizations' quest for legitimacy which explains why and how these tend to homogenize, such that the authors pointed to three different types of isomorphism: coercive, mimetic, and normative, each driven by three different stimuluses external to organizations: the State, the social environment (i.e., exogenous pressures broadly), and the professional community. In coercive isomorphism, organizations assimilate the demands of political influence (such as new regulations). In mimetic isomorphism, they instead respond to the uncertainties brought about by an external change by mimicking the behaviour of existing and established organizations. In normative isomorphism, finally, organizations adapt to the rules and/or structures of a given profession in order to fit into it more coherently, meaning that they will tend to adopt behaviours such as hiring personnel from the same industry. These categorizations of isomorphism all describe different ways in which organizations will tend to encapsulate new features to in order to acquire legitimacy. Whilst normative isomorphism seems to explain the professions' tendency to maintain legitimacy and power through closure (Weber, 1922), the categories of mimetic and coercive isomorphism can be used to account for organizational reactions to changes which threaten their status quo. By extension, these two categories are a way to conceptualise how organizations will seek to maintain power vis-à-vis their external environment.

This rationale also populates the debate on universities' managerial turn. In particular, it is often suggested that universities – just like other public institutions – have undergone a process quite similar to coercive isomorphism, as they were encapsulated by novel NPM logics pressed onto them by State reforms (Deem & Brehony, 2005; Deem et al., 2007; Slaughter & Rhoades, 2004; Palfreyman & Salter, 2014). Indeed, managerial governance within the State is an underlying theoretical pre-condition for explaining the rise of managerial governance in university institutions. This explanation rests on the State's own subjection to global capitalism, which, once faced with falling profits at the turn of the 1980s, was pushed into efficiency-driven cost-cutting practices borrowed from the business and management consultancy sectors. Public institutions faced with these changes could not stop them, for their full financial dependence on the State – accumulated throughout the post-war period – makes them highly vulnerable to and dependent on State demands. From this perspective, it is management's claim of efficiency which make it a new dominant "cultural archetype" (Deem et al., 2007). As an exogenous pressure, management is conceptualised as '[…] stand[ing] above, indeed outside, wider social, moral and political struggles' (Deem et al., 2007, p.6). Universities' managerial turn was therefore *forced* by the necessity to conform to new inescapable cultural imperatives coming from management and transmitted to them through the State. This is quite in line with the neo-institutionalist conceptualisation of State power as that expanding and extending logic which

molds institutional isomorphism. As I argue, the implication of this view of universities' managerial turn is that the internal process of institutional isomorphism is not explained in practice, but rather assumed and consequently described. This emerges in the way that universities' internal transformations are discussed.

In "Knowledge, higher education, and the new managerialism" (2007), Deem et al. explain how universities' adaptation to managerialism progresses as universities import managerial practices following State reforms (Deem et al., 2007). The authors explain institutional change by assigning all transformative power to processes and agents lying *outside* the university sector. They use managerial ideology as the key explanatory device which can account for the articulation and application of radically novel practices in universities' internal hierarchies (Deem et al., 2007, Deem & Brehony,2005). Although neo-institutionalism traditionally places significant attention to values as key features for understanding organizational behaviour (DiMaggio & Powell, 1983),institutions' universities' need of surviving by acquiring a new legitimacy dressed in managerial clothes becomes that greater necessity which encapsulates the entirety of the set of practices and behaviours existing in universities before the arrival of the exogenous, ideological-political pressure to adapt. With a functional necessity to maintain legitimacy in the eyes of a new type of power being exerted on them, institutions thus merely adapted to a new governance paradigm now filled with managerial values, such as efficiency and performance. Universities arguably implemented internal evaluation frameworks and abided to external evaluation in response to pressures of having to prove their worth and had no viable alternative to this *re*-action (Deem, et al., 2007). Institutional survival is pursued above all else, somehow silencing institutions' conflicting interests and values. The central issue of how institutional *change* occurs remains highly unproblematized, for it is assumed that external pressures were simply too great to be resisted or re-defined in alternative ways by university agents. This leaves no space for a detailed study of how management was internally articulated by institutional agents to achieve internal consensus in the first place. The development of internal consensus to radically new organizational behaviours is not really considered, for it is implicitly assumed that it naturally develops as universities change to gain legitimacy vis-à-vis their external environment. In a way, it is as if internal changes occur as a response to institutions' need of survival. From a retrospective position, (Deem et al, 2007; Deem & Brehony, 2005) the concept of institutional isomorphism is validated, because universities eventually adopted managerial practices. However, scholars do not concretely account for managerial powers, for their *de facto* rise. We remain unaware of which strategies and discourses were internally mobilised, and by whom, to redefine the traditional organizational behaviour of collegial governance.

An example of this is how British Vice-Chancellors' "Jarratt report" (CVCP, 1985a) is discussed. In the report, VCs advocated for the use of management practices and a management information system, setting out precise guidelines to their implementation. Deem et al (2007) discuss this document as evidence of universities' assimilation to the managerial archetype, although scholars of the time had previously suggested that '[…] soft and hard forms of managerialism correspond roughly to two phases of the managerial revolution in British higher education […] the *first* that developed *within the universities*' (Trow, 1993, p.11, emphasis added). Furthermore, British university administrators began researching performance assessment indicators before the first State selectivity cuts in 1981 (see chapter four of this thesis). Still, scholars' minor analysis into the Jarratt report evidently fortifies the verdict that '[…] most 'new managerial' permeation of universities can be attributed to external factors' (Deem et al., 2007). This significantly downplays an internally-driven change to universities' traditional functioning – a key historical insight – to '[…] a supplementary ingredient that can always be later added […] to produce a 'more complete picture', *if needed'* (Knafo & Teschke, 2017, p. 7, original emphasis). The process of isomorphism effectively becomes a theoretical device so strong that scholars replace the analysis of organizational processes with the analysis of isomorphism. The latter is treated as that process which can exhaustively explain university agents' behaviour, shaped indeterminately once the managerial logic was suddenly set off. Institutional isomorphism conceived in this way does not create the adequate methodological grounds for conceptualising agents' agency. The notion of adaptation to external changes extends to the study of institutional agents' actions to the extent that these become fully determined by new cultural values under which they are subsumed. In a similar vein, Fourcade (2006) argues that isomorphism '[…] serves as a near-universal metaphor for any process of widespread institutionalization' (p.154). In pointing out that  Powell and DiMaggio (1983) referred to the isomorphic *structure* that organizations tend to take – rather than dealing in detail with the process of isomorphism itself –; Fourcade has highlighted how conceptualising the process of diffusion as a fixed and linear transmission of new practices or ideas actually downplays the transformative power of the diffusion process itself.  In her own words, '[…] diffusability constitutes one of the defining elements of institutional or organizational change, not simply the other way round' (2006, p.54). Although Fourcade refers specifically to the diffusion of the economic professions – and shyly refers to power relationships – the point is very valid for problematizing how neo-institutionalists' emphasis on adaptation can become a reifying concept obscuring the analysis of how institutional agency plays out.  State reforms to the university sector can be defined as being dynamic; as events which university agents were trying to influence. This does not mean that scholars should disregard or underplay how the spread of norms, values and material conditions regulating institutions "from above"- or "from outside"- concretely influences

agents and their actions. Rather, external pressures pushing institutions to appropriate new practices are **constraints which still had an open-ended outcome** on institutions. Institutional adherence to constraints shouldn't be taken as a given outcome, even if  from a retrospective standpoint scholars can fall into this trap because they know that certain developments effectively occurred. It is an act of reification to study historical developments by imprinting them with the opinions and knowledge of the present. It is key to study history in dynamic terms,  to study the process of diffusion of ideas and practices by focusing on how agents act and participate during the diffusion of new political-ideological contexts. Only in a second moment, having taken account of these actions and interactions, of struggles and unintended outcomes, should we begin investigating if, and how, new ideas and practices become concretely institutionalised.

I have discussed in detail how and why the neo-institutionalist notion of institutional isomorphism is problematic for debunking the emergence of new institutional practices and how agents may have acted to legitimate them. This is even more compelling given the still open-ended process of universities' transformations, where new evaluation benchmarks are frequently introduced and existing ones re-shaped. By considering the legitimation path as the only viable outcome of the rise of new values and principles, we risk undermining the agential forces acting towards their institutionalisation, which in turn de-politicises the reading of the ways in which evaluation is re-structuring internal power hierarchies at the vantage of managerial powers. To summarize, theorizing through isomorphism is structural in two ways: firstly, in the identification of new pressures, located exclusively in "external" forces without an analysis of how novelties may emerge autonomously and/or internally in institutions. Secondly, agents are seen as inevitably reproducing such forces for the sake of survival, without consideration of their ability to redefine or reject given pressures. This second aspect in particular creates what Powell himself later defined as a "black hole of genesis" (Padgett & Powell, 2012a), '[the] analyst takes as given some constitutive features of the hypothesized individual or actor'  to explain agents' behaviour rather than historically investigating it. Indeed, scholars often presume that institutional actors "turned managerial" once managerial ideology and constraints emerged: hence the mere rise of management values suffices to set off radical organizational transformations. A dynamic and agency-focused analysis is even more compelling in the case of universities because in many countries managerial reforms were advanced parallelly with  "autonomy reforms" which sought to alter institutional agents' leverage over the running of universities. Below I present a stream of research directly concerned with understanding the relationship between university autonomy and managerial change and discuss its strengths and weaknesses.

## 1.3. Evaluation and autonomy: inextricably tied?

When States began to turn to university evaluation, they often did so with the premise of safeguarding institutions' autonomy and self-governance powers. This adds yet more complexity to linear assumptions on the rise of evaluation as a direct reach of State power, for it means that the latter needs to be balanced against universities' own self-governance powers, often safeguarded by law. Scholars have given a go at elaborating how different types of external authority influence internal university governance. An array of studies (Berdahl, 1990; Seeber, 2008; Capano, 2008; Shattock, 2009) use Burton Clark's notorious comparative "model of coordination" (1983). The model (see figure 1 below) relates governance at the systemic and institutional levels, observing how these change depending on different countries' external context education relationships.
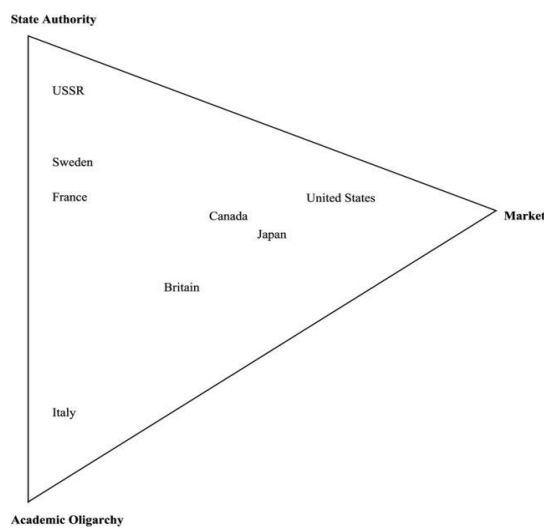


*Figure 1: Clark's model of coordination (1983)*

This has contributed to the definition of different organizational (i.e., "coordination") models in different university sectors: universities have been defined as "conglomerates" (Clark, 1995), as institutions with "tribes and territories" (Becher, 1989; Boffo & Moscati, 1998), or as a "decoupled" system which develops through a combination of adaptation to external pressures and independent internal dynamics (Clark 1983).

A key reason for the popularity of Clark's model is its ability to categorise countries' higher education sectors depending on the dominant type of authority exerted on them, namely State authority, market authority and academic-oligarchic authority (Clark, 1983). Broadly, the "continental model" is where State power is highly centralized, and universities have little internal autonomy in decision-making. This model is closely tied to what has been defined as universities' "oligarchic-bureaucratic" internal governance, for agents reproduce the formal roles set out by the legislator. Italy, for example, has been identified as belonging to this model for there was a significant degree of State centralization and also a strong role of academic oligarchic networks up

until the 1990s. The American and British models, instead, reflect two configurations where the presence of a strong market authority – variable in the two countries - leaves wide margins for internal governance initiatives. The collegial governance typical of the British "Oxbridge tradition", for example, enjoys State authority in terms of receiving funds yet also has significant freedoms for deciding their internal distribution or seek them from the market. In this model, State authority may be shared with the market's, leaving it up to universities to decide their type and level of involvement with third parties. The American, further, is placed at the utmost market-authority end of Clark's triangle, as financing and decision-making are highly marketized and entrepreneurial (see Locke & Spender, 2011). Clark's framework thus importantly acknowledges that universities' level and type of autonomy can differ depending on their relationship with the State.

As States turned to managerial reform of public-sector governance, authors have used Clark's approach to investigate the effects on university governance. Although implicitly, scholars suggest that managerial governance is best set-off with a specific configuration of university autonomy (Capano, 2008; Amaral et al., 2002, 2003), which "autonomistic politics" can regulate. These are in fact that device through which States can implement a degree of institutional autonomy coherent with the managerial turn and the use of evaluation for steering at a distance (Amaral et al., 2002).

For example, following State reforms of the mid-1980s British universities had a weaker autonomy compared to the 1960s-1970s (Berdahl, 1990), whilst Italian ones lived the opposite experience following the autonomy law of 1993 (Boffo, 1997; Boffo & Moscati, 1998)[9]. Scholars' frame for analysing  how institutional autonomy can change is insightful for thinking about how governance may have changed depending on different sectors' specific policy contexts. Similarly, to neo-institutionalist scholars, however, they do not investigate the university context in detail to account for governance, because they are more concerned with State steering. Their marked attention to policy analysis as a means to unveil policy successes and contradictions seldom dialogues with an equally attentive analysis of university-level developments. As I argue, university autonomy is categorised on the basis of its convergence with an 'ideal level' of autonomy which State policy needs to achieve if it wishes to implement managerial logics in universities. This emerges from their conceptual framing. Universities' process of implementation of NPM reforms is explained as follows:

> '[…] faced with a common problem (increasing universities' responsiveness), governments intervened with different strategies: in continental systems, significant autonomy margins

---

[9]That being said, university autonomy in both countries cannot be said to have followed a linear path since the reforms. To give an example, recent laws to the Italian university sector (the "Gelimini law" being the key reform) have been reversing the 'liberalisation' period of the mid-1990s to mid-2000s by conditioning a large part of university powers to the evaluation of the national evaluation agency, ANVUR.

were introduced, whilst in Anglo-Saxon systems there was a recourse to increasing State regulations and legislations (Capano, 2008, p.122).'

As it stands, governments recur to autonomistic policies to reach that right level of autonomy suitable for achieving universities' adherence to new public management reforms. The "common problem" is the ability of States to control universities at a distance and have them adhere to new instruments such as evaluation and performance-based funding, which are assumed to have been already developed by policymakers at the moment of reform. This perspective is an issue for understanding how university agents shaped their turn to managerial governance. It distorts university autonomy into a universal characteristic. In so doing, autonomy is de-socialized: it is not conceptualised as a form of power which universities may have developed throughout time and mobilise or protect, but as being highly vulnerable to political steering. As a result, the way that internal governance changes in different university contexts is reduced to a case of implementing the right policy. Framing the development of university autonomy in this way rids it of its meaning, because there is not consideration of how university agents' use of it throughout time may have established enduring behaviours and interests or a political role within the sector. When Guy Neave (1998) tackled the issue of university sectors' apparent convergence towards a type of autonomy through the reforms of the "Evaluative State" (1988), he later importantly warned that whilst reforms may on the surface appear to be reaching an international ideal and a similar form, they require an extensive historical analysis of contextual specificities. Therefore, suggesting that evaluative States can implement given autonomy policies to achieve managerial governance turns into a reifying act, for,

> '[…] de-regulation and self-regulation are very *relative notions*. If the former suggests a cut back in the control exercised through the legal procedure it is surely relative to the degree of close control previously exercised [which] varies enormously from system to system. The same reasoning applies to self-regulation, which is the institutional counterpart of de-regulation. In short, whilst there is evidence of these two principles begin taken up, the degree and extent to which they involve functions delegated or hived off from central government in the case of the first and taken aboard by individual universities in the case of the second, are very variable indeed' (Neave, 1998, p. 278, emphasis added)[10].

---

[10] Equally insightful was his suggestion that, with reference to the "Evaluative State", '[…] if we embark on it, we find ourselves having to account for those developments which led up to its full emergence. *We are, effectively, thrown back to analysing what led up to it – in short, to the transitional phase* […] If […] we take those earlier developments into account and bring them as additional perspectives to the current account of the rise of the Evaluative State […] we

Whilst scholars surely attempt to delve deeper into how the specificities of different university sectors influence their internal governance and their reactions to new State pressures, they therefore still convey a problematic theorization of institutional agency. Universities are generally categorised according to presumed levels of autonomy, seen as a product of State policy. Autonomy is not studied *as a boundary condition* (Neave, 1995), but reified. In contrast to Neave's approach, autonomistic policies are treated as a function of the rise of the evaluative State, without taking State-higher education relationship as valid elements for adding perspective on the successes and failures of State steering through autonomy. If by intervening in university autonomy States seek to influence the leverage of university powers, then changes to it shall, equally, be studied in terms of established powers which are both subjected to change. Failing to do so brings to the more intuitive conclusion that university transformations during NPM reforms depend almost entirely on States' ability to change the structure of the university sector rather than as depending also, and importantly, on the way that State-university historical interactions shape their relative powers in the sector. University agents' self-governance becomes in this way fully dependant on the State. We therefore get no insight whatsoever as to how university agents' powers made universities "unresponsive" to reform: how did universities attempt to shield themselves from the intentions of State reforms? Did all university agents act in the same way? What were institutional-level perceptions of the reforms, and what strategies were mobilised to protect autonomy, or negotiate it? These questions find no answers, for institutional agency is approached as secondary to State regulations. The way the rise of executive governance in Italian universities is described reflects this. Whilst acknowledging that political prescriptions were very ambiguous in providing clear policy for university reform, Vice-chancellors are still seen as having been subjugated by a strong pressure. They therefore arguably '[…] found themselves *forced* to exert a role much superior to their statutory one' (Capano 2008, p.123, emphasis added). This is in stark contrast to literatures which have indicated, for both the English and Italian cases, that internal, non-State level developments - which remain largely unaccounted for – played a significant weight in altering the roles of vice-chancellors (Vaira, 2011, Neave, 1998).

The different ways of conceptualising institutional agency as NPM emerged show that there are limits to how we account for institutional agents' specific articulation of management. This study tries to open this black box by investigating university agents' articulations of a changing university policy context and the actions taken to influence it in contextual and conjunctural terms. Indeed, regardless of the shared conviction on the verticalization of university powers, the

---

begin to appreciate the full weight and significance of the situation in which we find ourselves' (Neave, 1998, p.268, emphasis added).

historicization of how VCs practically *turned* into managers still awaits comprehensive research. As I have argued, this stems from a variety of implicit assumptions on the inevitability of the process of "managerialisation" at both the State and the university-institutional levels, united by a structuralist view of power. Importantly, it must be remembered that for both the English and Italian University sectors the process of a "verticalization of powers" involved, at a practical level, an extremely difficult process of dismantling, through consensus, the strong roles of academic senates and of collegial decision-making processes to the end of institutionalising a new governance dynamic revolving around evaluation. In the next section, I discuss how the political Marxist approach can overcome structuralist accounts of power by resorting to agency-centred historicism.

## 1.4. Theoretical framework and methodology

The "agency-centred" historicist method of this study is inspired by Political Marxism's insights (Wood, 1981, 1996; Brenner, 1976), first developed during the Brenner debate of the mid-1970s (Brenner, 1976). In the debate, Political Marxists (PM from now on) sought to re-dimension structuralist Marxists' rigid economistic reading of class struggles within capitalist development. Central to PM is the acknowledgement that transformations in capitalism are characterized by class struggles occurring at all levels of society, from the most central to the most marginal[11].To regain perspective on the relational process which power reflects, Wood (1995) for example defined class as a process and relationship. PM scholars thus seek to situate power relations at the centre of their method, and in particular those agents conventionally considered unable to influence systemic processes which are stronger than and above them. Indeed, the intent is precisely to rebalance a systemic understanding of capitalist development and return to an appreciation of how power struggles shape the construction of political-economic structures used to subject groups into specific logics such as competition or discipline. Within the PM tradition, scholars espousing the agency-centred method specifically take a step beyond the broader approach of key PM scholarship. Indeed, agency-centred scholars have highlighted how many PM scholars, whilst advocating for a relational understanding of how structures shape power asymmetries, still refuse to acknowledge the complex historical work that this implies. Accordingly, this leads them to recourse to abstracting those very historical specificities that make a concrete difference for understanding how agents deploy power through structures. (Knafo &Teschke, 2020). The agency-centred method therefore recognizes that

---

[11] Crucial in this was the attempt to enrich the notion of "political" from the increasing gap which scholarly literature was reproducing through abstracted definitions of political economy. As Wood put it, '[…] if the object of Marxist theory is to shed light on the terrain of political action, it can neither ignore these historical realities nor reify them by entrenching the separation of economics and politics that has served capitalism so well in theory and practice (Wood,1981, p.67)'. This has importantly engaged Political Marxist scholars in the attempt to reunite political and economic transformations within capitalism's developments.

society is moulded by unequal power relationships in which dominant agents tend to succeed. However, it does not remain limited to abstracting power relationships through an analysis of the structures which agents inhabit. Rather, it goes in more unfamiliar terrains to explore how, rather than being fixed, power relationships are instead constantly re-defined largely because they reflect a social interaction. As (groups of) agents constantly seek means to acquire forms of leverage over each other, they often have to confront changing constraints from each other. This means that tracing power relations in capitalism is a historical task which should be imbued with specificity and hesitant to abstractionism or generalization. A key insight of this methodology is in fact the acknowledgment that, precisely because it is (groups of) agents that inhabit structures, what are often seen as fixed "structural constraints" should be conceptualised as the reflection of attempts to exert power and therefore studied in terms of struggle between agents. This approach therefore rebalances the idea that structural conditions or contexts can be used generically to explain power relationships because it is concerned with better defining how structural conditions may have similar generic features but will tend to create very specific dynamics which depend on how specific agents in specific contexts mobilize to modify structures in their favour rather than merely reproduce unfavourable structural logics. Thus, the approach seeks to explore the ways in which history can inform our assumptions on the reproduction of dominant power structures and relations (Knafo, 2010). The method is guided by historical investigations of how agents relate to and act in their contexts before, during, and after a given systemic or structural change and constraint is said to occur. This allows to concretely politicise social change, unveiling the actions mobilized by different groups of agents to modify their surrounding context and ameliorate their position when faced with new constraints. Thus, capitalist development is studied in social terms, where '[…] it is the creativity of these actors, both powerful and weaker, which is the key to the story' (Knafo & Teschke, 2017, p.21). This method enables to study managerialism as a process imbued with complexities which call for a re-balancing of conventional notions through historical research on management itself. In arguing for more scholarly efforts at studying managerialism in *its own terms* (Knafo et al., 2019), scholars have made an important methodological step towards an agency-centred reconstruction of management's historical lineages. Whilst crucially aware that managerial practices are now key for debunking changes to global governance (Eagleton-Pierce & Knafo, 2020), scholars are hesitant in affirming that management's claim of efficiency is evidence of its intrinsic power, for example. Quite the opposite, discourses and practices around efficiency may reflect more the struggle occurring to acquire that power rather than an already possessed power. From this standpoint, the relationship between management and governance may acquire different shapes in diverse contexts. Crucially, generic managerial features cannot be used to deduce the relational dynamic that play out when management practices are mobilized. Before making claims on such a

process, it has to be investigated in historical terms. The risk, otherwise, is precisely to reproduce what the agency-centred method seeks to debunk, namely the conviction that we can make specific claims from generic categories. The effort to build on this theoretical framework is even more worthwhile given critical evaluation scholars' studies (Borrelli & Giannone, 2020), goal of analysing how '[…] the language of transparency covers obscurity' (Zarka, 2020, p.42). These scholars, also object, from a Foucauldian standpoint, to studying power in generic terms. Indeed, they argue that analysing the development of evaluation at face-value risks de-politicising how groups' actions and discourses often seek to use evaluation as a disciplining device. Aware that generically critical approaches to evaluation risk reinforcing the claims made by its proponents (Dahler-Larsen, 2020; Eagleton-Pierce & Knafo, 2020), this thesis thus adopts an agency-centred approach to recast the issue of the relationship between managerial governance and university evaluation into a historical investigation of how managerial groups have formed in different contexts and how they have attempted to construct new power relationships through evaluation.

# Chapter II

## Systems Analysis: a decision-making science

Having discussed what is problematic about framing managerial practices as inherently part of a neoliberal project, this chapter opens the historical research of this thesis, beginning with a historical investigation of managerial practices in their own terms. I here introduce key managerial developments occurring between the immediate post-war period and the mid-1960s in the US, contextualising the build-up of a management – educational evaluation nexus which was embodied by later educational evaluation indicators. As has been mentioned, this thesis explores the scholarly argument that the conceptual frame of performance assessment (Knafo et al., 2018) and educational evaluation (Bürgi & Tröhler, 2018, Tröhler, 2013) criteria has its roots in the Cold War period, when emerging ideas around social planning spread and were piloted to enhance educational productivity. This was done either directly through educational evaluation or less directly through new resource-allocation frameworks which allowed to tailor investments to the most productive – and to the most politically relevant – educational fractions. At a political level, the diffusion of such new concepts was enhanced by the Eisenhower and Johnson administrations' search for innovative and decisive policies and strategies (Amadae, 2003; Rudolph, 2002) – especially in the industrial, military, and defense sectors – which could transmit and strengthen the science-based industry and the 'radical disciplinary mixing' (Edwards, 1996; Wyn-Jones, 2016) developed during the second World War to the U.S.'s Cold War effort. Both administrations funded and seconded experimental projects – often devised by specific networks at notorious foundations such as Ford, Rockefeller, and RAND or through the university programs which these helped to fund and organise – on computer technologies (Edwards, 1996), industrial organization and educational innovation and curricula (Rudolph, 2002). As this and the next chapter show, in fact, education soon acquired an increasingly central role in the US's Cold War technological build-up as that sector capable of training modern and skilled future scientists and managers. It is largely during this thrust that US administrations thoroughly reformed education in the late 1950s and 1960s, and that intellectuals and researchers interrogated themselves on methods and theories for squeezing innovation and productivity out of it. At an intellectual level, the notion that society was a system whose complexity could be controlled and manipulated to steer society into specific political agendas was the child of a broader 'systems approach' which conceptualised reality like an organism or a machine; namely as a whole composed of multiple – and sometimes autonomous – sub-wholes whose interaction determined the functioning of the entirety. Amongst the various and highly diversified approaches which emerged from this one standpoint, this chapter looks at a branch of

systems thinking, the "systems analysis" framework, investigating its influence on the rise of new notions and methods around educational governance. For some (Bürgi, 2016a), systems analysis stimulated an "input-output" framing of educational governance, defined also as "output-governance" whereby educational decision-making came to rely importantly on quantifications of endowments (inputs) and intended vs actual outcomes (outputs). Borrowing from positivist methodology, this framework developed in no marginal way in the US defense sector and conceptualised progress as a case of expanding intellectual and mental abilities for deconstructing the way that the reality operates. It rejected the idea of a finite boundary of human action on and in 'the real world', heralding instead the view that reality is composed by a set of definable, quantifiable, and modifiable parts which, in their interaction and interlocking, form the entirety of reality itself, the "whole" and the "organism". As a corollary, human action could and had to better capture how these parts functioned on their own as well as when interacting with each other, and it was new knowledge which would guide action towards the observation and manipulation of social complexity. As this excursus will show, the methods proposed by systems analysists reflect their conviction that acting on society was, in broad terms, a matter of gathering sufficient information about it. Based on the conviction that human knowledge of reality is divisible from reality itself, systems analysts pushed for information gathering, insisting that this was the crucial perquisite for the construction of instruments – budgets, performance benchmarks, educational attainment levels – that could "return" information to decision-makers – managers, policymakers, educators. This chapter discusses how this rationale is materialised in two specific systems analysis models: the Planning, Programming and Budgeting System (PPBS), and the Context, Input, Process and Product (CIPP) model because of their influence on the construction of university evaluation criteria in England and Italy, which I discuss in chapters 4 and 5. Before dealing with these model's diffusion and articulation, however, I here first introduce their creators, technical features and conceptual axis. Whilst sharing the theoretical standpoint of systems analysis, these models represent different methods for steering education sectors. Indeed, simultaneously to experiments in the US department of defense, PPBS was piloted as a device for optimising the use and allocation of State investments in education (Hirsch, 1965), whilst CIPP invoked an information feedback loop between schools and policymakers through constant evaluation (Stufflebeam, 1968a,1968b, 1983; Guba & Stufflebeam, 1970). Both models pioneered practices which could, in line with the systematic-scientific approach, inform decision-making through the quantification of educational inputs and outputs. Whilst this view now appears close to standard practice, it was radically novel when it was first pioneered, as policymakers tended to finance education rather freely leaving education professionals the autonomy to operate without much scrutinization. As these models were piloted, they mobilised resources and new interests, acquiring momentum

I first discuss the PPBS model. Whilst CIPP emerged specifically for education, PPBS has an indirect relationship with education because it was first devised for military strategy purposes. After broadly laying out the key features of PPBS as a budgeting model, I discuss the key innovations of the department of defense which led to such features. Tied to this, I show how PPBS's intellectual authors at RAND – as well as that of scientific-military cadres more broadly – actively abstracted the model from its military origins in the attempt to diffuse its use within US educational policymaking during the Cold War period (Rudolph, 2002; Bürgi, 2017a, 2017b; Clowse, 1981). This was the case, for example, of the National Defense and Education Act (NDEA) passed in 1958. As I argue, this was a successful move in the context of the first decade of the Cold War, when a military-academic complex (Bürgi, 2017a) matured in response to the USSR's technological advantage (Amadae, 2003; Rudolph, 2002). I then turn to the CIPP model. I briefly consider its emergence  as a response to those very reforms enacted thanks to the role of scientific-military consultancy groups during the Eisenhower and Johnson administrations (Madaus et al., 1983), and look at CIPP within evaluation theory. I discuss how its proponent, Daniel Stufflebeam, used systems analysis concepts in thinking about CIPP (Stufflebeam, 1969a, 1969b, 2001), advancing information management at different levels as key to educational decision-making. As politics demanded educational innovation and forms of accountability, the CIPP model was constructed to help schools attain new funds coming from policy. Its proponents defined it as "program evaluation", a model first and foremost concerned with translating political demands into operational guidelines for schools. The chapter focuses more extensively on program budgeting than on CIPP. This is because differently to the latter, PPBS was also used to steer public administrations into efficiency across Western Europe between the mid-1970s and 1990s (Saint-Martin, 2000; Rebora et al., 2016). For this research, this means that analysing Italian and English policymakers' use of (or reference to) the model – as is done in chapters four and five – can help to test the dominant argument that public institutions' managerial turn was the effect of policymakers' own managerial turn (Hood, 1991; Pollitt, 2004,2016). Understanding PPBS's development during the Cold War period can thus unveil how the contours of new public governance concepts may not have emerged from neoliberal political cadres (Knafo et al., 2019; Bürgi, 2016a, 2016b) but rather from managerial ones; opening a historical revision of how the early stages of new public management concepts may have first manifested themselves outside the State.

### The development of Systems thinking

Systems thinking is a broad approach populated by ways of enquiring real-world problems by framing society (or nature) as a whole that is composed of multiple components which, when coordinated, allow a "system" to function. The notion of "system" for referring to society has

diverse theoretical parents, such as cybernetics, general systems theory (G.S.T.), and organismic biology (Checkland, 1983). In biology, systems thinking developed between the second half of the 19th century, whereas it evolved in engineering, sociology, and management between 1930s and mid-1960s. This latter wave was influenced by the methodological innovations of a group of U.S. defense researchers who re-formulated existing methods of military strategy design. Central to their work in systems analysis was the exercise of modelling new, rather than existing, logics for achieving a given outcome. This meant that organizations were conceptualised as systems composed of "sub-systems" (Checkland, 1983) whose functioning could be steered through the manipulation of information on their activities. These steps further evolved systems analysis's claim as a science of decision-making. Between the 1930s and the 1950s, in fact, problem-solving practices were re-defined to the point of becoming methods for designing new solutions through new means, rather than working with existing means, turning the conception of institutional governance into an exercise of *strategic* decision-making. I discuss these developments in turn.

Firstly, the rise of systems analysis occurred when defense researchers innovated on military operations research (OR) by modelling for strategic purposes and not for tactical ones (Knafo et al, 2018). Concerned with optimising existing endowments such as radars, British Royal Air force officials – who first developed operations research in the 1930s (Checkland, 1983) often carried out OR by simulating real-world operations in models and experimenting on them in order to then transfer the results onto real situations (Checkland, 1983). The key aim of these "on the ground" experiments was to optimise the use of existing military endowments. At the RAND corporation, however, problem-solving was re-framed as more than a mere optimisation practice (Knafo et al., 2018). In fact, as the methods of British military travelled to the USA, they were developed by a group of ex-veterans and researchers at the RAND corporation (Amadae, 2003). Rather than experimenting with existing means to maximise their use, the "defense intellectuals" (Amadae, 2003) built methods for creating new paths of action to achieve given military objectives, creating a gap between operations research and systems analysis. Although systems analysis also sought to optimise military capacities for defense actions, the key difference with operations research lies in the methods: systems analysis began to plan and design for the construction of *new* systems, whereby operations research was concerned with optimising *existing* systems (Checkland, 1983). The "new systems" constructed were initially mathematical models (Amadae, 2003; Knafo et al., 2018) which were much more complex than existing O.R. ones. If in O.R. models presented a means for simulating real-world situations "outside the laboratory" (Checkland, 1983) to then test for improvements in given military tactics, through systems analysis models became a way to design ex-novo military actions altogether. In this way, systems analysts extensively used mathematics and developing computer technology to construct possible alternative plans of action

and their associated alternative outcomes. This thrust began with attempts to go beyond the optimisation of endowments and developed into a concrete method of thinking in strategic, rather than in operational-tactical, terms. Whilst this difference may sound as a mere variation of a single method, it actually represents a radical departure from the idea that military operations were merely about making the best with given endowments at hand, reflecting instead an active search for ways of concretely controlling the multiple variables of military decision-making. The fact that systems analysis developed following the end of the second World War, further, led researchers to attempt to control broader variables related to military strategy which were conventionally controlled by US congress (Amadae, 2003), such as military resources. A direct result of this was that resource-allocation considerations became increasingly central to the modelling of system alternatives. As researchers gathered increasing amounts and diversified types of information, they elaborated complex models that tried to keep account of much more complex sets of variables – often constructing new ones as part of the search of alternative courses of action. Systems analysis models thus no longer needed to concretely reproduce existing military contexts as loyal as possible to "on the ground" situations but were complex matrixes which extended the possible decisions by intersecting and re-shuffling diverse military inputs and outputs. By expanding the horizon of ways for achieving given objectives, modelling developed into a concrete exercise of strategy formulation, where objectives could be re-formulated if alternative combinations of variables were considered more efficient. As RAND researchers presented their case for new strategies or considerations through systems analysis, they promulgated their new approach through their new evidence, bypassing intellectual debates over the validity or assumptions of new models because these could claim legitimacy on the basis of their scientific complexity and, thus, "superior explanatory power" (Amadae, 2003, p.72). The departure of systems analysis from operations research thus expanded the scope for the use of defense methods outside their context of origin, emerging as a fervent way to meet complex social needs thanks to the highly publicized and slowly crystallizing public belief that it was scientific ability which could best identify and assemble the variables of diverse phenomena to reach social, and not only defense, goals (Rudolph, 2002; Kevles, 1987).

Secondly, this departure from operations research methods turned information gathering and manipulation into a key practice feeding into decision-making, if not *the* practice of decision-making altogether. This is because the insistence on the consideration of alternatives justified the demand for more and more diverse types of information, giving analysts the power to reach out into the multiple divisions of the department and peek into the micro-dynamics of the department's functioning. Furthermore, as the number of possible variables and their combinations exponentially increased in the freer approach of systems analysis modelling, so did institutional decision-makers –

such as the secretary of defense Robert McNamara – see the opportunity of using *information modelling to acquire greater leverage for implementing their preferred plans of actions or resource allocations*. The way that systems analysis was perceived outside the defense context in the late-1960s confirms its gradual establishment as an instrument in the hand of institutional top-hierarchies, and, hence, as a new instrument for governance.

## 2.1. PPBS

Performance assessment practices usually associated to the rise of the NPM paradigm in 1980s have largely inherited the Planning, Programming and Budgeting System (PPBS) which developed from systems analysis in the late 1950s and early 1960s. PPBS was developed as a method for constructing a budget and plan resource allocation and/or investments, however it has historically acquired much wider scopes, amongst which a key influence on policymaking processes across public administration, beginning with Anglo-Saxon ones (Wildavsky, 1961; Saint-Martin, 2000). The relationship between PPBS and the rise of evaluation practices used for steering institutions, however, is not straight forward, largely because of the fact that PPBS itself was remoulded many times in diverse contexts, (Knafo, 2020), with new and different features fitting into the umbrella term of performance assessment. Although PPBS's application in State walls has been studied to capture changes to policymaking practices (Wildavsky, 1961), it has scarcely been considered a valid object of analysis for understanding changes in the governance of non-State institutions such as universities. This is even more surprising given that PPBS is widely recognized as a managerial framework whose principles are applicable to the planning processes of *any type of organization*, be it a private enterprise or a public institution, producing goods or services. In relation to its military origins, PPBS was quickly abstracted into a broad framework concerned with assisting the running of an organization's entirety by intervening at its core, namely the budget. What PPBS is and what it is not, to cite one of its key developers (Novick, 1973), is best understood through a comparison with a line-item budget, which it sought to radically modify, capturing the attention of old and new public administration scholars who debated on the strengths, weaknesses and implications of its implementation in public resource planning[12].

State budgets signal political agendas through the resources allocated to different public sectors. Line-item budgets were conventionally divided into departmental/ministerial budgets, designed and officialised by ministers responsible for social policy in health, education etc. A State's multiple arms would thus take policy decisions based on public resources available and account for them in resource terms, allowing politicians to ensure that resources spending occurred

---

[12] Novick (1973) suggests that PPBS was pioneered also between the mid-1920s and 1930s in American industry, namely by Ford Motors and the Chrysler Corporation, to face financial and management problems.

within given expenditure limits and making State ministries and ministers accountable for both the amount and

destination of public resources spent. The line-item budget carries this name because of its structure: items would be listed separately (i.e., personnel, maintenance) with their associated costs, often expressed in cash terms (although they could also be expressed in volume terms to signal resources not yet held by ministers (Wildavsky, 1978)). Conventionally, line-item did not involve a complex planning effort, as policymakers relied on the assumption that ministries could rely on yearly increases resources, meaning that yearly budgets were often predicted as incremental. The shift to program budgeting, however, set-off political struggles and resistances because it could lead to a significant reshaping of this traditionally incremental nature of State budgets through the way that program budgeting allows to reshape the goals of expenditure in the first place (Wildavsky, 1961, 1966). This aspect is attributable to PPBS's systems analysis lineage, which brought new methodological axes for approaching the concern of resource planning. I first discuss how the program budget allows for a redefinition of objectives and then turn to the military lineages of this legacy. In contrast to the line-item budget, the program budget '[…] seeks to define and achieve outputs, rather than appropriate inputs' (Novick, 1973). It groups resource use (i.e., costs) in terms of the activities which those resources involve and, thus the results that such activities *lead to*. This defines its "output-oriented" approach compared to how line-item budgeting divides costs generically in terms of what those costs *represent*. In other words, if line budgeting would convey the *sources* of costs, in program budgeting costs reflect all those activities undertaken to achieve a given outcome. Once again, what appears as a superfluous reshuffling of the same variables is a witting and deep normative difference. The example in the table below helps to clarify this difference by comparing a traditional budget with a program budget (see figure 2 below). Specifically, the table below is referred to as a conversion matrix or "crosswalk" which depicts the steps of converting a line-item budget into a program budget (see Novick, 1973).

As figure 2 shows, a line-item budget conveys salaries and wages paid by a federal agency by dividing salaries according to the departments where workers work, in this way accounting for the costs incurred by that department. So, for example, health department employees' salaries are listed in a distinct cost category to those of the federal water department employees. In the program budget, this "departmental division" is voluntarily overpassed, such that salaries of employees from different departments working on the same activity – the "program"– can be grouped into one single cost voice. Following the example, this means that employees in the health department and those in the water department which work on "water sanitation" reflect a single cost, that of "water sanitation personnel", incurred for achieving the final outcome, and not two separate costs for different departments. Interestingly, developers of PPBS pointed out that this division of costs was

*not* geared at efficiency: '[…] traditional budgeting is aimed largely at efficiency […] program budgeting sets its sights on larger purposes, the objectives of an organization' (Novick, 1973, p.12). Indeed, analysing PPBS in terms of efficiency can be rather misleading for it can shift the focus away from how efficiency was historically re-defined through program budgeting as a matter of defining objectives. Thinking of efficiency to understand PPBS is also a risky conflation of how its use differs in scientific management and in managerial governance (especially that in public institutions. See Knafo et al.,2018). For if scientific management pursued efficiency through the *maximisation of labour and/or capital production*, the use of PPBS for managerial governance frames efficiency as a *matter of coherence with planned organizational objectives*, meaning that in PPBS efficiency turns into an achievement dependent on resource planning and not on output targets or production benchmarks – as is the case in the classical example of timing production along the assembly-line. In its re-organization of activities in terms of programs, PPBS reflects the governance-oriented use of efficiency. Quite differently to an "accounting form" of the line-budget as an informative instrument (Sudama,1977), then, PPBS was devised by its developers as an instrument which would allow to influence an organization's core, namely its *decision-making process* (Novick, 1973;Sudama, 1977;Shick, 1966) seeking to intervene in defining organizational objectives, its dos and do nots. But how exactly can PPBS assist the formulation of objectives and thus decision-making on the allocation and efficiency of resources? Three key methodological novelties explain this: the

program format, a systems analysis approach, and the extensive use of information gathering and reporting (Novick, 1973).



FIGURE 1 Crosswalk: Traditional Line-item Budget to New Program Budget

Fig. 1 reproduced from *Management: A Book of Readings* by H. Koonz and C. O'Donnell, 3rd ed. copyright 1972 by McGraw-Hill Inc. Used with permission of McGraw-Hill Book Company
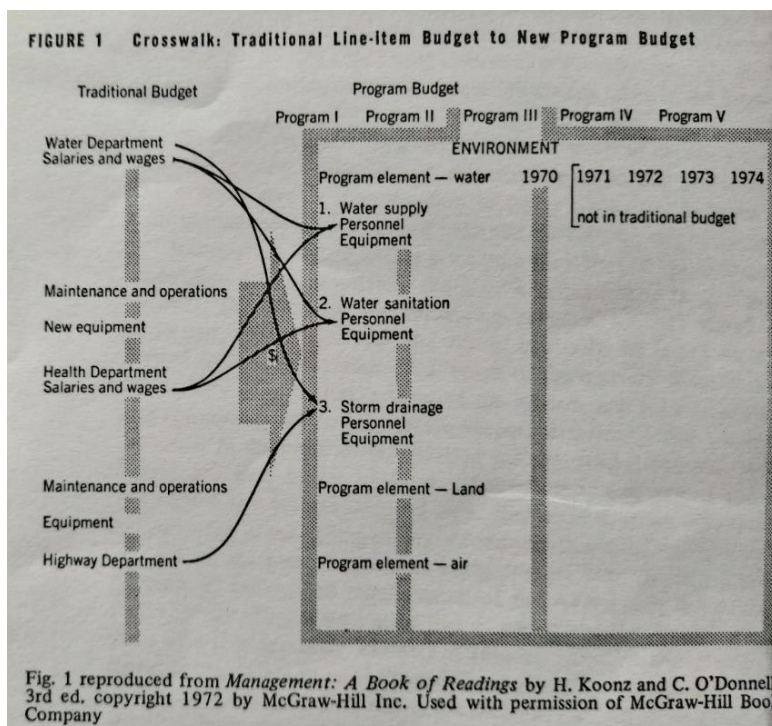
*Figure 2: A line-item budget vs a Program budget (Novick, 1973)*

Firstly, conceptualising costs as part of a given program – i.e., using a "program format" – allows activities to be grouped together under the program's heading, meaning that costs incurred for activities can be directly related to their intended output, acting as a more concrete device for analysing resource use  than if, to continue to example made above, salaries are expressed in more general terms. For decision-makers, this "revolutionary" (Novick, 1973) aspect of concretely relating salaries to activities can open the possibility, for example, of scrutinizing the extent to which salaries fully reflect the cost of achieving their intended outcome. This turns salaries (as other organizational expenditures) subject to potential manipulation by agents who define program objectives. Programming is thus that exercise of relating activities to their intended goals, and feeds into an organization's resource-planning decisions. In the budget, the program layout allows to decide resource allocations depending on the objectives, rather than through an incremental rationale.

Secondly, systems analysis in program budgeting denotes efficiency as the successful result of the right combination of organizational variables which allow programs to be achieved, rather than as their maximisation. In program budgeting, the military considerations discussed above are reproposed in resource-allocation terms. Indeed, an organization is framed as a system, its departments as sub-systems, and their interaction is analysed systemically through the analysis of the relationship between sub-systems' activities. Indeed, program budgeting follows the systems analysis approach in its two central claims: firstly, that any decision influencing the system must be the optimal one. Secondly –and tied to this – that the achievement of the optimal –efficient – decision strategy for pursuing a given program objective first requires a scientific assessment of *whether there are unexplored, more efficient, alternative strategies that could be pursued*. This latter aspect is key to how PPBS enhances the re-formulation of organizational resource allocation objectives and plans for their pursuit. Specifically, '[…] quantification is sought wherever possible' (Novick, 1973) so that no "predetermined path" is taken. PPBS developers advocated –along with a dose of managerial creativity – the use of cost-benefit analyses to assist the drafting of alternatives, because costs and benefits of existent and alternative resource allocations could be compared. This makes the constant re-shuffling of variables or endowments a key part of the process of drawing up an organization's program budget. The centrality of this approach for the purpose of optimisation explains why many PPBS proponents emphasize the importance of planning and, further, suggest that it is a lengthy process which can take up to a year (Hirsch, 1968; Lockwood & Fielden, 1972).

Once more in line with the systems analysis framing, the process of rationally finding and optimal resource allocation strategy calls a third aspect of PPBS into question, continuous data gathering. This type of data is much more comprehensive than that used in line-item budgets because it intends to cover every activity carried out within the organization, or as  many of them as

possible. This systematic gathering of information acts a supporting device to the exercise of cost-benefit comparisons. Yet, it also acted as an exercise of power on an organization's departments because data was used to construct cost-constraints to ensure that departments worked within specific budgets. This is the "budgeting" element of PPBS, where information on activities such as man-hours and resources used – allows budget planners to monitor organizational costs, and, secondly, extend top-down control over sub-systems' resource use. Critical studies (Knafo et al., 2019) show that budgeting was developed precisely to ensure that the department of defense's lower levels would align with the top-levels' articulation of military strategy, and that the transmission of cost information from the former to the latter was used as that instrument for doing so. This step went beyond the use of managerial benchmarking to reach given output levels, as information can be used to re-channel resources *away* from given areas and onto new ones in future plans – or in 'flexible plans' – if activities are considered incoherent with strategic objectives. In this way, information gathering in the PPBS framework can effectively function as that instrument allowing decision-makers or budget planners to exert (punitive) power. Budgeting thus did not limit top-levels to a monitoring exercise concerned with ensuring that a given set of resources are used in the best possible ways, it freed them of this constraining role precisely because it was framed within a systems analysis perspective. Through a discursive escamotage of cost-efficiency, budgeting can concretely turn resource-management into a power of strategy (re) definition – and thus not into a moment of monitoring or assessment of productivity. This crucially functions only if lower levels can be effectively pushed into providing information for top-levels to work with, and if they can be steered – through notions of accountability and participative management – to reporting through cost-benefit templates, so that planners can then receive information that is readily modellable according to *their* (changing) priorities.

Program budgeting thus represents a framework which can allow top levels of an organization to *govern* resources directly through the recourse to (resource) information, the modelling of resource use through different plans, and the subjugation of workers' resource use to scrutinization and evaluation through the program format. In this way, program budgeting reflects a model in which governance and management begin to converge, and the traditional boundaries of scientific management as an activity of production monitoring are surpassed (Knafo et al., 2018) through systems analysis. For these reasons, I critically argue that PPBS's 'program-oriented' budgeting framework is best understood as that prerequisite for legitimating decision-makers' ability to re-define objectives more arbitrarily. This underlying scope has its roots in the interests of its constructors, the RAND corporation's "defense intellectuals". They were able to portray the subjective judgements being made within the PPBS framework as a scientific expertise which was limited – until the mid-1960s– to its developers. The next two sub-sections explore how defense

intellectuals exploited this marginality to construct a new expertise around public resource management for themselves.

### 2.1.1. The military origins of PPBS

As mentioned above, PPBS is a military innovation of strategy optimisation techniques used in the US department of defense (D.O.D from now on) (Amadae, 2003; Checkland, 1983; Knafo et al., 2019). To be precise, it was at the end of the 1950s that RAND intellectuals began to '[…] challenge the division between management and strategy' (Knafo et al., 2019, p.10), starting from the defense sector and then spreading this innovation outside it, in the process developing into a new policy elite (Amadae, 2003). Previously to its establishment as a think-tank in 1946, the RAND corporation was a project wanted by the leader of the Douglas Aircraft company, General Henry Arnold, and the U.S. army's air force researchers. Aware that the end of World-War II would imply a declining space for their role, this group tried to maintain an active influence over the defense sector by acting as defense – and specifically air force – consultants (Amadae, 2003). The initial success of the project was evident when, two years following its establishment, the corporation was populated by 255 employees and endowed with an annual budget of $3.5. million (Amadae, 2003). At the D.O.D., RAND intellectuals consulted as cost-analysts, and by 1962 they were advocating for a more comprehensive use of program budgeting at the department, in government and in industry (Novick, 1962)[13]. When a new secretary, Robert McNamara, and his assistant secretary, Charles Hitch, entered the D.O.D., they mobilized RAND researchers to develop program budgeting. As I argue, this occurred for two key reasons. Firstly, to push lower levels to act coherently with the secretary's defense ambitions. Secondly, to free the department from funding constraints coming from the State. Both these interests were mobilized as the Cold War hit the globe in the late-1950s.

Concerned with optimising their military strategies, up until the 1960s the department of defense had tried to find new methods. Defense intellectuals working with systems realised that, although they could use systems analysis to decide on the ideal military strategy after having modelled alternative ones, they could not intervene on how lower levels then concretely operated when they were "on the ground" (Knafo et al., 2019). and this influenced the weapon development underway during the Cold War (Amadae, 2003). This was challenging, however, in a context where, once a military strategy was presented, lower levels could then somehow autonomously use the means put aside for concretely implementing strategies. This often caused mismatches between a strategy's' definition and its practical implementation, pushing defense analysts to try and more

---

[13] Developers defined planning in line with the systems analysis approach as '[…] the selection of courses of action through a systematic consideration alternatives' (Novick, 1962, p.5), and programming as '[…]the more specific determination of manpower, material and facilities necessary for accomplishing a program' (ibid).

concretely control how lower levels acted. This was first done at the turn of the 1960s by asking them to operate within cost constraints and meticulously report resource use in the hope that operations would play out close to identically to how they were strategically defined by top levels. It is in making this step that RAND analysts radically innovated on existing military strategy definition, constructing the very features of PPBS discussed above. Indeed, it was at the D.O.D. that program budgeting developers (Novick, 1962) first used the cost-analysis narrative to discipline the department's lower levels. The implementation of PPBS facilitated this, as it urged the definition of objectives and the constant monitoring of the information being provided to assess their achievement. In this way, PPBS was used to ensure that future military strategies would be implemented coherently with McNamara's personal view of the department's goals, which were devised with RAND cost-analysts such as David Novick. As he put it,

> '[…] in the past, the department of defense has often developed its force structure by starting with a budget and sending it off in search of a program. Our new system of program packaging has reversed this procedure, by first determining our over-all strategy, then fitting the hardware and the manpower to those objectives' (Novick, 1962, p.7).

Interestingly, developers further tried to strengthen strategic definition following this new hierarchy by emphasizing the importance of "program change". This was defined as a feature allowing for adaptation to external changes, yet, in practical terms, it was a step towards '[…] McNamara's objective [….] to have at all times an up-to-date, five-year force structure and financial programs' (Novick, 1962, p.7). Program change went beyond the definition of budgets, because it justified a continuous information inflow, with the program budgets themselves being revised three times a year in order to reach the top levels and '[…] meet the Secretary's summary and analytical needs' (Novick, 1962, p.10).

Borrowing from pragmatist insights in critical political economy which analyse power relationships by stressing that agents make instrumental use of structures (Konings, 2010), I argue that, following the launch of Sputnik 1 and 2 in 1957; the D.O.D. exploited the Cold War context to advocate for greater endowments and decisional autonomy. Whilst placing emphasis on cost-effectiveness, the real concern of program developers was to keep the program budget an open-ended and developing methodology which would reserve decision-making powers to the experts at the department of defense, rather than concretely responding to the management of costs of developing defense technologies. Had the D.O.D. been truly and radically oriented at cost-efficiency and thus concretely cost-efficient as its reports advocated, then it would not have had the

sufficient grounds for justifying its requests for more State funding in defense –although the rising Cold War tensions provided a good starting point. This explains why, eventually, the D.O.D. was able, through RAND researchers, to have budget ceilings removed (Amadae, 2003). By stressing on the national security needs and on its objective and scientifically vanguard decision-making tools, D.O.D. analysts could now actually develop defense policies for themselves without as strict Congressual guidelines on resources and their use (Amadae, 2003; Wyn-Jones, 2016). This incredible achievement owed much to analysts' narrative on cost-efficiency and goal-pursuit which empowered the planners and disciplined the military officers working for them. This explains how program budgeting became an instrument of policy analysis which could benefit the emerging class of planning experts at the department of defense. Additionally, this political dimension of PPBS importantly signals how efficiency itself is not a sufficient window for analysing how PPBS turned into a decision-making method.

### 2.1.2. Piloting program budgeting in US education

Simultaneously to these steps at the D.O.D., RAND analysts also reached outside it by suggesting PPBS be used as a new resource-allocation framework in non-defense departments (Novick, 1954, Hirsch, 1965), proposing to act as objective policy analysts who could help the State to steer resources according to its political priorities. As the military strategy scenarios of the defense sector turned into resource allocation scenarios in social policy, RAND intellectuals abstracted defense-tailored developments in the attempt to universalize the use of Program budgeting and extend demand for their consultancy. This meant pushing policymakers to ask themselves unprecedented questions on policy goals and policy efficiency which only RAND analysts were able to "rationally address". The systems analysis rationale of "considering all options" was in this way articulated as a matter of thinking about policy through a grid of multiple and diverse objectives each related to different combinations of inputs, i.e., costs.  RAND analysts succeeded at piloting Program budgeting in US education department not only by insisting on its objectivity, but also by taking advantage of the rising Cold War to foster the U.S.'s intentions of building scientific technology and manpower for keeping up with the arms race. The RAND corporation reported its vision and experiments with PPBS in State departments in a series of key reports. These were David Novick's *Efficiency and Economy in government through new budgeting and accounting procedures* (1954), Charles Hitch and Roland McKean's notion of "the economics of defense" *in The Economics of Defense in the Nuclear Age* (1960)*,* and Joseph Kershaw and McKean's notion of "the economics of education" *in Systems analysis and Education* (1959), amongst others (Hirsch, 1965; Novick, 1962, 1973). In the latter, education was defined as an "intermediate input" in the process of achieving economic objectives, and not as an output *per se* (Kershaw & McKean, 1959). Whilst the notion of the economics of defense focused on the transformation of economic power into military strength at

a macro-economic level –McKean, Kershaw and Hirsch advocated for the use of program budgeting within schools, seen as a step towards the reform of the school 'system' as a whole (Hirsch, 1965; Kershaw & McKean, 1959)

The RAND's earliest study on education traces back to 1959, when Kershaw and McKean published "*Systems Analysis and Education*" . Their analysis made novel, quantitative comparisons between multiple education systems, beginning with a study of Santa Monica's Unified School District – coincidentally close to the corporation's headquarters. Loyal to their systems analysis formation, the authors used information gathered in schools to model how the variation in inputs produced different outputs[14], mostly IQ tests on pupils[15] –. The study emphasized the definition of educational outputs, a task that the researchers themselves were pioneering for the first time and had no previous experience with, leading them borrow extensively from military strategy. This experimentation crystallized at the 10-day "Woods-Hole conference" of 1959, organised by the RAND corporation, the Air Force and the National Science Foundation (Bürgi, 2017a; Rudolph, 2002; Tröhler, 2011), which for some represents the genesis of a *military-academic complex* (Bürgi, 2017a) where military research methods were framed as a viable device for educational reform in a Cold War context where the US's production of defense technologies relied crucially on the ability to produce more and better scientists, engineers and technologically skilled students and workers. At the conference, in fact, RAND intellectuals suggested that its Santa Monica school district study was a valid example of a new path for the education sector (Bürgi, 2017a). Starting from new and radically different objectives for the education to those of the pre-war period, defense intellectuals "recycled" older, military decision-making methods (Rudolph, 2002) to have education respond to national security concerns. The organizers agreed that education *required*  military models and methods to set itself new objectives more apt to the new challenges of the post-industrial society and the urgent necessity of steering the country's priorities to the Cold War effort. In this way, the emphasis on objectives did not merely involve assisting educators in deciding how their capacities could be improved to achieve *their* educational objectives. Instead, at the Woods hole educational objectives were all-together re-defined to steer education to the State's existing intents of connecting the development of defense technologies to educational development. Not coincidentally, in fact, Eisenhower's administration welcomed the National Science Foundation's use of systems analysis to draw up the new goals of education (Rudolph, 2002), eventually

---

[14] The authors defined their research in the following way: '[...] our plan of *attack* involves determining the changes in output resulting from varying the inputs, estimating the changes in cost resulting from varying the inputs, and finally comparing the alternatives.' (Kershaw & McKean 1959, p.7, emphasis added).

[15] It is interesting to point out that authors insisted on using this indicator due to its easier calculability and quantifiability whilst acknowledging the 'volatility' of IQ tests and pointing out that it was a weak instrument for analysing the impact of educational processes on pupils' overall intelligence. However, their scientific method remained justifiable on the basis that the contradictions of IQ tests would dissolve given that '[…] our concern will be with large numbers where individual errors will balance out' (Kershaw & McKean 1959, p. 11).

articulated in the National Defense Education Act (NDEA) passed in 1958. This reform sought to construct "brainpower for the Cold War" (Clowse, 1981) by injecting of over a billion dollars in grants, loans, and services for all levels of education – particularly in the Science, Technology, Engineering and Maths (STEM) subjects. In the hope that this reform would radically intensify the future production of more and younger future scientists, engineers, and mathematicians, Eisenhower's administration responded to the US's lagging educational position vis-à-vis that of the Soviet Union where, for example, STEM students graduated at 22 and were quickly hired in industry (Onushkin 1975, Vol. IV). RAND's suggestions thus arrived as a direct response to the US State's own reaction to the launch of the Sputnik:

> '[…] Whether for State security or for the entire national, primarily economic development - education was stylized as a decisive instrument for defense, economic and social policy. This paved the way for the national and international dissemination of systems analysis' (Bürgi, 2017a, p.51, author's translation)

Shortly after the end of the NDEA's enactment period in 1961, RAND intellectuals stepped back into the State's education affairs. This time, they focused more in detail on the nexus between education and the capacity of defense technologies by peeking into the nature and quantities of State investments in education. In 1965, W.Z. Hirsch's '*Education in the Program Budget'* experimented the use of PPBS with the specific concern of understanding the extent to which the channelling of education funding in military and defense research could be re-designed to fit more tailored educational objectives (Hirsch, 1965) coherent with the previous reform. Hirsh consulted the department by calculating and quantifying its efficiency and assist it in improving its resource allocation according to its new objectives. According to Hirsch, it was by implementing the PPBS framework that expenditures could be re-allocated in line with education programs. Departmental use of program budgeting allowed to work on three main aspects: spending efficiency, objective setting (i.e., decision-making on what education ought to be steered for in the first place), and the tailoring of educational resources to specific objectives (Hirsch, 1965).

To improve efficiency, Hirsch advised the education department to centralize its highly dispersed funding –which was distributed through 42 agencies (Hirsch, 1965, p.2)–, for program budgeting would help to regain control over the funds allocated to them. Hirsch pointed out how the department was being financially "exploited" by receiving agencies, as '[…] many research and training funds are awarded to help specific agencies accomplish their missions.' (Hirsch, 1965, p. 26). It followed that program budgeting was presented as relevant for '[…] exclud[ing] from the education program budget those research and training funds that directly relate to specific

government missions other than education' so that non-educational programs would be '[…] grouped into the program budgets other than education' (Hirsch, 1965, p.27), becoming other departments' financial burden. Given that most of these agencies carried out defense or military scopes, it was therefore suggested that costs be transferred to the defense department. In other words, Hirsch proposed PPBS as a valid method for deflating the 'real' costs of education, allowing the department to make the case for more public funds - a direction which, in those years, the State was willing to take.

With regards to defining the department's objectives, the context of the Cold War was once again manipulated to publicise PPBS as crucial in defining the '[…] expression to the nation's position towards education' (Hirsch, 1965, p.28). As of 1963, the majority of the education department's funds were still largely divided according to the scopes of the NDEA. Of the total grants distributed by the department, 6.7% were destined to Higher Education, 49.5% of these were for military academies and veterans' education (Hirsch, 1965, p.22). The major funders of basic research and research facilities, covering over a third of such expenses, were the national institutes of health, the department of defense, the national science foundation, and the atomic energy commission, amongst others. This constituted a problem, namely that educational objectives were not directly and autonomously set by the department but were being "shared" with other major funders. If the department sought to better *control* educational expenditure, therefore, it had to regain some form of central leverage over educational funding. This necessitated, in part, questioning the extent to which third parties really needed such funding. When Hirsch presented a revised table of the department's division of education spending, he was able to better pitch the point that State education funding was largely defense-oriented, and that, importantly, it was not being controlled by the State as much as it could be. The use of PPBS, in other words, allowed to show how State steering of education was being obstacled through conventional budgeting methods.

For this reason, the education department needed to re-organize the way that budgetary expenditures were divided, a step which allowed to have a clearer picture of the situation and tailor expenditure. Firstly, this could be done by *dividing expenditure into programs* (primary education, secondary education, etc). Secondly, by specifying *which specific mission each program was funded to pursue* (such as vocational vs preparatory secondary education, for example). This, importantly, meant being able to exclude and eventually remove unnecessary expenditures from the department's total budget, or, in other words, rule out those educational missions which the department either did not want to pursue or believed should be financed by other departments (the financial burden of non-strictly educational goals could thus be shifted to other departments' budgets). As Hirsch estimated, a more precise account of the costs incurred *in* education and *for*

education emerging from this framework led to '[…] a substantially *smaller* education budget figure than the $3.5 billion […] possibly somewhere between $2.6 and $3.0 billion' (Hirsch, 1965, p.27, emphasis added). These considerations allowed the RAND corp. to insist on the importance of *setting clear objectives* (i.e., strategic decisions) for education through a different structuration of the voices of expenditure (i.e., the budget), bringing data on education expenditure to be more "accurate" and more suited to steering according to State objectives.

Finally, Hirsch's suggestions provided a framework useful to both the federal State and the education department. Through program budgeting, the latter would take more informed decisions on what it was willing to fund. For the former, instead, program budgeting allowed a direct quantification and visualization of all of its departments' claims for State resources. This would allow the State to take '[…] competitive decisions', namely, to decide between departments' resource requests.

> '[…] in light of a […] $93 billion Federal budget, and a $25 billion public education budget, is a $3.6 billion federal education budget of optimum size? Or would, for example, an additional billion dollars for education prove more beneficial to the nation than adding $1 billion to NASA's 1963 budget of $3.7 for space exploration, or to the DOD's 1963 budget of $57.8 billion? […] are we putting too much into our defense program and not enough into education? Or vice versa? […] partial answers […] require our estimates of […] their multiplier effects' (Hirsch, 1965, p. 30).

With respect to this last aspect, the report made an explicit policy proposal: that the federal budget for education be increased, relative to the proportion dedicated to it when it was compared to the sums being allocated to defense and military projects.

> '[…] *the additional federal educational dollar will carry extra duty* and this increment needs to be estimated and considered in a discussion of whether we would not be better off if the federal government were to increase its education budget.' (Hirsch, 1965, p. 30, emphasis added).

In these ways, Hirsch introduced marginalist thinking to the State's departmental resource allocation decisions, presenting in this way a trade-off view of how the State could decide to prioritise its many objectives which depended importantly on departments' competition for limited State resources. RAND's suggestions moved beyond mere technical guidance on budgeting methods, assimilating proper policy advice. Acting as a 'technical' consultant to the State's decisions on education thus simultaneously meant asking the State new types of questions which RAND intellectuals were keen to articulate for it. Crucially, this set off a new normative

understanding of social policy decision-making, where RAND exploited Cold War context to concretely enter the policy preoccupations of the time. By proposing its approach as viable for thinking about policy trade-offs, RAND intellectuals consolidated the demand for methods which effectively appeared to give both the State and the federal agencies more leverage in decision-making. Importantly, an innovative element of RAND's consultancy during this period lies in the fact that whilst it created new questions, it did not seek to outright answer them and in this way take a political stance on issues of expenditure allocation. In this way, RAND consultancy could claim its objectivity mantra whilst in practice acting towards a specific tightening of the defense-education nexus. Hirsch's decision to set education and defense in a trade-off relationship was a political move, which left RAND analysts with sufficient space for conceptualising educational and defense objectives as being either trade-off or complementary depending on the State's priorities. In this way, Hirsch did not try to overstep political authority as much as respond to its concerns, successfully making space for his consultancy.  Whilst Hirsch suggested to increase the education budget, the strength of his proposal lied in the ability to relate the two matters within a technical framework of decision-making in the hope that the corporation would be further consulted for help in addressing the very questions that it had raised to the education department and to the U.S. State.


## 2.2. The CIPP model

The expansion of educational evaluation in the US into an industry and a '[…] profession dependant on taxpayers money for support' (Stufflebeam & Guba, 1970, p.11) was largely the result of two key novelties and demands: firstly, those introduced as a result of the political influence of military-intellectual cadres; which led to the elaboration of both the NDEA and the new curriculum movement in the late 1950s. Secondly, those of the Elementary and Secondary Education Act of 1965[16] (Madaus et al., 1983), passed within Johnson's broader "war on poverty" reforms. As Rudolph has importantly highlighted (2002), Eisenhower's trust in the scope for scientific-military methods in the broader expansion and improvement of scientific manpower through educational reform explains the NDEA's and new curriculum movement's attention to educational innovation in science and mathematics. Seen as part of a broader strategy for strengthening national security vis-à-vis the Soviet Union's advanced scientific research, the reforms of the late-1950s compelled educational institutions to confront the changes being proposed and try to enact them. These changes included, amongst others, carrying out laboratory experiments in the classroom, and training teachers to work with new maths and science textbooks and films being proposed by the NSF (Rudolph, 2002). Further, defense intellectuals had also successfully

---

[16] See Public Law 89—10

managed to use a systems analysis approach for designing the objectives of the reforms also beyond scientific education, as was the case for "project English" and "project social studies" (Rudolph, 2002)[17]. The necessity that schools '[…] shift their concern for educational evaluation from the realm of theory […] to the realm of practice and implementation'(Madaus et al., 1983, p.13) was made more compelling by the new demands placed on schools with the passing of the Elementary and Secondary Education act in 1964. This Act not only confirmed the political intent to intensify educational innovation, asking local education agencies to ensure that new educational needs were being addressed in schools (Stufflebeam, 1968a,1968b); it also asked schools to make us of standardised test-data and assess the extent to which they were reaching the novel objectives set by political programs (Stufflebeam, 1968a, 1968b; Madaus et al., 1983). Furthermore, schools and evaluators were aware that coming to grips with these changes could secure important funding, pushing evaluation researchers to debate new practices and theories most apt to addressing the context of the 1960s, marked by significant increases in student enrolments and a less significant increase in teachers. Thus the 'age of expansion' in evaluation set-off between 1958 and 1972 (Madaus et al, 1983), where it was felt that education's illness – its weakness at devising its own evaluation methods rather than merely importing them from other contexts (Stufflebeam, 1983) – needed to be remedied through more and better educational evaluators.

The CIPP model ideated by Daniel Stufflebeam emerged within this expansion period and was marked by an attempt to provide schools with a concrete way for carrying out "program evaluation", which had two specific meanings. Firstly, it meant being able to assess the developments being requested by political programs such as the 1965 Act. Secondly, it meant also understanding how schools' new objectives could be articulated into programs and assessed through an evaluation framework, which, ideally, would elaborate information on schools' developments and inform institutional, federal, and State actors on relative successes and failures (Stufflebeam, 1968a, 1968b). In discussing educational-specific program evaluation, Stufflebeam and others insisted that education could not merely adopt existing models such as PPBS, for these had been designed to respond to specific – and in the case of PPBS, military – purposes. Aware of the functioning of PPBS and of the broader systems analysis movement, Stufflebeam attempted to draw out a model which would identify and respond to educational needs in their specificity, borrowing importantly from his research at the test development center of Ohio State university (begun in 1962), as well as from his experience as evaluator of educational projects being funded by the 1965 Act (Stufflebeam, 1968a). Whilst trying to distance itself from the PPBS model, however, I argue

---

[17]The new curriculum movement in the US has been defined as '[…]an experiment in applying innovative research and development techniques perfected by scientists during World War II' (Rudolph, 2002, p. 1), and Eisenhower's reforms have been critically conceptualised as the result of the rising influence of new scientific networks mobilized by Cold War concerns on national security rather than as a sudden, resurging political interest in public education (Rudolph, 2002).

that the Stufflebeam's CIPP model still importantly borrowed the theoretical framing of the systems analysis approach. Stufflebeam conceptualised CIPP as that model that kept federal and State decision-making in line with school-level decision making, such that CIPP eventually provided *both a broader guide on how information feedback-could function between schools and political funders & policymakers* and *specific operational guidelines for educators*. I discuss each in turn. The two key similarities between the CIPP model and the systems analysis framework developed at RAND are its insistence on evaluation as a tool for decision-makers through the constant feedback of information on activities, and the attention to objective-setting as being that key activity which defines and enhances evaluation. Thus, evaluation was not so much about reaching a judgment about the worth of a given activity (which in evaluation jargon is defined as summative evaluation (see Scriven, 1967)), as much as it was a process for helping decision-makers to collect data in order to formulate the activities to be carried out in the first place (Stufflebeam, 1968a; Guba & Stufflebeam, 1970). In his guidelines of an information-feedback loop between schools, State agencies and federal agencies, Stufflebeam argued that information gathering, and elaboration was crucial to decision-making. In line with a systems analysis approach, and quite similarly to PPBS, the CIPP model did not conceptualise objectives as fixed goals. Instead, Stufflebeam conceptualised the entire education system of the US as composed by the many schools in different States, such that system-wide (i.e., federal) educational decision-making was that process which could only occur through a systematic gathering of context-specific information provided by sub-systems (i.e., the schools). Thus, the CIPP model effectively conceptualises the education sector as a system whose objectives are interdependent on the sub-systems and, for this reason, objectives are subject to changes depending on how information is gathered and elaborated by both the single schools and the State and national decision-makers; just like in PPBS goals can be repeatedly re-drafted as new information on organizational activities is provided or as existing data is re-elaborated. Thinking from the perspective of State and federal agencies specifically, Stufflebeam thus encouraged schools to work on information gathering, organization and analysis, for this would have concretely allowed decision-makers outside schools to draw out new needs of the system (the education sector) by analysing the information received from schools (and, in the case of the federal agencies, the aggregate information provided by State agencies). As Stufflebeam argued, the strength of this highly interdependent process of objective-setting was that it allowed decision-makers to modify their political programs whilst these were being drafter, rather than surrender to modifying programs only following their implementation. Indeed, Stufflebeam saw this as the key strength of "continuous" evaluation (Stufflebeam, 1968a) because it surpassed the traditional exercise of evaluation as an activity that could be done only once the results of an activity were available. Indeed, if schools decided to implement the CIPP model they would be able to provide constant up-

to-date information on every step of the educational activities that they were carrying out, with the result that decision-makers would also be kept up-to-date and not have to wait for final results to modify political programs or begin judging the partial successes and failures of existing ones. The information feedback loop was therefore a key way to depict how decision-makers themselves could, through the participation of schools, carry out "program evaluation" of their political programs.

Additionally, the CIPP model allowed schools to carry out "program evaluation". For their lower position, schools needed to be able to translate political requests into educational activities (see figure 3 below). Whilst providing information on their educational activities was a way to inform policymakers and thus assist *them* in evaluating their programs, it did not really enhance schools' ability at unpacking how they had to change their practices to fit changing political demands. For this reason, Stufflebeam elaborated on the school-level dimensions of the CIPP evaluation model, drawing out its second key feature, namely the evaluation of *every step* of an educational process within schools – although he later moved away from schools specifically, suggesting that CIPP could be applicable to any type of institution (Stufflebeam, 2000). This conceptualisation informs the model's name: context, input, process, and product (CIPP) evaluation represent in fact the four key stages of the dynamic evaluation model. To explain how schools could carry out evaluation, Stufflebeam thus recast CIPP's principles in institutional terms. Information gathering and analysis was still central, and the CIPP categories do not reflect different moments of information gathering as much as different types of assessments being made on a school's data. Interestingly, Stufflebeam first used categories similar to those of PPBS for explaining how CIPP categories gave guidelines to the different types of decisions that schools needed to make.
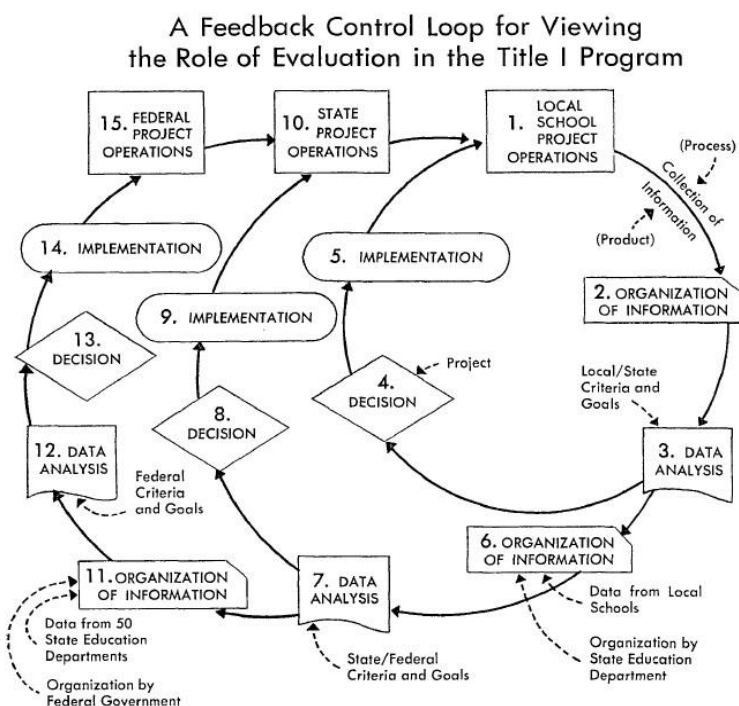


*Figure 3: Stufflebeam's Feedback Control Loop (Stufflebeam, 1968)*

As such, planning, programming, implementation, and recycling decisions can be taken through Context, Input, Process, and Product evaluations. Context evaluation consists in an assessment of the environment within which a school operates, and, from that environment, understanding the needs that educators need to integrate in education. In the context of the 1965 Act, such needs were

those of improving the attainment of disadvantaged children, for example. As Stufflebeam later added, context evaluation was not isolated to the beginning of the project, because new information on the context could be added to a project's planned development (Stufflebeam, 2000). Input evaluation, secondly, consists in analysing how endowments can become part of a program, and thus involves the setting out of criteria which can be used for assessing the school's broader achievement of goals. In this way, resource-use, schedules, and staff requirements can be analysed in order to program how activities can be carried out. Interestingly, Stufflebeam suggested how, similarly to the systems analysis approach used in PPBS, '[…] the end product of input evaluation is an analysis of alternative procedural designs in terms of potential costs and benefits.' (Stufflebeam, 1968a, p.36). This marked ideal outcome of input evaluation recalls the importance of alternatives modelling or hypothesizing also dear to PPBS developers, suggesting that also in CIPP there is the underlying conviction that social complexity can be reduced by quantifying it into costs and benefits. This evaluation stage concretely feeds into process evaluation, which monitors how inputs are being used and if there are issues in the way that their implementation was initially intended. In this way, process evaluation analyses activities as they are being carried out, allowing to take 'implementation' decisions which evaluators can modify if processes are considered incoherent with given goals or simply hardly implementable. As Stufflebeam put it, this stage was about identifying identify '[…]procedural barriers and remain[ing] alert to unanticipated ones' (Stufflebeam, 1968a, p.34), and could involve interviews with teachers on the running of a project, for example (Stufflebeam Guba & Stufflebeam, 1970; Stufflebeam, 1968a). Finally, product evaluation is that final stage which allows to decide whether a given process should continue or be discarded. This stage of evaluation is ideally a final one, for it compares educational process outcomes with their initial goals. Differently to PPBS, product evaluation was not intended to focus on cost-efficiency, nor did it implicitly refer to cost-assessment of schools' outcomes. Rather, Product evaluation – albeit the reference to "products"– intends to consider the multiple types of outcomes of a project, which in school programs can importantly allow to capture concrete changes in attainments (i.e., did pupils improve in maths/reading skills?), for example. Indeed, though of course accountability is a key dimension of the CIPP model, especially when seen from the perspective of middle and top decision-makers, product evaluation seeks to explore more than cost-accountability compared to how the PPBS model turns it into a concrete steering instrument.

## 2.3. Conclusions

The chapter has discussed the key aspects of systems analysis.  First, its post-war military roots, which suggest that NPM notions such as objective-setting and cost-analysis often associated to the

1980s actually emerged during the Cold War context as an attempt to strengthen the position of US defense intellectuals and the defense department. Second, the way that this military methodological innovation was a broader framework for using detailed information on activities to re-frame social policy decision-making, including educational policy. When studied in relation to the systems analysis approach, PPBS and CIPP portray, in different ways, the seeds of a managerial framing of educational governance. They were inherently geared at assisting the steering powers of decision-makers, either those within an organization (such as a school or an enterprise) or those within public bodies providing the funding for public institutions (such as federal or State agencies), or both. These models also shared features of a newly emerging managerial approach geared at assisting governance rather than at studying productivity. Both emphasize how the quantification of educational inputs and outputs is key for assessing the appropriate functioning of educational institutions. Secondly, they attach great importance to the process of goal-definition and goal evaluation. Thirdly, they conceptualise evaluation as that exercise which supports social policymaking, making it a governance device for its ability to influence decisions on educational expenditure. These features comprise key elements of a technical approach to public social processes such as education and formed the grounds of the development of "output-based" governance in education (Bürgi, 2016a, 2016b). Having discussed how the features of systems analysis relate to educational governance, the next chapter opens the historical inquiry of managerial educational evaluation indicators, beginning with an analysis of the diffusion of systems analysis at the international level.

# Chapter III

# The Cold War and the Managerial Revolution in Educational research

The rise of managerial experts in education tends to be associated to the rise of benchmarking in international politics. For western European countries, this was in the early the mid-1990s, when both the EU and the OECD advocated the use of performance benchmarks as key to the improvement of countries' competitiveness and growth (Giannone, 2019). This broad political turn to the use of quantitative indicators surely had a significant impact on the development of educational governance. The EU Delors commission, for example, insisted on educational improvements as key steps for employment-intensive growth. Similarly, the Bologna Process and Lisbon treaty encouraged the use of indicators as part of the EU's integration process. Taking these moments as a starting point, scholars often associate the rise of a managerial expertise in international education policy to the influence of technocratic networks of the 1980s and 1990s . The roundtable of European Industrialists (Pinto, 2012) and the OECD policy élite (OECD, 1997) are examples of those transnational élites who acquired legitimacy in evaluation matters thanks to their proximity to key international institutions (Apeldoorn, 2001), and, specifically, for their advocacy of managerial instruments as key pillars of public sector reform (Giannone, 2019). As experts such as business schools, think tanks, policy commissions, offered their expertise to governments, so did a firm theory of the state develop, with states agreeing to reconfigure their institutions towards efficiency with the ultimate goal of having them resemble markets. The managerial consultancy of such groups was functional to the construction of neoliberal governance itself (Davies, 2016). In an almost performative vein, scholars suggest that the international proliferation of educational indicators was a result of States' increasing adherence to the instruments proposed by these transnational technocrats: each time that benchmarks were implemented, managerial expertise found new spaces of legitimacy in educational governance, and gradually diffused to extremely diverse education sectors. However, when taking a step back from the 1990s, the international policymaking élites arrived after another group of influential educational experts. Whilst abovementioned accounts allow to understand the emergence of political interests in benchmarking, they fail to explain why some Western European universities were piloting the use of performance indicators as early as the mid-1970s (see OECD, 1968; Khan, 1971; Fielden & Lockwood, 1973). Indeed, the view that educational indicators be used for governing universities was already well developed and widespread in OECD by the 1990s. Drawing from critical insights, (Bürgi, 2016b), this chapter argues that recent developments in educational

benchmarking have a longer lineage rooted in a new managerial approach which developed at the OECD during the Cold War period (Bürgi, 2016b). In fact, the OECD has a long-dating history in the production of global socio-economic indicators[18]. To an extent, the OECD was a key constructor of the relationship between economic and social development (Bürgi, 2019).

In what follows, I explore how the epistemological grounds of 'output-governance' (Bürgi, 2016) in education developed through a complex intersection of multiple developments which occurred within and through the OECD. In fact, new concepts and their respective technologies *developed in* and *diffused through* the international organization. As the chapter explores, in fact, if on one hand OECD projects of the 1960s largely mimicked the US path discussed in chapter two, (Bürgi, 2019; Elfert, 2019; Elfert & Ydesen, 2020) on the other they also radically innovated on existing notions of university governance specifically. The diffusion of new ideas and practices from the US to the OECD, first, and from the OECD to specific national contexts, second, is thus here analysed as a process which underwent continuous re-articulation depending on the ways that groups of agents decided to use them. Concerned with capturing the possible reach of a new managerial expertise to Italian and English universities, this chapter discusses these developments and their intersection, in the process also tracing the international lineage of educational evaluation methods. Two key aspects will be considered.

Firstly, it is key to point out how the OEEC's very first activities inherited the US's agenda of building "brainpower for the Cold War" (Clowse, 1981), which spread in western European member States. By the mid-1950s, for example, the European productivity Agency (EPA) worked to steer countries into higher levels of scientifically skilled manpower, setting off the trend of *educationalization* (Depaepe & Smeyers, 2008; Bürgi, 2016b) of these new problems of modern society. This process fuelled the view that education itself was that engine of wealth gains and growth. European countries' adherence to these projects depended importantly on the US's strong global position in the post-war international order, with Marshall funds and other reconstruction funds coming with strings attached. The OEEC's role in the distribution and monitoring of reconstruction funds, further, gave the US an '[…] opportunity to reconstruct Europe in the American image' (Judt, 2005, p.93). An example of this is the establishment – set off by the EPA – of "national productivity centres" in member countries, a project which pushed countries to adopt liberal policies through macro-level industrial and economic planning (Elfert, 2019; Elfert &

---

[18] The OECD published one of its first reports on economic growth and education, "Economic Growth and Investment in Education", in 1959 and set educational targets for member countries in 1961(See Bürgi, 2017b). Between the mid-1970s, the organization worked on educational planning methods and indicators (see OECD, 1967,1974,1982). In a re-appraisal of educational planning, the organization defined it as an exercise of '[…] administrative sub-system rather than with the executive policy-making one'. The OECD defined the 'participatory educational planning movement' of the 1970s and 1980s as a way to mediate views on educational policy of agents at different levels, such as governments, on one side, and the increasing number of agents participating in local planning (schools, local authorities, enterprise).

Ydesen, 2020). The OEEC's closeness with the US also explains the significant funding provided to educational research by US bodies such as the department of defense, the RAND corporation, and the Ford Foundation amongst others. As Chapter 2 has discussed, these groups played a central role in the construction of PPBS's efficiency framework and methods. As this chapter further explores, their active support in OECD projects continued this trend, fortifying the experimental grounds of educational projects launched in the US by creating spaces within the organization for further theoretical and empirical developments in systems analysis and PPBS specifically. The US legacy in the OECD's educational research reflects more than just the reach of US power. As knowledge coming from the US context was further developed, the projects diffused a specific articulation of education as a strategic sector for the achievement of specific economic-political goals. Though this was conveniently framed as a technical-objective exercise, it represented a deep political re-definition of the very scope of education increasingly distant from a strictly social-pedagogical conception.

Secondly, the OECD was a platform (Centeno, 2018, Elfert, 2019; Bürgi, 2016b) where new networks of 'educational planners' formed and interacted. Defined also as 'technocratic-educationalizing networks', these were researchers and national delegates from diverse backgrounds (including sociologists, mathematicians, pedagogists, economists, administrators) who exchanged experiences, views, and projects; in this way creating new knowledge around how educational sectors of the "modern society" of the 1960s (Schmelzer, 2012) should develop. The key activity through which new members met was research, which developed through a multiplicity of projects and research centres of the OEEC and OECD, such as the Committee for Scientific and Technical Personnel (CSTP), the European Productivity Agency (EPA), the Mediterranean Regional Project MRP and the CERI. Most of this work was framed in positivist and liberal terms, and encouraged the production, gathering and elaboration of socio-economic data. A key aspect of the OECD's projects was their comparative-analytical dimension, often carried out through recourse to quantitative methodology and the production of hard empirical evidence. Such information crucially developed through committees' decisions to experiment new theories and practices – such as human capital theory– in specific contexts, treating the world as a laboratory. The specific kind of knowledge that these new networks helped to construct is hard to generalise because of the extreme array of projects and personalities cohabiting the OECD. Surely, the emerging network of university managers studied here eventually succeeded in eroding the legitimacy of social-pedagogical approaches in favour of quantitative, mechanistic, and managerial ones. As I argue here, the managerial revolution in educational research owed more to the actions of specific agents within these international networks than to the universalist or abstracting language of management itself.

Thirdly, the OECD was a locus where radically new research on university governance emerged between the mid-1960s and the mid-1970s. Interestingly, projects conceived the university as an organization requiring improved institutional management, particularly following the protests of 1968. This marked the beginning of *de-facto* assessments of universities' governance models, presented at the OECD by university agents seeking to ameliorate university management. As university governance was discussed in relation to the systems analysis framework, projects such as CERI's emphasized the usefulness of cost-planning procedures and of the PPBS framework more broadly. Research on – and in – universities took this turn due to the active advocacy of agents which were part of an Anglo-Saxon military research network, Alex King and Philip Coombs. As existing studies show (Bürgi, 2016a, 2016b) both were keen to continue developing systems analysis research, begun during or immediately following the second world-war. After having worked in British operations research during WW2, for example, it was King who brought systems analysis to the CSTP in the mid-1960s and, briefly later, at CERI. Similarly, whilst head of the Ford fund for the Advancement of Education, Coombs mobilized resources for RAND research on public education, and later founded the IIEP at UNESCO. As I argue, these intersecting developments fuelled a managerial revolution in educational governance which points to a much longer historical lineage of managerial practices in education which scholarly literature fails to adequately account for. This chapter it in detail, paying particular attention to how the managerial revolution occurring at the OECD walls played out among Italian and English groups of researchers. To study the transmission of output governance in education from the international to the two national contexts, I focus on how agents participating in OECD projects articulated the novel concepts and practices being developed, exploring whether and how they advocated them as viable innovations.

This chapter is divided into three sections. Firstly, it discusses how the rise of the "economics of education" at the OECD was closely intertwined with US educational policy goals such as the NDEA discussed in chapter two. As projects such as the MRP and the EPA encouraged countries to enhance their productivity, education was increasingly conceptualised as an instrumental sector. This emerges clearly in how a group of Italian researchers at SVIMEZ and CENSIS, who used the approach to advocate for major educational productivity in the early 1960s, discussed in sub-section 3.1.2. After having traced their research, the section explores their turn to the CIPP evaluation framework. CENSIS used the model as a key reference when it proposed a national school evaluation framework for Italy in the early 1990s. Section 3.2. returns to developments at the OECD, exploring how the rise of educational planning projects in universities was crucially shaped by an alliance between Kings and Coombs, who mobilized to introduce the PPBS framework in university-specific projects. As is explained, the methods devised at the institutions which emerged from this alliance, CERI and the IIEP, radically transformed previous

educational productivity methods into planning practices geared at cost-maximisation and performance assessment. After discussing this radical innovation for each research centre, the sub-section 3.2.2.1. explores how Sussex University's participation in OECD projects influenced its turn to managerial governance principles. As I show, a group of Sussex members significantly appraised the PPBS framework because it recognized how its use could empower administrative and executive roles in universities. Having seen how universities gradually became involved in projects on institutional management, section 3.2 explores how university VCs articulated their role in the context of this broader international research on university governance. I first analyse how management seminars at the Conference of European Rectors framed VCs as managerial actors, relating this to the CRE's longer-dating history in university policymaking. Secondly, I introduce the two-vice chancellor committees analysed in this thesis, the British one (CVCP) and the Italian one (CRUI), and briefly discuss their role in their university contexts. This is key for contextualising their role in developing a nexus between evaluation and university governance, discussed in chapters four and five.

## 3.1. The Sputnik launch and educational planning at the OEEC

Following the launch of the Sputnik in 1957, the OEEC's existing work on countries' economic productivity was further strengthened by an influx of human and economic resources from the US, which were geared at 'economizing education' and increasing European countries' levels of scientifically and technically trained labour (Elfert, 2019). With the support of funds from the US department of defense and from the Ford Foundation (Elfert, 2019), the organization mobilized to diffuse the key principle of the NDEA, namely that educational steering was – specifically in the hard sciences – crucial to the growth of manpower for the Cold War effort. The achievement of this goal not only necessitated key educational readjustments in favor of the hard sciences, but also the advice of planning and economic experts that could help countries to strategically steer their educational sectors. Against this backdrop, the OEEC founded a new body, the Committee for Scientific and Technical personnel (CSTP, est. 1958). The CSTP – which later became the OECD's "Directorate of Scientific Affairs" – was the key locus where the OEEC's educational research agenda first played out, aiming to achieve international comparability of countries' human (i.e., immaterial, educational) capital (Bürgi, 2016; Tröhler, 2013). Through this comparative outlook, the committee encouraged member countries to plan their economic investments in accordance with optimal manpower forecasts (Ydesen, 2019). The CSTP's concern was, just like in the US, that of developing a '[…] rational, technical-scientific preparation necessary to confront the new needs brought about by economic growth' (OECD, 1962, as cited in Alachevic & Granata, 2019, p.236).

This impulse set the grounds for the OECD's managerial revolution in educational research. As OEEC countries adhered to the U.S's Cold War educational agenda, their research developed through the impulse and methods being pioneered in the US through the defense intellectuals.

The committee was directed by Alexander King, as was the body from which it inherited its functions and outlook, the EPA. The latter has been defined as '[…]an octopus weaving a network of agents' (Bürgi, 2019, p.30) because it of its determination to form new 'change agents' that through activities and networks which, between 1953 and 1958, could convince member States that efficient economic production was the core driver of social progress. The EPA conceptualised productivity as a cultural phenomenon (Bürgi, 2019), and mobilized to diffuse this value through both practical and theoretical projects, such as data collection, the monitoring and reporting of countries' productivity, conferences, and management trainings (Bürgi, 2019). When King moved to the CSTP, he continued to encourage a similar *modus operandi* to the issue of European manpower levels. As former scientific advisor of the British central scientific office in Washington DC during WWII, King had strong faith in the usefulness of a scientific approach to social problems. Building from his work in operations research during the war, King had attempted, albeit unsuccessfully, to introduce methods of war planning into British civil service(Bürgi, 2017). His '[…] normative, technocratic presuppositions […]  in favour of a "monarchy of evidence"' (Bürgi 2016, p.170), however, found support among military and philanthropic networks which were simultaneously financing US educational developments. King's military networks and post-war aspirations allowed him to secure 40% of the OEEC's budget for the EPA, with another \$100 million funded by the US and the Ford Foundation (Bürgi, 2019). The CSTP, similarly, received half of its resources from the US department of defense, and the OSTP, the committee's "office for scientific and technical personnel" received \$500,000 from the US and further support from the Ford Foundation (Elfert, 2019). King's active engagement in the founding of the Club of Rome (Schmelzer, 2012; King, 2007) – initiated in the 1950s– reinforced his prestige among European members at the OEEC, an aspect which facilitated his later work in the OECD.

Besides being a key conduit in the transmission of  the productivity paradigm (William, 1982) to European countries whilst at the EPA, King sparked the circulation of human capital theory whilst at the CSTP. By the 1960s, as the OEEC was faced with urgent Cold War preoccupations, King took on US proposals to increment the organization's work on countries' manpower and productivity levels (Elfert, 2019) and mobilized his ties at the Ford foundation to set up a research initiative called the "study group on the economics of education" (Bürgi, 2016a),

located at the OSTP. Members[19] of the group came exclusively from the OEEC's "more advanced" northern countries (US, France, Sweden, Germany, Denmark, and the UK), who had since 1959 been working on a program which later merged with the study group's research themes, "economic growth and investment in education". At the Study group, it was John Vaizey's work (1961) on human capital theory, amongst others (see Schultz, 1963; Becher, 1964), which dominated (Murray, 2012). Vaizey was Professor of Economics at Brunel University in London and later elected member of UNESCO's IIEP (Murray, 2012). He was a key theorist of the first developments of human capital theory, and specifically he explored the correlation between economic investments in education and countries' actual vs forecasted scientific manpower requirements. Vaizey defined his work as "the economics of education" picking up the same concept introduced by Kershaw & McKean back in 1959 (Kershaw & McKean, 1959). Given the continuity of the Ford Foundation's financial support to these RAND intellectuals' work and, later, to the study group, the development of Vaizey's work at the OEEC can be seen as a middle ground between the systems analysis framework and the human capital approach. Similarly to the RAND intellectuals, Vaizey conceptualised education as an instrumental input. He too was concerned with the relative educational-technological advantage of the USSR, which '[…] has now for a period of over thirty years invested between twice and three times the average western proportion of its national income in developing education.' (1961, p.7). Further, Vaizey emphasized the centrality of (investments in) education in the increase of manpower. Still, there were also important differences between the two applications of economic calculus to the study of education sectors. Vaizey and the RAND intellectuals in fact studied two different levels of the production of human capital: the latter were concerned with its dependence on the local organization of educational factors, whilst the former studied the trade-off relationship between national investments in education and economic growth (Vaizey, 1961). Kershaw & McKean's work on the micro-articulation and organization of the inputs and outputs in education addressed educational institutions, whilst Vaizey conceived *the economics of education* as an approach useful to policymakers. Influencing national policy thus implied the quantification of education's productive capacity at an aggregate, macro-level, rather than the *optimisation* of given levels of national resources within specific educational fractions[20]. Differently to system analysts, therefore, Vaizey's approach solicited OECD member countries to importantly rethink their educational investments (1961). As Bürgi notes,

---

[19] The group was very small, composed by the Chairman, Hennis Friis, and Freidrich Edding, Seymour Harris, Raymond Poignant, Ingvar S. Vennilson, John Vaizey, Michel Debeauvais, Selma Mushkin and Jan Tinbergen (Papadopoulos, 1994).

[20] As discussed in chapter two, Hirsch also suggested that the US State rethink its education investments. However, and importantly for understanding the simultaneity of research developments occurring at the OEEC and at RAND, Hirsch's proposals only came about in 1965, thus 4 years later to Vaizey's and 5 years after Kershaw & McKean's pilot of systems analysis in education.

> […] the CSTP's establishment '[…] not only signalled the start of the OEEC's efforts to integrate science but also its engagement for natural scientific and mathematical subjects and *the first forms of educational governance (particularly for tertiary education) on the basis of extrapolating statistical forecasts and the country reports built upon them*' (Bürgi, 2016a, pp. 413-414, emphasis added).

The committee's work increasingly resembled an exercise of systematic –and to an extent, external– evaluation of countries' educational development against the benchmark of their ability to plan for economic growth and increased levels of scientists and engineers. From the very start, Marshall-fund beneficiaries were drafting annual reviews for the CSTP which reported levels of technical and scientific manpower and education. These were often formulated in collaboration with the CSTP's 'independent experts', who monitored member States' developments and discussed them with national officials. The Study group's approach thus intertwined with the EPA's advocacy of economic planning, and its increasingly policy-driven proposals encouraged countries to acknowledge the necessity of planning educational sectors' economic capacity to maximise wealth gains deriving from it, until eventually in 1962 European Education Ministers fully adhered to the approach (Papadopoulos, 1994). If planning was to work, however, countries needed to know exactly how much growth to plan for. This turned manpower planning and forecasting exercises into drivers of new developments in human capital theory. By the mid-1960s, the study group elaborated the "residual factor", a quantification of the wealth gains which given educational attainment levels brought to the economy, expressed as a percentage. Differently to the micro-analysis of the "rate of return" to education – which quantified *individual* gains from given levels of educational attainment (see Vaizey, 1961) –, the residual factor was an attempt to calculate and compare countries' correlations between GDP percentages of educational investment and levels of economic growth, a development which could have importantly legitimated the economization of education (Papadopoulos, 1994). The residual factor eventually proved too varied between countries to convey a general trend (Papadopoulos, 1994). However, it deepened the existing division between the OEEC's more advance members and its "underdeveloped ones". By the mid-1960s it was common view that advanced countries necessitated a social demand type of growth because these countries had already achieved high levels of skilled manpower and needed to focus on responding to the increases in demand for education sparked by the baby boom. The underdeveloped countries of southern Europe, on the other hand, required a "manpower demand" growth (Papadopoulos, 1994), where along with social demand, they needed also major educational planning to accelerate their production of skilled workers. Southern Europe's common disadvantage in skilled manpower owed much to their unemployment levels, low educational enrolment ratios

and still high percentages of agricultural labour force (Papadopoulos, 1994). As the next section discusses, Italy was part of those countries whose backwardness was directly addressed through yet another project which sought to steer the economy towards growth, the Mediterranean Regional Project (MRP from now on).

### 3.1.2. Italy and the Mediterranean Regional Project a driver for educational planning?

In Italy's case, the rise of educational planning owes much to its "underdeveloped" position in the post-War international context. Defined as a '[…] straightforward manpower planning model based upon an input-output view of the economy' (Williams, 1987, p. 335), the MRP was carried out between 1962 and 1965. Following the OECD's previous research on manpower levels in member countries, the MRP's key approach was to use manpower needs to construct educational targets, strengthening the idea that education was an instrumental input to economic development and encouraging countries to speed up their production of skilled labour. The central aim was to align the underdeveloped Southern European countries' (Italy, Greece, Spain, Portugal, Turkey, and Yugoslavia) national quantities of technologically skilled labour force to those of the advanced European economies. Through funds from participating countries and the OECD, the MRP asked countries to follow what Villari has defined as 'the US way of civilization' (1972). Indeed, the thrust for Italy's economic progress intertwined with its developing planning expertise, as US economists visited the country as part of its intent to train future generations of liberal, Southern European scientists, economists, and planners (Elfert, 2019). As a matter of fact, the MRP intervened in those countries living dynamic anti-capitalist social upheavals, and the program sought in no marginal way to make use of the Mediterranean's strategic socio-political significance for the US's Cold War strategy (Granata, 2022). Italy's own political climate in the 1960s was fervently left-wing, yet increasingly distant from the communist forces of the immediate post-war period and increasingly tending towards a moderate socialism. Indeed, the communist party had a short-lived leadership in parliament, such that following the De Gasperi III government it faced consistent challenge from the social democrats of the Democrazia Cristiana. Throughout the 1950s and the 1960s, the country was guided by the governmental majority formed by alliances of the socialist party and the Democrazia Cristiana, with the communist party in opposition. Since the establishment of the Republic in 1946, the two parties had taken different sides on the suitable economic planning strategy for Italy's reconstruction (Granata, 2022). The social democrats were willing to devise a liberal form of State planning which would openly collaborate with large portions of the industrial élite, whilst communist members feared the fascist legacy of the north's

key industrial actors, and were reluctant to negotiate with them (Granata, 2022)[21]. Eventually, liberal planning succeeded as ministers seconded the post-war economic-political designed by Pasquale Saraceno, a key social-democrat intellectual. More important still, his views on a moderate interaction with business for the planning of the country's economic development found fertile ground in the country's "planning barycentre", a social research association which Saraceno coordinated (Granata, 2022), namely SVIMEZ[22].

As an "association for industrial development in Southern Italy", SVIMEZ had been founded in 1946 and was foremost engaged in research on the country's inequalities, proposing economic and social policies for remedying them directly to government. SVIMEZ was all but new to international developments in economic planning. In 1957, for example, its researchers travelled to Greece through the EPA program to '[…] identify structures suitable for designing […] economic planning' (Granata, 2022, p.15). SVIMEZ was also recognized as a fervent centre developing well-rounded educational research by philanthropies such as the Ford Foundation and Rockefeller foundations, which in 1960 jointly organised the Bellagio meeting on "Economic aspects of educational development" (Granata, 2022) and invited SVIMEZ to take part. By the time the MRP was launched in Italy in 1962, SVIMEZ readily stepped in well-equipped to initiate research on Italy's future manpower building path, using the OECD's economics of education approach as its key example (Spalletti, 2008). SVIMEZ's research eventually produced "snapshots" of national education, using these as variables for forecasting projections on the required macro-economic investments. The education sector – both secondary (SVIMEZ,1961) and tertiary (Martinoli, 1962) – was analysed in detail in the attempt to calculate its long-term occupational potential in the necessary fields. SVIMEZ's suggestions were "modern" and encouraged those "human attitudes" (Granata, 2022) to educational planning which in its view had made Northern European countries' achievements possible. In observing the cases of the USA, Norway and Finland, SVIMEZ argued that their high GDP rates were the result of a well-planned "coordination of human activities*"* (SVIMEZ, 1961). This organizing and organized, rational and rationalising, managerial capacity was seen as being the core of their high and growing *modern* manpower levels, and, hence, of their economic growth. Along these lines, SVIMEZ later turned to forecasts, following the CSTP's planning legacy, encouraging policymakers to reform education in order to achieve major increases in scientific and technical personnel.

In a study on tertiary education (Martinoli, 1962), SVIMEZ insisted on universities' strategic role in shaping the country's ruling class ("la classe dirigente"). Building skills around

---

[21]Italy's industrial development was largely unequal between the North and the South, and the "cassa per il mezzogiorno" –established by law in 1950 – was precisely an attempt to create specific funds and industrial development policies for southern Italy.
[22] For a history of SVIMEZ see Vitale, 2000.

economic planning was key, for it was seen as the […] essential guide in the complex life of modern society' (Martinoli, 1962). In line with the MRP's ambitions, SVIMEZ's studies forecasted future manpower levels starting from its assessment of graduation growth rates in different university macro-disciplinary fields[23]. Similarly to its study on schools (SVIMEZ 1961), it forecasted growth rates over a 15-year period, up to 1975. For all the disciplinary fields analysed, SVIMEZ concluded that universities would rapidly become overcrowded due to the rapidly increasing enrolments (particularly between 1956 and 1961), up by 80%. (Martinoli, 1962). The predicted overcrowding was based on the analysis that enrolments would continue to grow, but, also crucially, that graduation rates would continue to remain, in proportion to the new enrolments, weakly constant. In the economic disciplines, for example, only 27.7% of the enrolled graduated (Martinoli, 1962). This 'bottleneck' was a central worry in the context of the study's introductory preoccupation: making universities the right places for the construction of a new, modern, ruling class – echoing how Coombs argued that "the right kind of people" needed to be educated in the post-industrial society (Bürgi, 2016a). Indeed, no ruling class could be formed without graduates. For the researchers, the universities were to blame for the bottleneck, teachers were inefficient at graduating students on time (Martinoli, 1962). Universities were also unable to provide society those graduates '[…] more equipped with given attitudes' who could […] exert rational and efficient entrepreneurial functions' (Martinoli, 1962, p.35).

Engineering faculties were less attacked, as this was that educational segment which CENSIS considered strategic in the enhancement of economic development, in line with the MRP's "manpower approach". Whilst the economic sciences were treated as 'explosive', with '[…]faculty look[ing] like an evening school, where only the most determined obtain their degree' (Martinoli, 1962, p. 20) the technical field was treated with less reproach. Having abundant engineering graduates was in fact a more pressing issue than sustain growth in the social sciences, for it was the former who would practically lead the country into rapid technological development. Indeed, the report set high targets for the hard sciences with the aim of '[…] obtaining an overall increase of engineering graduates equivalent to the growing income of the industrial sector, which is around 8% per annum, the new graduates in 1975 should be around 10.000 per year' (Martinoli, 1962, p. 16). CENSIS highlighted how the university sector's estimated ability to produce only a maximum of around 2.000 new graduates per year by the year 1966 was unsatisfactory. It proposed an 'urgent measure' for filling the gap of 8.000 graduates, namely 'fast track' for engineering students that could be attained by establishing *additional* engineering faculties teaching shorter and practical, excellence-oriented, courses. These would be aimed at the best students, those who '[…]

---

[23] It analysed faculties ("gruppi di facoltà"): the scientific macro-area (mathematical, physical, and natural sciences), the technical (engineering), the juridical, the literary and the economic (including political sciences).

demonstrate greater capacity and tenacity to access PhD positions and then work experiences' (Martinoli, 1962, p.17). As research turned into policy proposals, however, SVIMEZ was downsized because it was considered too autonomous by its founders (Spalletti, 2008), leaving this research in a halt for a few years.

In 1964, the Centre for Social Investment Studies (CENSIS from now on) was born from what had been the SVIMEZ's sociological research unit (Alacevich & Granata, 2019; Spalletti, 2008). Shortly after its establishment, CENSIS launched its journal, *Quindicinale di note e commenti* (see CENSIS, 1965), where educational issues from all over the world were discussed and presented along with the centres' own research findings and critical perspective. The centre largely continued SVIMEZ's educational research and was indeed well adjourned on educational planning specifically. Its founders, De Rita, Martinoli & Longo, continued to encourage and engage in studies on the relationship between education and research and economic development. Specifically, CENSIS presented a new methodological framework to SVIMEZ's, which paid more attention to the *sociological and transmissive dimensions of education* (Spalletti, 2008) when formulating proposed national investment plans. This allowed to relate economic and sociological insights from educational research rather than fuel educational growth merely through the axes of demography and social demand. In line with a human capital theory approach, CENSIS felt that whilst politicians acknowledged the value of research, they only recognized its productive and social capacity in abstract or generic terms. To remedy this political approach, researchers tried to support policymakers by empirically substantiating the view that education be considered similarly to an economic multiplier within society. They thus framed university research as '[…] a factor of technical and organizational transformation for many sectors' (CENSIS, 1964, pp. 29-30), hoping to initiate a debate on the necessity of more tailored national planning for research investments. It played its part through sociological analyses supported by econometric studies of educational productivity. This approach culminated in 1966 and 1968, when it presented Italy's "residual factor" (Spalletti, 2008). Researchers placed a great deal of significance to this instrument's explanatory power, suggesting that '[…] the residual factor was responsible for between 70% and 94% of the growth of Italian GDP between 1960 and 1961' (Spalletti, 2008, p.235).

**From educational planning to educational evaluation**

In the mid-1970s, CENSIS researchers turned to educational evaluation methods. Following a European resolution to improve countries' ability to integrate school graduates into the labour market[24] (European Education Ministers, 1976), the centre was appointed as the evaluator of Italian

---

[24] A key goal was also to ensure that education was not missed as a social opportunity, to '[…] give school and educational experience to youth who entered the world of work early in their lives' (Bucciarelli, 1982, p.44)

schools. The resolution had a strong social scope, seeking to identify the needs of school leavers, the reasons for lacking job opportunities, and strategies which could improve students' personal development. Member countries agreed to set-off pilot studies, such that Italy initiated 18 pilots in 11 regions – although no southern region was involved (Bucciarelli, 1982). The evaluative dimension of the project involved assessing how and why education leavers struggled to enter work and had a strongly sociological scope compared to the OECD's which aimed to construct indicators. The project sought to more closely comprehend the complexity of the issues being analysed in the first place (Bucciarelli, 1982). As a result, CENSIS researchers began thinking about how to understand and compare educational *contexts*, searching indicators that could transmit this information in the CIPP model. Contextual information was gathered in order to have a rounded understanding of those social, economic and territorial factors (socio-economic background, central/peripherical location of schools, schools' resource endowments, etc.) which distinguished schools, and which could provide background information for exploring the differences in educational attainment levels (Bucciarelli, 1982). The use of CIPP for thinking about school evaluation continued following 1988, when Italy participated in CERI's International Indicators of Education Systems (INES) project. The broader aim of INES was to comparatively analyse countries' correlations between educational attainment and the make-up of the labour market (National Research Council, 1995). Italy's participation, as the international coordination of the project, were guided by Norberto Bottani, who was a researcher at the OECD since 1976 (and until 1997). Italy's timid yet existing engagement in these projects allowed researchers to try and practically catch up on missed past opportunities that could have helped to provide policy and evaluation instruments for the school sector and improve its attainments (Bottani, 2009). For CENSIS, the project allowed to develop its existing research. It also showed the importance of establishing evaluation in Italy. Researchers saw education quite pragmatically, namely as '[…] an organization that converts its resources in services for its students, and its success lies in the results of this operation.' (CENSIS, 1991, p.110). Preparing future students to a modernizing society thus meant using evaluation as a means to equip 'organizations' with instruments for 'confront[ing] the difficulties of a given social order, […] transform [ing] these difficulties in processes of autonomy and strengthening.' (CENSIS, 1991, p.110). The centre was very clear on the point that modern society was advancing more quickly than educational sectors, and, similarly to how OECD projects framed the purpose of evaluation, 'the right kind of people'(Bürgi & Tröhler, 2018; Bürgi, 2016b) needed to be the 'results of this operation', of education. This was also very much in line with the feeling that a post-industrial society was behind the corner and that older generations needed to find the way to confront the surge of radically developing technologies and the implications that their diffusion would have on the composition of the labour market. Indeed, no evaluation at all further

slowed down this 'preparation' process and was seen as further aggravating Italy's already lagging position (CENSIS, 1991). For these reasons CENSIS was very outspoken in its intentions to provide practical evaluation guidelines.

By 1991, CENSIS proposed the founding of a national, independent evaluation agency for the school sector (CENSIS, 1991). This proposal was the result of previous projects on education and evaluation and, specifically, of Giorgio Allulli's own field studies in US and Australian national school sectors. At the time of the proposal, Allulli was a researcher at CENSIS. With other colleagues, he drafted a rounded and thoroughly informed proposal for the government which included critical considerations of other countries' experiences with evaluation. The centre used CIPP's principles when outlining the framework. The proposal was warmly welcomed by the then school minister, Sergio Mattarella,[25] who asked CENSIS to organize a crucial national conference on the school sector during which he personally insisted on the necessity of implementing a national evaluation agency. Regardless of such approvals, the court turned down the proposal on the basis that a national accountability exercise of such wide scopes could not – as CENSIS advocated – be assigned to an external, private, and autonomous agency, for fear that Ministerial financing could be abused and that political control over it could be lost[26]. To maintain State control over national school evaluation, the implementation of a future evaluation agency was delegated to the Education Ministry, yet the proposals were only partially implemented by the suggested body, namely the European Centre for Education (CEDE), then headed by the pedagogists Aldo Visalberghi and Giovanni Gozze. CEDE was later transformed into the National Institute for the Evaluation of the Education System (INVALSI[27]) in 1999. This distorted the initial proposal, for the way that evaluation was carried out was very detached from the contextual focus of its sociological imprint. CENSIS's more holistic model hoped to make extensive use of questionnaires to circulate to staff and pupils and parents alike in the scope of better grasping the (differences in) social composition of schools across the country[28]. The agency which CENSIS had proposed was indeed envisioned as a double exercise in which educational processes (and outcomes) were un-detachable from their context as Stufflebeam proposed (1966, 2000), using surveys as key samples and *in loco* visits by teams of evaluators that could directly witness the contexts being analysed. In CENSIS' model the role of the evaluator was not very distant from that of the researcher seeking to use information for improving schools' educational planning; and the construction of benchmarks was not proposed as the priority of evaluation but as the culmination of the research conclusions of the proposed agency. Whereas CENSIS attached great importance to these factors for explaining how to improve

---

[25] Minister for education between 22 July 1989 and 27 July 1990, in the Andreotti VI government
[26] Conversation with CRUI members
[27] Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di formazione
[28] From conversations with CRUI members

schooling contexts and learning outcomes, INVALSI abandoned the context element and has, since its establishment, primarily acted as an increasingly detached evaluator of skills acquired in given school subjects, mostly Italian, maths and English[29].

To summarise, SVIMEZ's and CENSIS' long-dating involvement in educational planning and evaluation research help to understand how international developments between the 1960s and 1980s may account for Italy's path of educational governance. CENSIS's international perspective importantly influenced its efforts to introduce the CIPP model (CENSIS, 1991) for school education. Researchers were convinced that a comprehensive evaluation framework would sooner or later emerge, just as it was emerging and quickly developing in other countries. An Italian articulation of the CIPP model was thus seen as a necessary, valid, and concrete path. When in 1993 CENSIS researcher Giorgio Allulli was hired by the CRUI committee, this perspective further ramified to the university sector, as chapter five discusses.

## 3.2. From manpower planning to management: Institutional management in universities at the IIEP and the CERI

In relation to the OECD's broader work, the MRP really reflected a success. Its projects were in fact stimulating quite rapid changes in countries' approach and direct application of planning programs (Papadopoulos, 1994). At the same time, however, these successes owed more to research developments in member countries than to concrete changes to national policies (Bürgi, 2016a). The OECD monitored developments in countries' educational planning through "educational development" reports, often also in a critical vein as was the case with its judgment of the UK's educational planning machinery, which was considered as "closed" and insufficiently attentive to the social effects of planning methods (OECD, 1975). Through the way that it influenced the MRP program, the study group had thus made important steps in developing human capital theory. As Italy's abovementioned use of the "economics of education" framework shows, the nexus between educational planning and economic growth was indeed embraced with little questioning of its deeper underlying assumptions. In many respects, the OECD did indeed turn planning into its key tool for uniting member States once the major inflows of US funding began declining (Bürgi, 2016a). When the "Educational Investment and Planning Programme" (EIP) was fused with the MRP in the mid-1960s (Bürgi, 2019), for example, planning was more insistently encouraged, to the point where, ideally, '[…] planning groups have to be located within Ministries of Education' (Papadopoulos, 1994, p.47). Further, the OECD's goal to raise its combined GNP by 50%

---

[29] To see how the agency operates, consult the INVALSI website at https://www.invalsi.it/invalsi/index.php

throughout the 1960s (Schmelzer, 2012)[30] acted as a key benchmark against which countries' developments could be assessed and further work on growth justified (Bürgi, 2016a). What began as a response to the Soviet Union's growth of the 1950s eventually led to a surge in the organization's combined GNP, which by the late 1960s increased by 55% (Schmelzer, 2012). If these successes contributed to the diffusion of the OECD's projects outside its walls, however, they later struggled to keep internal legitimacy for further projects intact. As historical studies on the organization point out (Schmelzer, 2012; Bürgi, 2016a), by late 1960s the OECD had lived a double crisis: an intellectual crisis of the growth paradigm (Schmelzer, 2012), and a legitimacy crisis of human capital theory (Bürgi, 2016a). As I argue, this moment turned the OECD's previous managerial revolution in educational research into a managerial turn in university governance.

Firstly, the attacks of the "ungovernable society" (Chamayou, 2018) emerging in 1968 also reached the OECD walls, where members were debating the real successes of the growth paradigm (Schmelzer, 2012) being embraced since the turn of the 1950s. As students and workers criticized the acceleration of capitalism and the legitimacy of those institutions fuelling it, the OECD raised internal questions on the "problems of the modern society", namely on the risks of replacing welfare with economic growth and ignoring the rise of environmental, political, and social issues tied to capitalism's extractive model of production (Schmelzer, 2012). Though not void of an instrumental re-articulation of critique, the OECD's "intellectual crisis" (Schmelzer, 2012) addressed the social risks of pursuing economic growth as an end per se, leading the new secretary general, van Leppen, to argue for a qualitative turn in the OECD's growth paradigm. This shift to the "quantification of quality" was meant to push the organization into using social indicators which could convey the social benefits or disadvantages of growth policies. In this way, it was argued, the organization would confront the by-products of growth and assume a more responsible and positive influence on global development (Schmelzer, 2012). As the way out of the problematic growth paradigm was found in the re-articulation of the methods through which it was framed, rather than as a long process which necessitated a deep questioning of its contradictions; the growth paradigm itself was thus all but abandoned. Alexander King's role as the 'intellectual author' of this debate importantly transformed the issue of the effects of economic growth into a case of their containment, with the organization eventually deciding not to deal with eradicating its causes. King insisted on a 'reorientation of the objectives' of growth (King, 1968 as cited in Schmelzer, 2012,

---

[30] During the 1960s OECD countries experienced phenomenal rates of growth, raising the combined national product by over 55% (and even 60% when including new members such as Japan and Finland). This period is thus consequently remembered as the height of the 'golden age'. The optimistic and technocratic outlook that had dominated the 1960s however began to crumble. Questions were raised, such as 'To what uses should this growth be put? If increased growth does not create improved conditions of life, will not growth become an illusion? What is the point of more unless more means also better?' They capture the insecurity of economic policy makers in the early 1970s about the goal of economic activity. (Schmelzer, 2012, p.1001-2).

p.1008), for he was convinced that a qualitative perspective on economic growth was possible. A qualitative take to growth would arguably allow to tackle its emerging social contradictions if the organization made the right efforts and studies to understand such contradictions, for example by gathering data on environmental impact and social inequalities. Indeed, King worked closely with the OECD's newly elected secretary general, Van Lennep, to develop a new series of social indicators, patching the intellectual crisis of the growth paradigm with a methodological escamotage which reinforced a positivist view of society (Schmelzer, 2012). This change reinforced the solutions which King was simultaneously proposing as a remedy to human capital theory's legitimacy crisis.

This second key turbulence in the OECD's work was the result of internal ruptures within the CSTP's study group. Countries' weak empirical results on the positive correlation between economic growth and national educational planning worried economists and educationalists alike. Vaizey himself, for example, exposed doubts on human capital theory (Woodhall, 1966) as further research showed inconsistencies with the theoretical underpinnings. Although member States were witnessing growth and devising educational plans, in fact, they were still unable to provide sufficient quantitative and qualitative data substantiating the success of human capital theory. Furthermore, countries' research on educational planning developed along the economics of education approach, scarcely and hardly translated into its implementation in national educational policies (Bürgi, 2016a) or in reformed planning sections in education ministries (Bottani, 1990), questioning further the theory's core assumptions and its ability to concretely influence social and economic change. The study group eventually dissolved (Bürgi, 2016a), and the CSTP responded to its lack of hard evidence through a managerial review of its research methods which re-appraised the usefulness of planning (OECD,1964). Differently to previous projects, the organization's future research would turn away from studies on national educational plans and shift to studying how planning functioned *within educational institutions*. This step defended planning research and methods as well as the CSTP's previous projects by blaming countries' lack of political will for the insufficient implementation of investments for manpower and economic growth. At the same time, the OECD argued that political hesitancy could be remedied by giving policymakers a much more detailed picture of how educational institutions concretely handled resources and produced graduates. This new rationale veiled the CSTP's underlying attempt to advocate the use of systems analysis as means to keep its legitimacy intact. Specifically, the idea that there be more research in institutional functioning was a way to try and revitalize the usefulness of planning in a setting different to those experimented until then, namely the institutional setting. When, in 1968, the CSTP was reinforced by the newly established Centre for Educational Research and Innovation (CERI), the two bodies' work picked up the operational research practices being developed by King

and others at the OECD's Directorate (DAS) (Bürgi, 2016a) to initiate a new working agenda on university governance. King's networks played a crucial role in the establishment of CERI: the project was seconded by the deputy secretary general, Michael Harris (Papadopolus, 1994), and readily materialized thanks to one million dollars from the Ford Foundation – and later also from the Royal Dutch Shell Foundation (Bürgi, 2016a; Papadopolus, 1994). The Ford foundation's resources not only crucially allowed to speed up CERI's establishment – which may have otherwise suffered from member States' possible resistances to its funding[31] – they also allowed the foundation to assert its own interests in the development of a new approach to educational planning within the OECD as well as outside it. These interests, in turn, were influenced by Philipp Coombs' work whilst head of the foundation's Fund for the Advancement of Education (FAE) and his active advocacy and support of a systems analysis framework in educational planning (Ydesen, 2019).

Between the late 1950s and the beginning of the 1960s, Coombs had mobilized diverse groups involved in the development of systems analysis. In 1959, for example, Coombs invited RAND intellectuals to develop work in education also at the foundation (Bürgi, 2016a), and, in 1961, he encouraged US and OECD planners to collaborate in the construction of greater educational expertise by hosting the notorious conference "Economic Growth and Investment in Education" (see Bürgi, 2016a for details on the conference). By 1968, Coombs was the director of UNESCO's IIEP as well as author of a very influential book on educational development, *The World Educational Crisis* (1968) and on educational costs *What is Educational Planning?* (1970). From Coombs' perspective, CERI's establishment responded to the concerns that he was raising on educational planning and reflected an adequate continuation of the IIEP's work between 1963 and 1968. Just like King, Coombs argued that human capital theory increasingly appeared only valid at a theoretical level and used this as a premise for justifying a managerial turn in educational planning. Writing in a context of intense educational expansion, Coombs argued that growth would hardly be achieved if educational institutions failed to adopt 'practical solutions' to the management of increasing educational costs, the priciest of which were teaching salaries (Coombs, 1968,1970). As he argued, costs would continue to rise without cost-efficient methods alternative to the widespread solution of increasing the supply of teachers, an approach that was seen as coherent with a humanistic yet conservative view of education which merely functioned to '[…] make a more effective claim on national budgets' (Coombs, 1970, p.22). Indeed, Coombs connected the student revolts of 1968 to schools' and universities' failures at changing the traditional means for dealing with student expansion, suggesting that if educational institutions merely accommodated

---

[31] Bürgi (2016a) points out how key members such as the UK and France were in fact contrary to CERI's establishment. As we will see, however, this did not exclude UK and French universities from actively participating at CERI's projects, pointing to delegates' and university representatives' contrasting interests in educational planning and their autonomous initiatives to do so.

higher numbers without rational management methods, then education would be unable to concretely fuel social and economic progress (Coombs, 1968). This narrative elevated the importance of institutional planning over that of national planning, similarly to King's own reappraisal of planning at the CSTP. The rationale of this argument lay in the interest of legitimating a systems analysis approach that would further solidify his years of work and experience, veiled with a discourse on the inadequacy of educators' ability to improve educational productivity. Coombs thus argued for major changes by elevating the importance of educational cost-efficiency –amongst other proposals such as curriculum innovation and technological advancement in the classroom–. Coomb's managerial critique relied on the belief that *internal planning effectively determined the outcome of national investments much more than the investments themselves* (see Coombs, 1970). In framing the success of investments as dependant on institutions' ability to distribute resources in the most cost-efficient activities, Coombs argued for a meticulous scrutiny of institutions' activities, creating the grounds for the use of a systems analysis approach to educational governance. Education's productive potential was brought back to a local-organizational issue – just as Kershaw and McKean had done a decade earlier in US schools (Kershaw & McKean, 1959) – and away from the macro-level, quasi-Keynesian approach used by human capital researchers of the Study group and the MRP. To further advocate planning, Coombs insisted on how Soviet success depended in no marginal way on its ability to frame education as a system where separate component parts need accurate and detailed coordination (Coombs, 1968,1970).

What eventually emerged from the valid preoccupations on the presuppositions of human capital theory was, I argue, a pragmatic use of discourse (Konings, 2010) which allowed King and Coombs to reassert their influence in a context which threatened to side-line it as the OECD's methods of pursuing educational planning and economic growth came under pressure. Rather than leading to a concrete problematisation of the theoretical assumptions underlying an instrumental view of educational governance, the OECD's planning crisis of the mid-1960s led to a reinforcement of output governance through King's and Coombs' ability to recast the issue as a *case of a widespread inexperience with meticulous planning methods* (Bürgi, 2016a). In so doing, the two agents crucially innovated on previous conceptions of educational planning. Whilst they still considered it key to the achievement of high manpower levels and consequently growth, King and Coombs mobilized resources and networks that could practically support – especially through funding – a new narrative of systems analysis planning as a still insufficiently explored field with considerable potential which hadn't yet been given the chance to develop. Their critique of planning was thus an opportunistic, discursive manipulation of educational planning's concrete weaknesses which was geared at entering the debate on the issues of educational planning to then actively shape

future outcomes. If agreeing on failures allowed the agents to side with OECD members during the planning crisis, proposing a different perspective on those failures enabled them to depict themselves as those agents of change which would bring about the necessary solutions. King and Coombs acquired in this way a leverage of power rooted in their ability to turn a significant constraint into a moment –and to an extent, an instrument– for legitimacy-building. As I argue, these agents' key innovative action lay in their success at *shifting the focus away from what had been done wrongly until then to what had not yet been done*. As the next two sub-sections discuss, this emerges clearly from the evidence that CERI and the IIEP immediately initiated systems analysis research on university governance. This step was an innovative reappraisal of planning, which articulated the usefulness of military systems analysis methods not in universal terms, but in a specific context which the OECD had still left unexplored: universities. The OECD's crisis thus resulted in a considerable shift in how educational planning conceptualised. Increases in educational productivity would now develop thanks to inward looking research on universities' resource management and goal achievement. The IIEP had already initiated this turn, and in 1968 CERI quickly seconded Coombs's view of educational innovation (1968, 1970, 1972). Eventually, both centres developed the idea that future educational planning necessitated a new group of educational planners working directly within educational institutions. As the next two sections explore, these planners would not be the educators, but rather administrators, who had the adequate skills for resource management, and university representatives, who had the right attitude to change. Together, these agents would succeed in reshaping the conception of university governance to the need of having efficient and well-performing institutions able to plan for and deal with constantly changing external environments. As the IIEP and CERI encouraged projects on performance assessment, resource planning and cost-analysis, they actively created the grounds for a managerial path to university governance rooted significantly in the principles of program budgeting. This practical development is explored below, where I analyse the two bodies' research on institutional management and how it, too, sowed the seeds for a managerial understanding of university evaluation specifically.

### 3.2.1. The CERI and the performance budget in universities

As mentioned above, the OECD's institutional turn in educational planning materialized as its work of the late 1960s engaged in '[…] training people in the new skills required to plan, administer and control rapidly changing educational systems' (CERI, 1969 as cited in Bürgi, 2016a, p.419). This occurred simultaneously at the Directorate of Scientific affairs and CERI, two bodies which in practice implemented the CSTP's managerial review of educational planning methods (OECD, 1964) in universities. Following the path of the management sciences, CERI welcomed program

budgeting advocates who could demonstrate its usefulness for university governance. This was no difficult task for King, who invited RAND intellectual Werner Hirsch to present his experience with program budgeting at the Directorate (Hirsch, 1968). Hirsch's work is particularly relevant, amongst the others presented at the Directorate, for two key reasons. Firstly, because it introduced the notion of performance assessment by proposing that universities implement the "performance budget" (Hirsch, 1968). Secondly, because, as I argue, it shaped the grounds of CERI's later projects on the use budgeting principles and practices in universities (Khan, 1971). Thirdly, it set out the idea that vice-chancellors needed to take charge of this process of transformation.

At the Directorate, Hirsch added three new elements to his previous work on program budgeting in educational planning first piloted in the US in 1965 (see chapter 2). Firstly, he now proposed that PPBS be applied also to higher education institutions. Secondly, he suggested that policymakers and institutions had to coordinate their plans. Thirdly, and most interestingly, he discussed the performance dimension of the program budget. This was part of his own experiments with the performance budget at the University of California, where he was professor of economics from 1963 till 1990. As he discussed at the Directorate's meetings on *"Budgeting, programme analysis and cost-effectiveness in educational planning"* in 1968 (OECD, 1968), the "performance budget" was a way to '[…] facilitate the judgment about […] the least-cost method of undertaking a given activity' (Hirsch, 1968, p.94). In so doing, his work stood out from other works which attempted to balance costs and benefits, for, in a radical way, Hirsch was fully focused on cost-efficiency. In his own words, the performance budget:

> '[…] is quite different from the concern of the program budget, which is to facilitate decisions about "what activities should be undertaken?" In one sense the performance budget is an ex-post budget whereas the program budget is an ex-ante budget. Thus, the performance budget seeks the least-cost method of undertaking a given activity; *it ignores the benefits, and the question of what activities should be undertaken.*' (Hirsch, 1968, p.94, emphasis added).

This conceptualisation of university governance as that task which 'ignores the benefits' of university activities was considered viable within a managerial framing of the university, which was first and foremost made up of cost-incurring activities which necessitated attentive monitoring. Indeed, university planners needed to mobilize adequate means for gathering an array of detailed information such as '[…] unit costs, e.g., man-hours […] often arranged by departments.' (Hirsch, 1968, p.94). ). The gathering of cost information was central to the achievement of cost-efficiency because it allowed to identify the most cost-saving and the less performant departments. At the same time, however, most universities of the time had such big departments (or faculties) that cost

information was too generic to fit well into performance budget. As Hirsch highlighted, universities would benefit importantly from a decentralisation of powers, for example by creating more but smaller departments from bigger ones. This would allow to gather cost information more accurately, simply by gathering it on a smaller scale. The performance budget would thus be eventually constructed by gathering as specific information as possible on staff use of costs, i.e., their performance. This matrix of decentralised powers and cost-information gathering, in turn, was key to incentivising workers into cost-efficiency, because in this framework they would gradually become part of a competitive struggle for the university's resources (Hirsch, 1968)[32]. This understanding of performance travelled to CERI's first projects, headed by British professor Hasley and organised by a restricted group of members including Alexander King. The "Programme on institutional management in higher education" was set-off in November 1969 (Khan, 1971), in a first meeting in Denmark. The program was discussed and developed over 5 meetings in which universities from the 'most advanced' Western European countries (France, Sweden, Denmark, UK, and later Yugoslavia) participated. The program did not aim to build '[…] a *generalized model* of university operations because of doubts concerning the practical usefulness of such work' (Khan, 1971, p. 14). Rather, it sought to initiate a series of bottom-up initiatives coming from universities, which were willing to test new planning and management principles. The aim was also to see if and how universities' own projects could enrich the work in progress with new inputs and case studies. Overall, eight universities were selected for CERI's program; with each carrying out team projects decided autonomously within the universities themselves. The British universities involved in CERI's pilot projects were Bradford, under the leadership of J.A. Bottomley, and Lancaster, under the leadership of M. Simpson. Bradford's project looked at the costs of 'producing' graduates across disciplines, with specific attention to marginal costing and the construction of forecasting models. The aim was to quantify how variations to a series of inputs (student numbers, teaching methods etc.) impacted the university's "production costs". Lancaster's projects also sought to analyse costs, in particular the project here sought to introduce principles of cost-constraint and budgeting (Khan, 1971). The program's final report saw the involvement and interaction of vice-chancellors as its key success. This was seen as evidence that the right group was being targeted, for VCs were the most influential agents in universities, able to bring about change. As the report Stated:

'[…] The most important turning point in the evolution of the programme has been the success of a conspiracy to get the direct involvement of the Rectors (now called presidents

---

[32] Here, the integration of new costing procedures in how universities used State funds received was therefore key, as well as a novelty compared to the CSTP's previous work, focused on extrapolating information from education sectors to project future manpower levels.

in France) and Vice-Chancellors of the universities that had set up the experimental projects under the programme. It is true that in many cases the executives had got involved […] but interaction among themselves as a group and the interaction of this group with the project leaders was brought about only after the organization of the steering committee in January 1971' (Khan 1971, p.11).

As I critically argue, King's determination to maintain his legitimacy at the OECD drove his attempt to institutionalise the use of military practices in university governance through the Directorate and CERI. King's choice to identify universities as the *loci* for building legitimacy around his critique of the OECD's previous unsuccess with educational planning is probably related to the fact that he had ties with the US academic world, though future research needs to investigate this claim in historical detail. Similarly, his efforts in the establishment of CERI and the development of its work were seconded and supported by an array of academics and OECD members – including the deputy-general Harris and his own ties with US academia and the Ford Foundation (Padapoulous, 1994) –, meaning that it would be distortionary to suggest, in a deterministic vein, that the rise of projects on performance assessment in universities were the result of a single man's power. What remains certain, however, is that by 1968 the Directorate was well-informed on the role of military networks in the intellectual development of PPBS and was determined to introduce the framework into its future research on higher education. At the Directorate's meeting in 1968, John Vaizey's contribution importantly attempted to reconcile previous methods of educational planning – which he had himself fostered at the study group – with the Directorate's decision to turn to PPBS. In his view, the OECD's turn to PPBS could be explained as the acknowledgement that educational institutions and public funding bodies needed to work together. As such, whereas in the research of the study group education was functional to the increase of manpower levels; it was now framed as strategic to countries' new resource allocation methods, with Sweden and the US as key examples. Thus:

> '[…] educational planning activities should be oriented towards the provision of information relevant to budgetary decisions. Budgetary structures and procedures, on the other hand, may need systematic and continuous revision to be in line with current policy objectives. […] Developments in this direction would also facilitate the more extensive application of such management techniques as cost benefit analysis, cost effectiveness analysis, the application of cost models, policy instrument analysis, and so on. These techniques, although still in the early stages of development, show considerable promise as tools for more rational decision-making.' (Vaizey, 1968, pp. 16-17).

### 3.2.2. The IIEP and institutional management

UNESCO's IIEP also used program budgeting to develop new institutional management practices. Concepts of university management which were arguably borrowed from the corporate world (Deem et al, 2007) were being articulated in universities, through the IIEP, much earlier than the rise of the NPM paradigm in the 1980s. Key to this was the series of seminars taking place between 1968 and 1975 (see IIEP, 1968; Jones et al, 1971; Onushkin, 1969, 1971, 1975;), entitled 'Planning the development of universities' and collected in four volumes (see Onushkin, 1971). Similarly to CERI, universities participated directly by presenting the State of the art of their governance models, with the very first being a comparison between the highly centralised university of Leningrad and the highly autonomous University of Sussex. Later reports looked at internal decision-making processes, the evaluation of teaching programs, sources of university finance (Vol. II) and specific case studies on the USSR, Germany, New York, Australia (Vol. IV). This extensive reach reflects the systematic efforts made to analyse planning structures and processes within extremely diverse university sectors through PPBS. Here, I focus on the research concerning universities' internal management.

The first Volume of the series (1971) described the project's birth at 5-day seminar in Paris between 7-12 July of 1969. The seminar discussed participants'[33] views of the most pressing issues in higher education to then gather applied case studies which sought to find solutions to the issues identified within universities. Participants agreed to draw up a manual for university planners and administrators to be used as guide throughout the program. Specifically,

> 'Two recommendations for the manual were made. Firstly, that if it is to indicate alternatives and options in management and planning, *different national structures must be taken into account*. Secondly, the manual should, if possible, deal with factors to be considered in *planning at every level*. Every effort will be made to meet these two specifications.' (Onushkin, 1971, p. 30, emphasis added).

Whilst CERI developed performance assessment in universities, IIEP research discussed, in a way, its potential in strengthening institutional governance vis-à-vis State policy. By emphasizing universities' dependency on State resources, the centre's research made use of program budgeting to elaborate on the notion of universities' accountability.

Firstly, the 1969 seminar discussed the concept of *accountability* in how universities used national resources.

---

[33] These were international organizations and university representatives from more than a dozen countries

'The university must fit the needs of the economy, and of the social system as a whole, if it is to justify the large sums of money being spent on it and to be spent on it in the future.' (Onushkin, 1971, p. 34)

It tied the issue of responding to social demand (i.e., growing demand for university places, increasing demography of younger people) with that of accountability. Importantly, it was discussed that a common issue facing all universities was having to deal with the limited resources that they received, often much lower than the high demand for their work. As a result, the principle of a funding constraint was turned into one of the program's key challenges, as was the case of the budget constraint in the more general PPBS framework. The program also discussed the quality of research -or, more precisely of 'research output'- where quality was intended as the usefulness of research to '[...] the economy and themselves [the research graduates]'[34]. Following a series of comparisons with the Eastern world that had already been studied in previous OECD and UNESCO projects, participants from the Soviet Union were also welcomed at the IIEP. By 1971 –and in contrast to the first steps being drafted in Western Europe – the Soviet Union was at a further stage of internal university planning. Some Soviet universities were ahead, experimenting more advanced practices such as curricula planning, making cost considerations on an equilibrium between independent study and university study time.

A second key aspect discussed was the leading role of the manager as that executive figure pushing for, and monitoring, university efficiency. Planning (Onushkin, 1971) was discussed as a very complex issue, because it was considered difficult to draw a strict separation between national plans and institutional planning[35] if the two had to develop in symbiosis, as the IIEP advocated. This seminar asked participants to consider the ideal managerial figure in a university, asking the following questions:

'How are the managers of university planning to be trained and who are they to be? Although it was generally agreed that this was an important question, few answers to it were

---

[34] Further, 'If we accept that of the knowledge gained by a university student during his studies about 10 or 15 per cent remains useful and directly applicable and that the rest is gained afterwar during his working life, what is the best attitude towards curriculum planning? Two solutions suggest themselves. Either one may attempt to include in the course as much concrete knowledge as possible, thereby slightly increasing the percentage of useful knowledge, or one may devise a curriculum which offers a wide general background. These two contradictory points of view are always represented when programmes are worked out, but the second attitude would seem more likely to produce students who are both useful to society and themselves satisfied with their own education. This opinion seemed to be shared by the majority of the participants.' (Onushkin, 1971, p.40)

[35] Further 'The essential, in planning terms, is to be clear whether we are talking of national aims or whether we are talking of mechanisms at the level of the individual institution. In discussion, it is possible to separate these elements, although in practice it is not. (Onushkin, 1971, p. 53)

forthcoming. It would perhaps be fair to say that there was a general feeling that an academic training was not a bad background for an administrator, but that aptitude and some further training is also necessary.' (Onushkin, 1971, p. 53)

University managers were thus ideally trained as administrators rather than academics, but answers were still being developed as to the skills required and the procedures of how managerial training could be promoted in universities. In part, the question '[…] how is it possible to move the focus away from academic prestige towards planning prestige?' (Onushkin, 1971, p.53) was answered by pointing out who the managers would *not* be.

### The legacy of a military lineage in research on university governance

These examples show that both Coombs and King "targeted" universities as those institutions through which planning methods could be institutionalised and eventually stick. Concepts of university management which were arguably borrowed from the corporate world were being articulated in universities, through the IIEP, much earlier than the rise of the NPM paradigm in the 1980s. The existence of these concepts in a radically different time frame and context emphasizes the necessity of carrying out further research on these international lineages. So far, it emerges how universities' experiments with these methods through practical pilot projects intervening in institutional organization increased the awareness of the planning models and references to them. This not only institutionalised the use of PPBS across diverse sectors, but it also legitimated the critique through which Coombs and King had advocated for more scientific approach at the OECD. The dissemination of PPBS across universities was a process which gathered a large and very diverse amount of "hard evidence" on university management and internal governance. By gathering networks of "educational planners" at the two international organizations, Coombs and King effectively constructed important steppingstones towards a managerial revolution in educational planning, and, more specifically, in university planning. By the early 1970s, educational planning in universities was beginning to take an increasingly 'internal' and 'institutional' nature, developing from macro-sectorial projections based on enrolment estimates to local administrative and cost-analysis planning. Thus, *university governance through cost and performance management* concretely began to take shape in these projects. This attracted those universities and agents interested in developing the administrative-managerial functions. More research would need to investigate the extent to which, at the participating universities, planning practices and groups of planners were still in their initial development or well well-elaborated. This would help to understand if these seminars were effectively facilitating participating universities' first engagement with such topics, or if there were yet other networks or lineages connecting these specific

universities to the projects. The projects discussed so far surely reflect the circulation of a management culture both within the international institutions and outside them, i.e., directly in the universities participating. Planning was being articulated as the management of university activities, as *institutional management* – a concept which, as chapter three will elaborate, gained important momentum in English universities of the mid-1980s.

Crucial to institutional management was the figure of a leader, identified in the vice-chancellor and administrator roles. Indeed, institutional management required non-academic knowledge, but rather financial-administrative one. By emphasizing this point, advocators carved the managerial expertise from within the university walls. Indeed, the managerial tasks developed at the projects did not present abstracted notions of management, (nor did they use business efficiency methods). Efficiency was articulated through PPBS's performance budget, resting importantly on the previous developments in RAND research and those at university planning in the US (Hirsch, 1968, Balderston, & Weathersby,1972). The association of costs to academic activities meant that the performance budget would be used to monitor academics' use of resources, in this way turning non-academic figures into the evaluators of the institution's financial health. This created a *distance of expertise* between university agents, where – similarly to military hierarchies – top hierarchies exerted a strategic – commanding – function over middle and lower workers rooted in defining the objectives and ensuring they were met.

Considering the research question driving this thesis, one of the most interesting elements of these projects is the presence of three British researchers and the work they presented in 1971 in the first Volume of the IIEP's publication series. This was a case study of Sussex University's management structure, presented at the IIEP by Mackenzie, Jones, and Lockwood (1971). Their presence in the project can be better grasped by considering Lockwood's use of a systems analysis framework in his broader research on university management, which has been defined as 'militaristic' (Lockwood, 1981). Of significant interest is the fact that Geoffrey Lockwood, administrator of the university, was part of the IIEP projects since the very start. His international interactions probably sparked, and certainly sustained, his PhD research on university management (Lockwood, 1981), a piece which, I argue, pioneered the use of systems analysis in British universities.

### 3.2.2.1. Sussex as a case study at the IIEP: the emergence of a British class of university managers?

During his PhD research, *An Analysis of the Planning Process in 1968-73 in the Contexts of the History of the University of Sussex and of the Management of Universities* (Lockwood, 1981), Lockwood published ongoing research (Lockwood, 1972a, 1972b, 1979; Lockwood & Fielden,

1973) on university management and on the changing role of administrators and registrars in planning for change in the context of English developments. This thread runs along his work, accompanied by the conviction that a flexible management structure in universities could help them to plan for periods of increases in State funding (see Lockwood, 1981) as well as for periods of decreases (see Lockwood and Davies, 1985). As he also emphasized, however, the success of a new management structure rested on an institutions' adherence to the principles of planning, budgeting, and performance assessment. Lockwood's' work was crucially influenced by his participation at the IIEP, re-articulating the framework of systems analysis to fit the British University context specifically.

In his PhD dissertation Lockwood (1981) carried out a detailed analysis of Sussex's planning and governance structure, Sussex being the first of the plate glass universities established in the 1960s[36]. Lockwood appraised the management structure at Sussex, and made clear recommendations for further improvement, which he tried to implement through his role as registrar. Lockwood's proposals argued to *improve the efficiency of central administration* in a context of the university's expansion. To this end, Lockwood was determined to enhance *the role of administrators* who had '[…] turned into planners' (Lockwood, 1981, p.95, §5.3.3), and had to be more acknowledged for '[…] all university costs […] can be calculated by known ratios and costing allowances from the number of students to be provided for' (Lockwood, 1981, p. 95, §5.3.3). However, a greater role for registrars and administrators could only be achieved if universities recognised that they needed to redefine their governing hierarchies and accommodate to simplified planning practices which could overcome the traditional committee structure where functions often overlapped. He insisted that through new planning structures, universities could become more efficient at allocating resources internally. In turn, this could lead to a chain of positive effects, most importantly, it would allow to accommodate for more students and, by extension, receive more external (i.e., State) funding. This was particularly felt amongst "new" universities, which felt unfairly treated. Whereas governments had advocated that these universities expand quickly upon their establishment, in fact, the introduction of external parliamentary audits scrutinizing and at times reducing universities' income projections from 1967 onwards (Shattock, 2012) risked leaving them in a situation of underfunding. Conscious of the thrust of change being brought about by these adverse developments, Lockwood argued for the necessity of devising university-run and long-term costing plans which could effectively compete with governmental plans and projections, thus

---

[36] University of Sussex (1961), University of East Anglia (1963), University of York (1963), University of Lancaster (1964), University of Essex (1964/5), University of Kent (1965), University of Warwick (1965). The term 'plate glass' referred to the modern novelty which these universities represented, such as their curricula, but especially their architecture, with plate glass, steel and also wide green spaces. Some of these were built from scratch, whilst others were colleges which were 'upgraded' to university status. (see Harold, 2003).

substantiating requests for more resources than what was being planned by parliament and the then funding body, the university grants committee (UGC):[37]

> '[…] if a university wishes to maximise its own influence on its future development it will be necessary for that university to plan to chart a path which takes advantage of national opportunities and changes' (Lockwood & Fielden 1973, p.111)

Upon the proposal of the then Vice Chancellor, Asa Briggs, Sussex University had undergone a management review known as the "Review of the University Government and Administration", carried out under the aegis of the McKinsey consultancy.[38] The review confirmed Sussex's privileged position in financial planning amongst other universities in England. Sussex was indeed '[…] less compartmentalised than the more traditional UK structure [which] avoids the narrowness of the old-fashioned department and provides more flexibility and cross-fertilisation'(Lockwood, 1981, appendix §7.4). The university was considered managerially mature enough to continue its managerial path. Further, the report proposed that Sussex could develop further in its managerial structure by combining financial and academic budget decisions (McKinsey report as cited in Lockwood, 1981). These efforts necessitated better defined and more incisive managerial roles, such as those of '[…] accountants, systems analysts' (Lockwood, 1979, p.318) which would support administrators. In agreement with McKinsey's suggestions, Lockwood[39] proposed the establishment of a planning committee and officer at Sussex (Lockwood, 1981) who consulted and advised directly to the Vice-Chancellor on university planning matters. The conclusions of McKinsey's review were coherent with Lockwood's and enhanced further managerial developments. Lockwood's active proposals for improving Sussex were thus welcomed by Asa Briggs, who possibly asked Lockwood to participate in the IIEP program (Jones et al., 1971) precisely because of his existing participation in higher education seminars of the IIEP (see Onushkin, 1971), which was radical in its approach to university planning and governance. As further analysis should consider, also Briggs' own military past may explain his awareness of developments which were then occurring at the IIEP as well as his secondment of the new PPBS principles at Sussex. Sussex university was publicised as a university with '[an] institutional atmosphere conducive to change and managerial innovations' (Lockwood, Mackenzie & Jones in

---

[37] Lockwood Stated for example that the UGC's quinquennial settlement for years 1967-72 was £650,000 short (1981, p.102).

[38] 'The review is of major importance in the history of the University of Sussex and […] the resultant structure at Sussex becomes a highly publicised model throughout the world (Lockwood, 1981, p.106).

[39] Lockwood defined McKinsey as '[…] the name in the public eye. Its list of clients covered major public institutions (The Bank of England, the post office) as well as the cream of British industry (ICI, Lever brothers, Shell, BP). It had a reputation as the renovator of British boardrooms' (Lockwood, p. 107, Section Two, chapter six "creation of the planning process").

Onushkin, 1971) in international circles, and Lockwood's research at international circles continued also beyond Briggs' stimulus. In 1972, the publication "university planning and management techniques" was published as part of a study of the OECD's Directorate, then headed by British war veterans Alex King and Ronald Grass. Here, Sussex's budgeting and planning system was compared to California's[40], the most advanced university with regards to experimentations on PPBS (Lockwood, 1972b). During these moments of interaction with OECD institutions Sussex thus entered a community of researchers and aspiring policy-makers keen to transform the university from within. As research on managerial governance in universities legitimated more research on performance assessment, planning and budgeting, these concepts became increasingly interwoven in the higher education context and gradually less-so by the military one: university research and research *on* the university importantly innovated on the RAND model which it borrowed from – creating a new articulation of the methodological axis of systems analysis. The 'managerial turn' in HE thus relied importantly on these figures and their ability at connecting each other's research to build more objective and universal models, whilst simultaneously bringing concrete "hard evidence" of change happening within their universities:

> 'Many members of universities seem to believe that the roles of their institutions are diminishing; *if they can be convinced that the opposite is the case*, and that the responsibilities and responsiveness of universities must increase as their importance grows, then the members themselves *may begin to understand, accept and encourage* the responsible development of institution-wide objectives and systems and the concept of continuous internal change.' (Lockwood 1972, p.96, emphasis added)

Just as Lockwood brought to Sussex the insights of the IIEP seminars, so did he bring to the seminars new evidence and arguments substantiating existing projects. Registrars and vice-chancellors could in fact quote the successes of other universities, and touch on fears of lagging behind or becoming politically instrumentalised by the State to accelerate a process of internal, voluntary change. Consensus building was, in other words, key. Effectively, Lockwood could be defined as what social network scholars call a (bidirectional) conduit[41] between international research on institutional governance and developments occurring in the British university sector. In

---

[40] 'The details of the systems, processes or models being developed at California, Sussex, Toronto or Yale may not be relevant to all universities, but the attitudes and concepts under-lying those developments are universally relevant' (Lockwood 1972b, p. 96).

[41] The research carried out so far has identified Geoffrey Lockwood, but it is possible that, amongst the British vice-chancellors, also other agents participated actively in both their own universities' management and the international circles. Lockwood's work remains particularly relevant for this thesis because of its influence in shaping a proper British current of organisational and management research in higher education which was specifically geared at the modernisation of British universities in the changing funding context of the 1970s and 1980s.

his dissertation, the systems analysis framework was treated as *the* rational instrument for coordinating the allocation of resources and the "output objectives" of HEIs (Lockwood, 1981), following similar U.S. experiments (Balderston & Weathersby, 1972) that he learned about at the OECD and IIEP.

The suggestion that ideas and principles on managerial governance using performance models travelled through agents in this specific context can be developed by turning to one of Lockwood's other publications: *Planning and Management in Universities: A study of British Universities* (Lockwood & Fielden, 1973), rather like a manual for British university administrators, was co-written with John Fielden, then consultant at Peat, Marwick, Mitchell & Co. This book tried to build consensus on university governance as a management activity. When compared with the book's guidelines for British higher education management, parallel British political discourse on educational efficiency[42] of the time appears embryonal and only vaguely *managerial*. The proposals then made by Lockwood and Fielden importantly clashed with conventional university governance, and three key points made in the book reflect their radically new approach.

The first is that universities be defined as productive organizations, in contrast to their traditional definition as social institutions[43]. This meant conceptualising universities as having precise productive goals (such as the production of a given number of "products", i.e., graduates, researchers, and research) to meet. A corollary of this shift in definition is the scrupulous analysis and monitoring of university activities to ensure that no costs went wasted if "production" was to be successful and maximise a university's capacity (Lockwood & Fielden, 1973), an increasing alienation in the name of '[…] objective and quantitative decisions, which are more realistic' (Fielden, 1973, p.17). Through this approach universities would prioritise strictly managerial thinking to their traditional social function of providing for all – the notorious 'social demand' principle (see Williams, 1982). From this organizational perspective, management and administration were skills facilitating the transition towards a different university, which could no longer be rejected or only partially implemented. The authors supported this view through a "mechanistic theory of management", where:

> '[…] administration is something like a production process. You define what you want to produce, and you devise means for efficient production. The production process is broken down into a series of measurable operations […] which ensures that the end product emerges at the minimum cost in resources expended' (Jones et al., 1971, p.273).

---

[42] see Prentice's statement in 1975, reported in the times Higher Education Supplement, 23 May 1975
[43] See Williams, 1982 on the conflict between the social demand function and the output function of a university.

Management practices were crucial coordinative instruments, gluing the operations together towards efficiency, conducive to the very understanding of the production process. This view made universities conceptually comparable to productive organizations, efficient universities depended on efficient management in place. Tied to this first element are *the other two key propositions* of the book.

A second key proposition was that of gearing universities towards implementing forward planning: the making of plans for the present by keeping future and changing circumstances into account. This was a form of planning that tried to control future uncertainties by pre-considering them; and was starkly different to how universities had been used to carrying out their financial matters. Throughout the 1970s, universities' financial plans were rather simple, partially because governing bodies could rely on yearly incrementing resources. As such, plans were often very similar to each other, with each plan keeping account of constant or increasing intakes of students each year. Further, financial plans were largely detached from specific academic considerations – for which senates were mostly responsible – and a ley criterion was that of allocating resources in a relatively balanced way amongst departments or schools. (Lockwood et al., 1973). Once drafted, plans were then quickly implemented, and made public to the rest of the university community (such as academic staff) only in a second moment as a sort of Statement of "what had been decided" in terms of financial allocations. Lockwood and Fielden, instead, sought to reformulate planning as a continuous process. In many ways, their proposal sought to extend planning beyond the administrative sphere. Planning, it was argued, was not a mere exercise of accounting for resources spent in previous years and use this as a framework for estimating future expenditures. Much more than that, planning was a *process* (rather than a "moment") which encapsulated financial and non-financial considerations, and thus depended on a harmonization of objectives between university groups to be successful. The crucial advantage of somewhat "expanding" planning in this way was to increase the functions of administration and provide grounds for its interference in non-financial aspects (Lockwood & Fielden, 1973). If planning was that combination of projects and sub-plans being devised by multiple groups – an exercise which could take over a year considering the drafting and approval phases – then it had to reach out to more university activities; and do so more meticulously than before. In this way, it was argued, all the university groups' objectives could be considered and included in the university's financial plan. Coupled with the insistence on establishing five-year plans as a more suitable structure for such a comprehensive method, their proposal created a convincing narrative of good and ideal resource management. Thus, for example, the decision of how much of the university's resources had to be allocated between different departments would no longer be detached from the function that those resources would then be used for. Quite the opposite, the authors advocated that all "units" produce "projects" to

report how they expected to use resources. Conventionally, information on departments' use of resources was 'predicted' or 'intuited' in general terms by administrators (i.e., the finance officers distributed resources taking for given that these were used on students and university resources) rather than being substantiated by micro-information on how a department addressed resources[44]. This intuitive approach is precisely what the authors sought to overcome, for, they argued, it made internal resource allocation excessively based on subjective predictions and assumptions about departmental spending, leading to a situation which impeded *real* cost-efficiency. It impeded real cost efficiency because the limited degree of information on units' activities, it was argued, (caused by reporting gaps) brought university decision makers to pay more attention to distributing a fair share of resources to all rather than to the elimination of potential 'wastes' or 'misuses'. In other words, universities were ill-equipped at *budget optimisation*. Forward planning was insightful and useful because it could turn the university into a flexible institution, ready to steer for positive and negative financial changes coming from 'the outside'. Establishing yearly planning was a way to begin constructing forward planning and familiarising committees with thinking in financial /economic terms, something they were not used to. However, a crucial difficulty of a planning cycle as that envisioned by the authors is that it would only be truly "forward looking" and "flexible" if it could gather an exhaustive amount of information and data from all the university's segments. (i.e., carry out forward planning). The opposite scenario ran the risk of leaving important information behind. The ability of the university to systematically gather information, and thus participation is a basic pillar for the functioning of the planning cycle.

The third element which the book saw as a necessary principle of organisational change is a shift in power balances within universities. Defined with the more inclusive term "*participative management*"[45] the concept conveyed the idea of planning as a collective action, a system in which all components ought to collaborate by drafting reports and gathering "essential information" on each other's activities. Participative management is the idea that thinking about the university in planning and optimisation terms needn't be an exclusively administrative task, quite the opposite, it ought to be extended to "units and levels" (i.e., faculties or schools of study and departments). What is interesting about participative management is that, on the surface, it appears counterintuitive to the conventional imaginary on the role (and privileged position) of the manager as the sole expert fit for taking organisational decisions. The manager is, typically, that person having the limited expertise for coordinating staff and resources to optimise their use. Hence it appears puzzling to find a managerial model in which it is argued that this skill – as well as the principles which lie

---

[44] For example, questions such as 'how did the department make use of the extra resources which were left over?' were not really made, for planning did not have a scope of monitoring as much as distributing resources broadly.

[45] Also referred to as "coordinative devolution" (Lockwood 1972, p.100). Interestingly, this concept was also used by factory managers at the turn of the 1970s as an attempt to dilute workers' rising dissent over exploitation (Chamayou, 2018)

behind it – be extended to 'all levels'. From a manager's perspective, the answer of this puzzle lies in what the extension of the participation to the planning process allows to obtain: information for decision-making. In fact, working with increasing funding constraints – thus working towards the optimisation and cost-efficiency of a university's resources – could only occur through a detailed analysis of all university activities where costs are incurred. The crucial reason for having participative management in decision-making is less inclusive than the term conveys. It is, rather, a rhetorical escamotage for the underlying aim of disciplining lower levels of academic staff through the control of information on its activities. The inclusion of all university agents in a "system" where cost-information is a primary input relies on these agents' responsibility of reporting such information in the first place, and of doing so following precise, pre-established planning templates. This participative aspect of the management model is more directed at dissipating adherence to a management culture within a university's "units", regardless of whether they cover managerial, financial, or administrative roles. For example, it was suggested that deans (which traditionally compose the academic senate) could also be appointed from outside the university sector, and hired not necessarily from academic contexts (Lockwood & Fielden, 1973). Additionally, Fielden and Lockwood argued that deans carry out cost-considerations and welcome lay members within their committees for major advice and expertise on financial management.

For each level of the university, and for each unit, a hierarchical composition was suggested so that each agent within a unit would have someone above her or him to report information to. However, whilst practices of information gathering extend through participative management, this does not translate – in Lockwood and Fielden's proposals – into an equivalent extension of the privileges enjoyed by the 'chief manager' to all. This emerges from the way that the authors treat the tenure of university staff. Senate's academic staff tenure was considered as supposedly harmful for staff performance, to the point that the authors suggested for its elimination. The justification was academic staff's natural conservatism, where '[…] the only scope for change is the vacancy of a post' (Lockwood & Fielden 1973, p.22). This was part of a more general attack towards those who did not see the benefits of objective, realistic information which management information systems could provide, indeed the attack was based on a traditional criticism which had previously been advanced to civil servants in government, namely that these, as academics and middle management, were jealous of their posts and tended to predominate subjective decisions to rational and objective ones. Further, all the information gathered throughout the months of the planning cycle would be received by the planning committee in order to be considered and judged by the vice-chancellor, whose key responsibility was to decide whether the units' planning proposals and information was relevant and suitable for the university's broader entire plan.

The managerial element of top committees' prestige lay in the vice-chancellor's team, composed by a planning committee and planning officers intended to oversee and monitor the overall information being collected, helping the vice-chancellor to interpret, elaborate and use the data for institutional-wide allocative priorities. Participative management, thus, was geared at the redeployment of powers between the groups within the university (traditionally the vice-chancellors, council, senate, the faculties, and the departments) such that hierarchies remained quite similar to traditional university governance structures, with the difference that planning committees would have more key responsibilities. This however implied major 'participation' of reporting work from traditionally academic figures, such as deans, originally working for the development of departments' teaching and research contents and organisations, as well as curricula, developing student wellbeing etc. Thus, these functions were secondary for the authors when proposing that deans act as different types of leaders.

Considering the content of the proposals, and Lockwood's active participation at the IIEP seminars in the early 1970s, it emerges how Lockwood and Fielden's publication on university management was their articulation of the PPBS principles and their attempt to convey the use for the British university context. When analysed within (and in part also through) this framework, the proposals reveal the outright attempt of legitimating a new institutional governance model in universities, one depending crucially on the roles of Vice Chancellors and their teams in decision-making, at the expense of the roles of academic staff. Indeed, through the management model proposed the authors radically broke with how universities of the time functioned. They sought to modify the key organization of administrators' roles through a more centralizing – yet decentralized – directionality of decision making (embodied in the planning cycle). Academic staff was asked to reshape its participation in decision-making, the importance of academic committees was effectively downplayed by asking academic staff to act as information- gatherers and providers. Considerations on students and staff were seen as relevant exclusively for cost considerations, with few digressions on academic themes or, similarly, the impacts of these mechanistic processes on the quality of academic teaching and research. Lockwood and Fielden treated information collection in institutional governance management as a nodal point. When considering the lineage of the PPBS model, the reasons for this become more comprehensible.

### Strengthening the role of administrators

The empowerment of the administrative-registrar figure was, as discussed, key to improved planning and cost-considerations. However, this was significantly dependent on an adequate "decentralized" organizational structure, without which information on costs and activities would have simply been transmitted to planners in an aggregate and hence useless form. For coordination

and planning at all levels to practically work in a compartmentalized view of the university, a management information system was necessary, or, more specifically, a "pyramidal structure of information". Generally, this pyramid is shaped by a large base of data, from which administrators at the top of the pyramid could extract necessary information. Administrators' elaboration of management and planning data would then form a more restricted and selective amount of data for "strategic" decisions to be made by them. Although advocating for coordinative devolution, the practical functioning of the new planning model required the concentration of data on costs, staff, activities, and student numbers in the hands of planning administrators/officers. These had the right to demand more and clearer information to be presented and organized frequently by lower levels such as lower-administrative staff and academic staff. This made it possible to develop resource allocation decisions which would create '[…]substantial behavioural and political effects in an institution' (Fielden 1973 p.18), such as for example facilitating greater scrutiny of the use of staff time and manipulate the balance of student-staff ratios. The fact that devising an internal management information system of this type created new administrative burdens on academic staff, was considered a superfluous aspect defined as '[…] the opposite side of the coin to participation' (Lockwood, 1971).

To summarise, ***the relationship between decision-making (strategy) and the execution of decisions (operation) was rearticulated into a practice of (budget) optimisation, one where it is not the power to define objectives that distinguishes the top levels from bottom ones. Instead, it is the capacity to evaluate the extent to which there is an assimilation (by lower levels) of the new criteria for defining academic objectives (such as new costing exercises) which creates status and power for top-level planners within a managerial model of university governance.*** Fielden and Lockwood brought a loyal interpretation of systems analysis in their proposals. Similarly to RAND analysts, they surpassed the traditional boundaries of scientific management where functions depended strictly on hierarchies. In this latter framework, overseeing the direction of a firm was the function of a top executive, whilst producing the goods was the function of the bottom, with managers overseeing the productivity of workers. Instead, notions of "flexibility" and "adaptability" to changing funding policies were used as to legitimate the view that running a university was no longer solely about fulfilling the charter's principles. The authors argued that running a university was a *modern* task concerned with quantifying its productive capacity and maximising it through planning. To this end, they worked to build consensus on the notion that universities necessitated a complex yet efficient feedback system – a 'process of projects' (Lockwood, 1971)– in which all university agents would depend on each other for their work. What stands out in this interdependent model (figuratively very similar to a cybernetic view of how a machine functions) is the framing of the university as a "system". The modern task was indeed that of co-ordinating university groups

and their activities similarly to how engineers co-ordinate the component parts of an electrical circuit to ensure its functioning. Clearly, planning a social system posed important challenges, for people can hardly be controlled as objects. Thus, the new planning model required an elaborated articulation in order to function. Interestingly, Fielden and Lockwood proposed a model in which they extended the participation of planning. At the same time, however, this was a veil for masking the determinant part of planning, a power which would rest in the hands of vice-chancellors and planning officers. Indeed, whilst lower university groups would be planning and justifying activities – the extension of planning – VCs and administrators would retain the power *to evaluate the extent to which plans and objectives and budgets could be approved or amended at the beginning and at the end of planning cycles*. This is crucial for understanding the type of power gradually which vice-chancellors and administrators gradually acquired. Fielden and Lockwood's book elucidates current researchers with a key insight into the beginnings of the managerial turn in university governance, one occurring *within* universities. As of 1973, VC's managerial power was not about *taking decisions on the operation of lower levels* based on the construction of a top-down, disciplinary, hierarchical structure. Rather, it was more about acquiring *the power to determine the space of action of lower levels*, to define and set the boundaries within which decisions had to be taken and formulated, not their content. Units and levels within devolved coordination had more formal than practical significance. Though presented in participative terms, the vice-chancellor's role is redefined rather similarly to an evaluator who judges how other committees' actions are justified in relation to the criteria and decision-making templates that these are given. These considerations allow to comprehend how performance assessment first permeated into the new features of university staff's actions through a redefinition of systems analysis inserted within a framework of planning, and not articulated as performance assessment. The advocacy for new planning models involved a reshuffling of power relationships within the university based on principles of objective setting and data gathering, crucially geared at handing top-levels the ability to evaluate whether lower levels were adhering to a system of planning as a project itself. From a wider perspective, this publication represents the beginning of a process in which the degree of alignment with administrative/bureaucratic functions was used as the justification for major executive leadership, and it is in *this* way specifically that the turn to quantification within universities was first used as a governance mechanism. Quantification was used a governance mechanism not solely because vice chancellors demanded university staff to achieve given targets (i.e., a 10% increase in student numbers, a 10% decrease in costs) but because they asked staff to use quantitative methods for defining their own objectives, with punitive repercussions if they failed to do so. The VC, thus acquired, within this new managerial model, two types of new powers. On one hand, the power to decide criteria and times of activities and to evaluate staff performance on

the basis of such criteria rather than on those of quality of teaching & research. Secondly, VCs could in this way retain a constant form of control through the power to frequent re-define criteria, justified on the basis of an externally changing and unfavourable context to which the university ought to adapt.

The intellectual roots of the institutional management model proposed by Fielden and Lockwood at the turn of the 1970s thus lie for an important part in the post-war international research on education undertaken at both UNESCO and CERI. The active collaboration between international organizations and national academics, vice-chancellors, educators, and economists fostered the development of proper national re-articulations of the systems analysis model strongly wanted by Coombs and King. In the UK, a first instance of the managerial turn in universities was sparked off at Sussex University in the early 1970s, when the PPBS framework was effectively articulated for the British university context and seemed to provide a guide for future administrators and vice-chancellors. These would now be, in many respects, legitimated to argue for technical analysis and performance measurement practices in universities. In Fielden and Lockwood's model, systems-analysis principles were used to articulate planning as a monitoring and evaluative exercise, bringing evaluation increasingly close to governance concerns as a means for aiding cost-efficiency judgements. If the development of research at the IIEP importantly informs conventional assumptions on the roots of the British lineage of university evaluation criteria, this is less the case for the Italy's turn to educational evaluation, which as discussed owed more to developments in the use of the CIPP model in schools.

## 3.3. The Vice-Chancellor committees in the new European community

### 3.3.1. The Conference of European Rectors

The development of institutional management also ramified to universities through vice-chancellors' trainings at the Conference of European Rectors (CRE). As these seminars developed with time, the CERI and IIEP soon became only some of many channels through which managerial governance spread to university walls. Between the mid-1970s and mid-1980s, in fact, vice-chancellors began appropriating notions of managerial governance crystallizing the notion that they were first and foremost executive leaders of universities – increasingly defined as "complex organizations". The CRE's engagement with notions of institutional management set-off in 1969 in the immediate aftermath of the student movements, whilst its "management seminars for newly appointed executive heads of universities" (CRE, 1994) were launched in 1977. The latter were central to what I conceptualise as *the first instances of the emergence of a managerial expertise in the university context*. Differently to the CERI and the IIEP, VCs' participation was the core of the

CRE's functioning, making it a locus where university issues were debated from their specific perspective. Although this occurred also long-before the mid-1970s, the management seminars are exceptional in that they represent the CRE's first attempt to influence university developments not solely through its conferences and debates; but by influencing vice-chancellors' behaviours, starting from a re-articulation of their role within universities. In seeking to train them as "agents of change" (CRE, 1994), the CRE seminars *extended new knowledge outside CRE walls by concretely transmitting new practices to VCs*, and, through them – it was hoped – to single universities. Exploring the CRE's management seminars allows to grasp two key aspects: their emergence in response to the student movements and the intent to transmit strategic management through the agents of change, the vice-chancellors. Management emerged thus as a response to governance turbulences, and vice-chancellors as those figures taking the lead in turning universities into modern organizations.

As the communist uprisings of 1968 set-off, fears of de-legitimation by a student body determined to disrupt the crystallization of capitalist university model were addressed in the 1969 meeting in Geneva. Here the CRE really began to work on shared principles[46]. This meeting was organised by the CRUI, initially meant to host the conference in Bologna (Corradi, 1998). It was significantly centred around university autonomy. The latter was not just a concern for the CRUI, but also a way for university leaders to reflect on building a new type of university governance that could resists to the tumultuous context of the 1970s. For the CRUI, the CRE debate on university autonomy allowed it to develop its views on how to achieve it in the Italian university sector. For the CRE, it also marked a first step of its turn to issues and practices of institutional management in universities. Autonomy in fact ignited discussions over the types of institutional powers and ways to secure them. University leadership and appropriate organization of its resources could importantly guarantee that universities remain relatively free to set their own development paths. The 1968 student protests fuelled these reflections and symbolized a broader conjuncture of political, social and economic events for many university sectors living left-wing dynamism and revolt in Western (especially in France and Italy) and Eastern Europe . They fuelled university heads' preoccupations over sectors' radical expansion, the integration of Eastern universities, and newly-emerging State demands – with many States asking universities to equip themselves to quickly accommodate exponential expansion (Barblan, 2002). Vice-chancellors tried to picture their role in a future, much larger and diverse, global higher education context. The massification of higher education could not be resisted, particularly in the solidifying democratic and social-democratic context of the reconstruction period. At the same time, however, vice-chancellors needed to come to grips with

---

[46] As Barblan notes (2002, p.5): '[a]part from General Assemblies, and until 1969 and the repercussions of student unrest, there was no discussion of common problems shared by all CRE members.

how to manage the process, for, it also symbolised the arrival of new competition between more-established universities –those most represented at the time within the CRE – and future "modern" ones. Massification and the eradication of élite tertiary education implied that States would become more involved in the sector to sustain the educational expansion necessary to keep industrial development and economic growth going. This meant that vice-chancellors needed to consider how welfare State logics could potentially downplay their traditional and collegial-oligarchic governance more typical of small to medium universities. Vice-chancellors were, in other words, directly concerned with how their university sectors were developing, turning university autonomy became into both an important principle and instrument for future university governance. On one hand, it supported the expansion and massification of the university sector through principles of freedom and academic governance as the 1968 movements and various student collectives demanded. On the other, the debate on how vice-chancellors should deal with expansion was also about how university autonomy could be that device for re-gaining or maintaining control over academic and student bodies in a context of social uprising. VCs could, for example, claim autonomy from governments by insisting that they were apt for leading universities away from the expansion crisis and into 'operational growth' (Barblan, 2002). This became increasingly possible as vice-chancellors acquired managerial training, and, thus, new knowledge that could assert their role in universities.

Although the protests were harshly attacking the view of vice-chancellors as executive leaders, management principles allowed the CRE's vice-chancellors to practically build counterarguments and actions to the attacks being raised to their figure within the university. In 1969 the CRE's sessions began to practically set up its views on university management as a VC's restricted responsibility, turning to '[…] comparative academic development […] coupled with thematic seminars […] in which key questions of university governance would be discussed.' (Barblan, 2002, p.6). The CRE's managerial turn in 1969 specifically, when the CRE decided to detach itself from the Council of Europe (Barblan, 2002), can be considered as VCs' reaction to the context of protests. The seminars began as trainings, and later developed into concrete moments of comparisons between universities' use of management models. As Barblan, secretary general of the CRE between 1976 and 2001, has explained, the CRE's first seminars:

'[…] were described by some as "continuing education" for academic leaders who, after the reforms resulting from the 1968 student troubles, were asked by new regulations to get more and more involved in the detailed management of their enlarged institutions. This also had the consequence, very often, to change academic leadership, many of the prestigious scientists of older days being reluctant to commit four to eight years of their life to

university administration – at the risk of jeopardising their scientific career.' (Barblan, 2002, p.6).

As a new '[…] breed of academic leaders, usually younger staff […] interested in the university as an enterprise' appropriated notions of strategic management in the 1970s (Barblan, 2002, p.6), so the CRE's management seminars developed and became more detailed and incisive within universities. Notions around setting off the attentive monitoring of university activities in terms of their costs, performance, organisation also diffused. In later seminars, VCs were indeed trained to act as the formulators of such questions within their universities (CRE, 1994).

### Training new "agents of change": principles and methods

In 1978, the timid debate on institutional management took a decisive turn. Thanks to resources mobilised by the INSEAD management school in Fontainebleau (Paris) and the collaboration with CERI, the CRE piloted its first (and small) management training seminar with just twelve participants (CRE, 1986c)[47].The OECD's *Institutional Management in Higher Education* (IMHE) project[48] had begun in 1969 within CERI (IMHE, 2002), and by 1972 was defined as '[…]an independent and decentralised project' (IMHE, 2002) which was mimicked outside the OECD through similar program outlines and content. What most interested the CRE was indeed the program's attention to internal management procedures and practices, which to many vice-chancellors were still new. The research being developed represented a valuable knowledge pool from which rectors could be encouraged to learn and debate. For the organisers, one of the key strengths of the seminar was in fact its "multiplier effect", whereby the diffusion of new knowledge on university governance could occur as individual rectors reported the trainings back at home, developed context-specific case studies using the new skills learnt and translated the seminars in their own language (CRE, 1994). The reasons for training VCs as managers were multiple. From the CRE's perspective, the seminars were a chance to mobilise rectors around new notions which were developing, which allowed the conference to become a key reference point and a legitimate actor in international university debates. From vice-chancellors' perspective, the seminars were indeed very experimental, and the fact that they were developed to suit as many diverse organizational structures as possible meant that vice-chancellors could learn from common similarities and differences. Also, the seminars gave a special place to rectors as those who were entitled and able to bring about changes for dealing with expansion and changing State policies. As

---

[47] Vice-chancellors came from the universities of Ireland, Greece, Norway, Switzerland, Germany, UK, France, Turkey and Austria.
[48] This project later developed into a Journal on Institutional Management in Higher Education. See https://www.oecd.org/education/imhe/

such, in the trainings  vice-chancellors could develop skills that would allow them to argue for the use of new practices in universities. The fact that seminars gradually became more popular made the diffused practices of institutional management more easily justifiable in home universities on the basis that other universities were doing the same, stimulating what in Powell and DiMaggio's terms (1983) would be a homogenisation trend within the field. Between 1978 and 1995, the management seminars experimented with different topics around the role of the rectors. In the 1990 seminar in Norway, for example, the attempt to equip vice-chancellors with consultancy skills and help them 'shop for solutions to their own management problems' (CRE, 1994, p.17) failed to stick, as did the attempt to introduce TQM principles, seen by participants as a form of evaluation rather than as a device more strictly concerned with the specificities of management processes. The 1986 seminar held in Maastricht (CRE, 1986a) is a good example of the CRE's work on managerial training.

The 1986 seminar was headed by John Davies, who had already participated as a management consultant in the CRE's Fontainebleau seminar. At the time, Davies was professor of education management at Essex University as well as co-author with Geoffrey Lockwood of *Universities: The management challenge,* published in 1985 (Lockwood & Davies, 1985). Davies had an organizational perspective of higher education institutions and was particularly concerned with exploring how they interacted with their external environment. Much of Davies' own work borrowed from the US context, where the "entrepreneurial university" model was much more developed. Following these lines, Davies explored university constituencies, developing managerial strategies which could best help them to interact with and stakeholders and simultaneously allow universities to establish themselves within larger university contexts where '[…] the supply of higher education outstrips demand' (Davies & Melchiori, 1982, p.88). Suggesting that universities needed to surpass their traditionally 'closed' relationship to the non-university world, he encouraged vice-chancellors to think of themselves as "policy formulators" who needed good management skills for tying relationships with different stakeholders and raise universities' public reputation. In the "five-stage managerial scheme for image building" (Davies & Melchiori, 1982, p.94, see figure 4 below), for example, he outlined how institutional management increasingly depended, first, on university's engagement with market research, namely research on other universities' position within the sector as well as of the key stakeholders in the 'outside world'. Thus, the university market could be used similarly to a source of information gathering, and much less as a platform for competition. By participating in market research especially, institutional managers would have more information for adjusting university planning in order to be more visible to the rest of universities and stakeholders, of course. At the same time, the market was not part of the 'university constituencies' (see appendix A) and in the image building model the market was

also being framed as an instrument of data gathering. This behaviour would also improve their scarce ability to carry out long-term resource planning. As universities developed these skills, they could importantly improve their position vis-à-vis that of other universities, attracting relevant resources from stakeholders, for example. In this approach the actual development of a university's institutional program came at stage four, following stages of data gathering and analysis, and was carried out by central management. For Davies, university development was therefore about balancing internal objectives with the necessity of developing public relations with groups outside the university (including other universities). Similar to how PPBS broadly conceptualised institutional management, these methods allowed universities to reduce uncertainty about the future through planning, for their changing external context – and mostly expansion, changing State funding and changes to university policy – could be better controller if universities had information about their environment and engaged in its networks and activities. If public relations could develop mimicking US universities' entrepreneurial character, internal plans and objectives could improve if vice-chancellors learned to find those management strategies more suitable to their university's governance model.
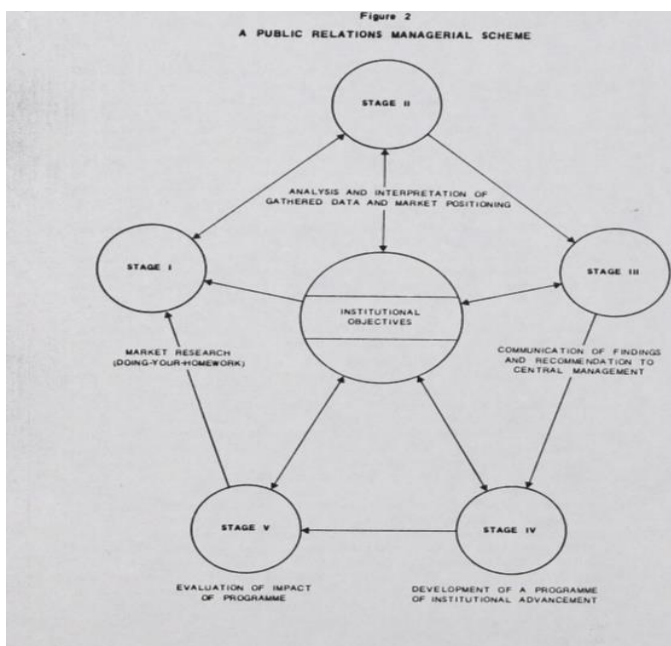


*Figure 4: The managerial scheme for image building (Davies & Melchiori, 1982, p.94)*

Participants were asked to provide detailed information on their governance structure – such as groups, financial and non-financial endowments, procedures of election (CRE, 1986b)– as this allowed them to identify different types of decision-making accordingly. To guide the exercise, Davies used a management textbook, *The project planning situation*, published by the Human Synergistics consultancy. The consultancy had devised a "Human Synergistics Circumplex",

model[49] to identify organizational performance. Davies used the circumplex model to categorise different organizational structures in universities: bureaucratic, collegial, political, organised anarchic and business enterprise. Through these categories, vice-chancellors were encouraged to reflect on the type of meetings and 'organizational behaviours' which dominated at their university, amongst management, council, committee or bargaining types of meetings (Human Synergistics, 1985, as referenced in CRE, 1986c). As vice-chancellors located their university in these axes, they could thus find their specific leadership style. So, for example, a bureaucratic university worked best if vice-chancellors focused on information control and thorough reporting, whilst in a political university its role needed to focus on mediation and negotiation, and in a collegial university it was advised to create consensus for specific goals prior to meetings or to set up working groups for carrying out goals. What was at stake in these seminars was giving vice-chancellors concrete instruments for, broadly, 'getting things done', to help them mobilize 'tools of power which can be displayed or created for the particular model' (CRE, 1968c, p. 153). Indeed, vice-chancellors were being trained on how to achieve consensus for their role within their governance structures. This also involved helping them to advance new management questions and procedures by tailoring their way of doing so to the diverse power dynamics surrounding them. The Synergistics textbook outlined key management activities which vice-chancellors were encouraged to prioritise, such as measuring progress and performance, develop strategies and coordinate activities, amongst others. If interested in taking up these roles, vice-chancellors could not merely assert them within governance bodies, but needed to show their importance (improving public image, for example), and find means to convince other university agents that they were able to carry them out. Although institutional management notions were diffusing in the 1980s, in fact, this was still an ongoing process, which often involved universities carrying out new organizational experiments. Interestingly, the seminar did not display management as a single and easily definable activity, but rather as a varied and adaptable skill which could be tailored to different contexts. This ubiquitous declination of  university management gave the seminars an extensive reach. Different universities would not feel left out from a dominant view of university governance, for the seminars gave a managerial response to many types of organizational dynamics and obstacles to managerial leadership. In this way, the seminars did not attempt to have universities converge into a single type of management model as much as encourage them to set up management procedures in the first place. The universal appeal of managerial skills in other words helped to advocate for their establishment in universities, vice-chancellors from different contexts would not be left out if their university did not fit the typical US axis, for example, for they could find answers to their specific situation.

---

[49] See the consultancy's circumplex at https://www.humansynergistics.com/about-us/the-circumplex.

The seminar also borrowed from CERI's research to discuss universities' 'goals and objectives, program coherence, and course contents', and new indicators on grants, PhD students' success rates, number of applicants and consultancy contracts (CRE, 1986c, p.278). These served both internal and external goals: internally they gave management the tools for developing programs and monitoring how agreed projects developed. Externally, they helped universities participate in ratings and signal their "market positioning" acting also as an incentive for improving internal resource management and activities in order to be more competitive and attractive to stakeholders. Here too, the discussion on indicators remained rather broad for some vice-chancellors were being introduced to their content and scope for the first time. The role of the vice-chancellor in 'diagnosing the effectiveness of institutions' was that of spreading awareness on the types of indicators possible and the ways to implement them. Further, the vice-chancellor was that figure which would use them concretely to assist the formulation of university goals. Thus, rectors needed to act as 'guardians of quality' (CRE 1986c, p.281) who could spread an evaluation culture and use indicators to develop ways for strengthening the university. Instead of encouraging the use of indicators for the scope of resource allocation as Fielden and Lockwood did, Davies focused more on how evaluation allowed rectors to improve the university's image and use evaluation as an instrument to set-off institutional change. Thus goal-missions were importantly fed by indicators and followed through thanks to vice-chancellors who encouraged their use for achieving missions in the first place. Rather than concretely developing the indicators themselves, vice-chancellors needed to work with administrators and external agents to identify them to then diffuse their use within the university walls.

Through seminars such as the Maastricht one, CRE vice-chancellors asked themselves new types of questions regarding the universities that they represented – or, rather, managed. The new questions being raised necessitated new solutions in the form of rectors' behaviours and practices. This encouraged a process of testing of new experiments on university governance – similarly to that taking place through the IIEP and CERI. Though often defined as an a-political group, the debate as to whether the CRE 'represents vice-chancellors or universities' (Corradi, 1998) is still open, albeit less ambiguously than in the past. If on one hand the CRE represented vice-chancellors' views; on the other these often also spoke in the interests of their universities' development. The CRE's expansion into the EUA's new policy network has blurred the lines of this debate: the EUA defines itself as the representative of over 19 million students and places itself at the centre of formulations of university policy. At the same time, its role as a university consultant conveys a specific normativity as to how universities best function which rotates around a managerial conceptualisation of university development. In its recent participation in the NEWLEAD project, for example, the EUA '[…] aims to build the capacity of university leaders in steering change and

in addressing new priorities on the institutional transformation agenda'[50]. The EUA is thus building on to the narrations which the CRE first developed throughout the 1980s in its management seminars. In leaving the epistemological grounds of output- governance unproblematized, it takes a clear political stance on university governance, one which remains unaffected by the increasing international contestations to managerialism in the sector. The EUA's proliferous work on evaluation standards and multi-level governance, further, signals the association's support of managerial logics also outside the institutional realm, implicitly validifying the principles of the NPM paradigm. It carries out numerous projects on internal quality assurance systems geared at homogenization of quality standards across the universities involved, such as EQUIP[51] and the European quality standards (ESG)[52]. The EUA's reach extends also to non-European universities, with project MIMIR[53], for example, geared at the development of institutional practices in Arab universities. For the EUA quality assurance is a thus a key device for constructing and developing new vice-chancellor networks, a process which strengthens the EUA's prestige and reputation. The EUA's recurring attention to these issues and its extensive reach has effectively turned the organization into a key partner of the European Union as well as of countries' education ministries, partnerships which are mediated through vice-chancellor committees and/or research networks. If we take a step back, the CRE's very birth was closely intertwined with European politics.

When the Brussels pact[54] evolved into the Western European Union in 1954, the CVCP and the Duke of Edinburgh – both part of the pact's cultural committee – organised two conferences which would eventually mark its foundation. Invited to Cambridge, ninety-three university representatives and eighteen national representatives from fifteen countries (Barblan, 2002; Corradi, 1998) agreed to set up a university cooperation independent from both national and supranational governments, where university heads would lead initiatives with '[…] complete freedom to express their opinions on any topic regarding universities, without political control' (CRE statement as cited in Corradi, 1998, p. 205). Still, the CRE's birth was imbued from the start with specific political goals. The CRE's first steps were driven by the post-war military alliance's four D's - decartelization, demilitarization, deindustrialization and de-Nazification (Ruegg, 2011). These concepts had been elaborated as part of the Potsdam agreement's education program with a

---

[50] https://eua.eu/resources/projects/793-newlead.html
[51] See https://eua.eu/101-projects/569-equip.html for a report of the project
[52] See the report on the ESG standards here
https://eua.eu/downloads/content/standards%20and%20guidelines%20for%20quality%20assurance%20in%20the%20european%20higher%20education%20area%20esg%202015.pdf
[53] See the project at https://eua.eu/resources/projects/585-mimir.html
[54] Signed by Belgium, France, Luxembourg, the Netherlands, and the UK on 17 March 1948

particular focus on "educating the Germans"[55]. The victorious powers '[…] were to 'allow' and 'control' a successful transition from nazist and militaristic doctrines to the development of democratic ideas, but they were not to bring it about themselves' (Rüegg, 2011, p.78). Between 1948 and 1955, university representatives had attempted to take-on the goals of the educational program. The British had played a key role in the de-Nazification process, formulating lists of German professors to be fired and/or monitored depending on their previous actions during Nazi rule (Ruegg, 2011). Further, VCs had attempted to construct a more specific program and turn universities into new value-dissipating institutions by proposing a "European association of university heads" in 1949. However, the Council of Europe's refusal of this on the basis that it excessively marginalised member State delegates in its own proposals through the Committee for Higher Education and Research (CHER) led representatives to further engage in bilateral agreements to the point that member States' political involvement was close to marginalised[56]. The 1955 conference – held in Cambridge from 20 to 27 July – therefore officialised earlier attempts to facilitate representatives' influence in post-war educational policy. The liberal-democratic reconstruction inextricably blended with emerging views on post-war university governance, feeding what later became the dominant view of the CRE as an a-political organization. Following the Potsdam agreement's research bans on industrial and military-intensive sectors (physics, chemistry, aviation, and shipbuilding), and the consequent removal of research staff and facilities (Ruegg, 2011), the new CRE tried to articulate shared principles to reduce unintended outcomes. In its first decade it discussed types and degrees of convergence between different university contexts, reflecting '[…] the differences between a closely knit Europe of sovereign nations [….] and a Europe whose member countries were ready to abandon part of their sovereignty to achieve common aims' (Barblan, 2002, p.4).

If on one hand the CRE made use of the IMHE's work by directly collaborating with CERI, this does not imply that individual VC committees followed a homogenous path towards the use of new institutional management practices simply because they may have participated in the CRE's trainings. Indeed, the turn to managerial governance is more complex and depended importantly on national, context-specific factors, too; such as university policy, VC committees' leverage in the sector, and VCs' openness to new governance strategies. Before discussing the configuration of these specific features in detail in chapters four and five, the next two sub-sections briefly introduce

---

[55] Further, the alliance monitored the developments occurring in German universities, for example by organizing regular meetings with the heads of German universities between 1952 and 1957 (Rüegg, 2011), driven also in this case by the crucial push of the British, through the control commission.

[56] The culmination of an autonomous political influence over universities was indeed successful as the CRE maintained exclusive powers. To give an example, '[b]ilateral agreements between the national rectors' conferences on equivalences of diplomas therefore remained until the 1990s the usual procedure for the mutual recognition of studies in foreign countries' (Ruegg, 2011, p. 24).

the British and Italian VC committees and then contextualise their actions in the context of international developments so far.

### 3.3.2. The Committee of British Vice-Chancellors and Principals

The Committee of British Vice Chancellors and Principals (CVCP) is characterised by two key features. These are its exceptional autonomy and its active engagement in articulating the post-war university sector. The English University sector is internationally renowned for its traditionally autonomous character (Shattock, 2012; Shattock & Berdahl, 1982). Between 1906 and 1988, universities were funded through an intermediary body, the university grants committee (UGC), composed of an academic majority. The UGC was dissimilar to ministries in European university sectors. It was never fully responsible for the management of State resources, it decided the funding sums for each university and then leave universities to freely decide how to distribute its resources internally. For this reason, university vice-chancellors in the UK traditionally have a strong degree of control over State resources, which, as chapter 4 explores, was threatened in the mid-1970s.

Up until the sector's expansion of the 1960s, the CVCP's role was based on an elitist view of higher education, which, it has been argued, is a thread for comprehending the committee's stances in higher education. When universities' monopoly over the power to award degrees was challenged by labour governments in the 1960s, for example, university vice-chancellors felt 'de-classed' (Tapper & Salter, 1978) Accordingly, the sector's transition towards mass higher education was characterised by two clashing: the "traditional elitist" ideal of the university as a locus for a restricted and high-class group serving the interests of society; and a "social demand" view which threatened this organization, as political attempts to radically enlarge access to higher education grew. Occurring particularly following the second World War, new political thrusts such as the expansion of technological colleges' and demands of efficiency created a situation of "elite ideologies in conflict" (Tapper & Salter, 1978). These broad changes put vice-chancellors' traditional, 19th century autonomy under pressure, revealing its boundary dimension (Neave, 1982).

The CVCP was formally  established in 1930 upon the proposal of the then Universities Bureau of the British Empire (Silver, 2003). Up until then, vice-chancellors were informally grouped, since 1918, closely to the Bureau as the "standing subcommittee of Vice-Chancellors and principals". The reasons behind the replacement of the standing subcommittee with a new body, the CVCP, concerns an array of developments which occurred between 1903 and 1930 that do not concern the scope of this thesis, and have been discussed in detail by Harold Silver in his book '*Higher Education and opinion making in twentieth century England*'(2003). An important insight from Silver's analysis which must be pointed out, however, is that the CVCP began to actively intervene in political decisions on the construction of a nationally supported higher education

system starting from the second World War period (Silver, 2003). This happened as universities and their unions began to acknowledge the existence of political intentions to make an instrumental use of universities in the post-war period. This resonates with other studies (Tapper and Salter, 1978) which have grasped the CVCP's active attempts to articulate the actions and responsibilities of universities in relation to the post-war 'national interest':

> '[…] universities may properly be expected not only individually to make proper use of the resources entrusted to them, but collectively to devise and execute policies calculated to serve the national interest.' (CVCP, 1946, as cited in Salter and Tapper, 1978).

Indeed, the CVCP did all but sit and wait as politicians set out new grounds for an expanding university sector, Well aware of the Percy report[57] – which paid significant attention to education's role in augmenting the country's technical productivity in light of its lagging position in relation to the US and the USSR (Percy, 1945) –, the committee '[…]was interpreting the 'climacteric' in financial terms, but with a strong sense of purpose and policy' (Silver, 2003, p.85) and advocated for radical increases State funding. Through more resources, in fact, the committee could try to retain autonomy and reduce its future dependency on State policies. By insistingly asking universities to engage in 'co-operative planning of the University system' (Silver, 2003), the CVCP gradually became a 'national negotiating body' (Silver, 2003) which actively advanced policy proposals to government and spoke out in journals such as *Universities Quarterly.*

### Steering research policy in a context of change

The CVCP looked at US developments to follow its position. This played out also for British universities' research policy. The committee was witnessing how US research policy was strategically focusing on the technological and engineering sectors. When drawing international comparisons of research capacities within the most advanced Western countries, the CVCP concluded that whilst the UK was an important "producer" of technical and scientific researchers, it struggled significantly to employ them in industry (CVCP,1968). Its response was to develop British committees and key scientific research funders such as the Committee for Scientific Policy (CSP) and the Committee for Science and Technology (CST). The key attempt was to exert leverage over these groups in order to follow the US example and reduce its lagging. It was Sir Frederick Dainton in particular – chemist by formation and university administrator – who

---

[57] 'The position of Great Britain as a leading industrial nation is being endangered by a failure to secure the fullest possible application of science to industry' and that 'this failure is partly due to deficiencies in education' (Percy, 194, p.5)

participated in a key debate on the relationship between State and university responsibilities in shaping university research.

The CVCP had a copy of the Adams report (Adams & Murphy, 1967), a publication in which NASA officials discussed pathways for implementing research contracts between the Agency and universities' atomic and aerospace research. In the report, Adams & Murphy (1967) described the US's Sustaining University Programme (SUP), which intended to increase the quantities of trained scientists and engineers, remove interdisciplinary barriers in space research and make universities more conscious of their '[…] societal responsibilities in the attainment of national goals' (Adams and Murphy, 1967, p.12). It aimed to maximise the efficiency of university research for strictly political goals. The were the government's and NASA's efforts to increase the production of advanced technologies during the Cold War. As I argue, the CVCP used the American example to further explore the UK's global position and understand how UK university research could be steered towards higher levels of employed engineers and scientists.

Interestingly, it is in meetings held with the Italian records (CVCP,1968), that the CVCP expressed preoccupation for the UK's lagging position, suggesting that a '[…] technological gap exists between Europe and North America in favour of the latter' (p.6). The number of UK *trained engineers* (71%) was superior to the US's (54%), but its quantities of industrially employed engineers it lagged the US as well as the three major western European countries – Netherlands, Belgium, and France. Dainton highlighted how this aggravated existing issues of scientific policy in Britain, arguing that universities create "centres of excellence" for scientific research which would carry out high-quality applied research to facilitate later employment in industry. However, the UK first needed a '[…] more "directed" national science policy' (CVCP, 1968, p.14) for its  funding. Further, raising the status of the hard and applied sciences in a context where the admissions and demand for the social sciences were taking over was no simple path within a government committed to the social demand principle. In a report published whilst working at the committee of scientific policy (CSP) – of which he became chair between 1970 and 1972 –  Dainton explained how admissions in the hard sciences had fallen by about 5% (from 45.9% in 1962 to 40.6 % in 1967), replaced by a doubling in the admissions for the social.  Dainton's insisted that universities needed to act more decisively in national scientific policy, to '[…] develop their own views about scientific research in universities' (CVCP,1968, p.15).

This was particularly relevant in the changing funding context between the late 1960s and early 1970s. Whilst chair of the CSP and parallelly VC of Nottingham, Dainton expressed his position (Dainton, 1971) on the matter vis-a-vis Lord Rothschild's new proposals for research funding (HMSO,1971). Dainton was concerned with Rothschild's evident intention to erode the independence of the then existing research councils (Medical, Agricultural, Environmental, Science

and Social research, the last two founded in 1965) through the "contracting principle". Here, funds normally sent to research councils would instead be channelled to new "customers" of scientific research, namely the government departments (HMSO,1971), who would "employ" and pay only for the specific research which they required, rather than unconditionally sustaining the research community. This evidently signalled that research councils themselves would become almost fully-dependent on political –and not social nor academic/scientific – demand. Even worse, councils also risked becoming "unemployed" if government decided either to fund other organizations or not to rule out the use research findings for policymaking altogether.

In a context where scientific university research was funded by UGC and the research councils and coordinated by the CSP – which divided funds received by the Department of Education and Science (DES) amongst the councils – Rothschild's suggestion also translated into possible reductions in the funding of PhDs and Masters. The CSP's role was being readjusted into a mere advisory one, shifting decisional powers over research funding and content from the academic & research community the DES. Dainton strongly opposed this *de facto* erosion of research councils' independence and argued that research could be steered through less radical adjustments, such as the establishment of a "statutory board of research councils" (Dainton, 1971) which could strengthen councils by giving them declared rights and say (Dainton, 1971). This struggle between Dainton's and Rothschild's proposals is emblematic of how the CVCP's intent to steer the hard sciences pushed it speak out, through members such as Dainton, in the attempt to retain research management powers. Dainton in fact proposed that a new advisory board would ideally be composed of a wide array of representatives, accommodating political demands without eroding the role of an intermediary group which could act in the interests of academia. This new negotiation space could allow the CVCP to more closely intervene by having one of its members in the board alongside rotating representatives from departments, a cabinet office member and the UGC chairman (Dainton, 1971). The CVCP thus tried to directly interact with political groups and win political support for an alternative path to that being proposed by Rothschild, reflecting how the committee saw itself as a political group whose actions were all but limited to the university walls. As chapter 4 will explore, this was often the case when political demands presented new constraints, and led in no marginal way to very important, bottom-up innovations for the university sector.


### 3.3.3. The Conference of the Rectors of Italian Universities

Considering that until 1964 Italian universities did not have a national Vice-Chancellors' committee, it is important to briefly introduce the committee's history to grasp its development. The CRUI's position has significantly moved away from the margins in which it originally stood upon

its establishment. Today, the CRUI is a key consultative organ to the Higher Education Minister, a political role which, further, has often been covered by ex-Vice-Chancellors (Pinto, 2009). As chapter 5 argues, this escalation of power was largely facilitated by the CRUI's work in evaluation, in turn importantly shaped by its connection with key research centres such as CENSIS. The issue, however, is that reconstructions on this specific development – and of the committee's history prior to the 1990s more broadly – are extremely scarce. One of such works is that of Sofia (1998), drawn from her role in the committee as its secretary in the 1970s and as key conduits with international political circles in the 1980s and 1990s. Corradi worked on educational policy at the UN throughout the 1960s, and her engagement has been such that she carries the name of 'Mother ERASMUS' because of her role as key founder of the ERASMUS program[58].

The first part of the CRUI's history follows a specific thread, namely the intent to acquire sufficient political influence over national university policy. Intuitive as it may sound for a committee which represents the interests of the university sector – and can by definition be considered as naturally inclined towards this goal – political influence in the sector is by no means a given. In fact, projects such as ERASMUS owe more to the committee's engagement with international networks than with national ones. The CRUI was informally established in the academic year 1948-9, in a first meeting in Milan initiated by the committee's then first (informal) president, Giuseppe Menotti de Francesco (Corradi, 1998), just as Italy was beginning its Post-War reconstruction path in which the expansion of education was a key priority. The committee met up occasionally until 1955, when the formation of CRE pushed it to take a more proactive turn (Corradi, 1998). In 1963 the CRUI wrote up its charter, articulating its aims, amongst which:

> '[…] to develop studies on universities' key issues and represent universities' needs to the political authorities […] to propose provisions, also law provisions, oriented at bettering teaching and research regulations, […] maintaining relationships with teaching staff and associations and organisations which can have interaction with the Universities' (comma 2 of the CRUI's Charter, as cited in Corradi, 1998, p.53).

Following approval of the charter, the CRUI was officially founded, with Guido Ferro – rector at the University of Padova – as its first president. The committee sought to act '[…] in tight connection with the State, which turns to it [the committee] for advice, just as it then turns to the State for legislative proposals' (Carnacini, 1976, as cited in Corradi, 1998, p. 52). Although the

---

[58] See http://www.sofiacorradi.eu/ . Although the ideation of ERASMUS dates back to the late 1960s, it was not until the mid-1970s, specifically in 1976, that the EEC approved a resolution with the principles set out a decade earlier, and it was in 1987 that the program was officially launched (Corradi, 2015).

committee's meetings often attended by ministers and prime ministers (Corradi, 1998), this did not suffice to evolve into a model which it looked up to: the CVCP.

**The CRUI looks up to the CVCP**

The CRUI admired the CVCP's ability to intervene on university issues at a sectorial level, with its proposals reaching MPs and discussed also through ad-hoc commissions when necessary. The CRUI was well aware of the roles and influence of the politically autonomous yet politically active UGC and CVCP relationship and advocated to emulate this in the Italian university sector (Corradi, 1998). Similarly, the CRUI was up to date on key British reforms such as Robbins', such that it attempted to influence the Ermini commission's articulation (law n.2314 of 1965) of how Italian universities could expand, drafted along the lines of the British binary divide (Corradi, 1998) model. Like the CVCP, the CRUI was hesitant about expansion for the effects that this could have on prestigious élite universities[59]. However, the CRUI was aware that the CVCP's ability to speak out on reforms importantly depended on the UK's exceptional configuration of university autonomy in the UK, and that it needed to achieve a similar relationship with the State if it sought to become more influential. Between the 1960s and the mid-1970s, the committee's jurists (such as Carnacini, president of the CRUI between 1972 and 1976) advanced legislative formulations which encouraged politicians to concretely define and acknowledge the CRUI's position and role in the sector, they emphasised university autonomy as both the sector's and the CRUI's founding principle (Corradi, 1998). These actions, however, did not lead to an effective political recognition of VCs' power, VCs were still considered secondary to politicians, it was politics that kept the power to administer teaching & research procedures. Tired of these national constraints, the CRUI mobilized to develop its own knowledge and networks, organising conferences with the CVCP, and engaging directly in VC circles such as CRE.

At a second conference held by the CVCP in Nottingham in 1968[60], the two committees discussed various key topics: autonomy, finance, expansion, governance. Specifically, the CRUI sought to learn from the CVCP on how to develop into a prestigious and more politically influential committee. One of the CRUI's key political goals at the time was avoiding the outright massification of higher education and, instead, achieve the establishment of a binary HE sector. Its political position was influenced by the 1968 student protests, with the *Pantera* movement

---

[59] 'The establishment of new universities faces some issues […] first, if we are talking about universities in the strict sense of the word, the number of new institutions can only be very limited, for the very nature of the élites destined to their use. Thus, expansion should generate necessities of expanding existing universities rather than creating new ones.' (copy of the presentation of Del Prete, VC of Bari, as consulted in CVCP, 1968)

[60] The first conference was in fact held two years earlier in Padua, between 7-8 October (CVCP, 1968) following Mr. Ferro's request of establishing a '[…] permanent cooperation […] for an exchange of information, experiences, ideas and the establishment of concrete forms of collaboration' (CVCP, 1968).

importantly involving also large segments of the Italian academic body. Concerned that massification could de-legitimate them and threaten the committee, Italian vice-chancellors explicitly discussed the issue of academic majorities in governing bodies with the CVCP (CVCP,1968). British governing boards were in fact more diverse than Italian ones and were for example open to lay consultancy (CVCP, 1968) and advice from the private sector. The CVCP proposed that the CRUI could consider a British initiative being developed, namely that of attempting to control staff and students by tracing their careers through the collection of information on their movements and achievements (summary of the CVCP, 1968 meeting (CVCP, 1968). Further, the CVCP advised that student numbers could be monitored through analysis of '[…] departmental costs, and the cost per student in different subject areas in relation to the size of the university, department and the nature of the work undertaken' (CVCP,1968, pp.3-4). These moments of interaction provide interesting details as to how the broader changing context was being understood by the CRUI, but, more importantly, the reference model which it sought to emulate. The developments of this network throughout the 1970s have been difficult to trace given scarce archive and academic material. The discussions between the two involved considerations on exchange programs between British and Italian universities in the hard sciences (CVCP, 1968) and generally in the construction of a more solid collaboration between the two.

### 3.4. Conclusions

This chapter has presented the OECD's research on educational governance between the post-war period and the mid-1970s to highlight the organization's role in the development of output governance in education. As pointed out throughout, the OECD's influence developed in two key ways: through the research which was produced within its walls, and through the ramification of such research and its respective networks to national contexts such as Italy and England. A first process concerns the immediate-post war research on educational planning and human capital theory . This research stream developed largely in relation to the US's own liberal agenda of manpower building and diffused to Italy through the MRP project. A second key process was the OECD's shift, at the end of the 1960s, from educational planning research to institutional management in universities. As discussed, this culminated in the work of CERI, and, outside the OECD, in that of the IIEP. These bodies' use of management science principles and of the PPBS model specifically was significantly advocated by Alexander King and Philipp Coombs, who mobilised the RAND corporation's research and networks when setting-off the two centres' research agendas. Importantly, it is in this way that a nexus between management and educational evaluation began to develop, as non-educational instruments coming from the military context were developed to suit educational decision-making issues through the frame of institutional management

in a context of a 'world educational crisis' (Coombs,1968). Further, it is in this context that managerial turn in international research on university governance which reached into university walls through universities' participation in international research projects. The case of the University of Sussex's turn to institutional management through the use of the PPBS framework shows the crucial role played by both international research, on one hand, and national-specific research, on the other. Conceptualising the development of a new managerial approach to university governance as the result of an intersection of national and international research networks allows to decipher, further, how university agents importantly mobilised emerging concepts to reach given interests. This signals the importance of historicising through an agency-centred approach. Indeed, analysing context-specific articulations of the rise of educational evaluation broadly allows to unveil the processes through which new ideas and practices were concretely mobilised. Crucially, this chapter has shown that managerial developments in educational governance began developing much earlier than the 1990s, contrarily to what most scholars assume. Iin different ways, ***both Italian and British researchers were working on educational planning and evaluation by the mid-1970s***. ***This older lineage emerges in the third process which I discuss in this chapter, namely the development of managerial framing of university governance inside vice-chancellor committees***, as the case of the CRE's launch of management seminars in 1977 shows. Combined, these insights point to a very specific historical lineage of educational evaluation practices in the university context, one which is more related to the development of a planning movement during the Cold War period than to the rise of the neoliberal political cadre at the turn of the 1980s.

# Chapter IV

## The Rise of Performance Evaluation in English Universities: 1970 – 1992

Having discussed how educational planning emerged and developed into educational evaluation in the international context, the focus now turns to the national sphere. This chapter opens the analysis of the two case studies of this study, starting with England. The British State is often considered, along with the US's, the cradle of new public management principles. In part, this is because Anglo-Saxon capitalism is seen as developing on faster tracks as the anticipation of other countries' developments. This convergence towards Anglo-Saxon capitalism is often re-proposed by NPM scholars, who suggest that the retreat of the British State was accompanied by performance reforms in central government. Although this argument is convincing at a general level – given that European States also began to reform their central governments and dismantle the public sector through privatization – it is problematic when confronted with specific historical evidence of the British State's use of management practices throughout time, of its specific NPM trajectory. To give an example, the entrance of managerial consultancies in the British State occurred in the midst of the 1960s (Saint-Martin, 2000) rather than during its "rolling back" in the 1980s. This, arguably, suggests that the turn to efficiency in government and in the governance of the public sector more broadly began earlier than the State's presumed "roll-back" in the 1980s. Coupled with the fact that it was Wilson's labour government to first advocate for the entrance of consultancies in government, this indicates that the turn to 'management' was not necessarily driven by neoliberal doctrines. These aspects reveal to the necessity of better historicising managerial governance through evaluation, as the use of management practices for policymaking may not have diverse explanations which are not directly correlated to assessing the public sector specifically.

This chapter begins by discussing how, though specific performance-oriented reforms to central government were made in the 1970s, they were rather unsuccessful and quickly rejected by Thatcher upon her election as Prime minister in 1979. Building on existing studies which have problematized the definition of management as an umbrella term (Saint-Martin, 20000), the first part of this chapter thus problematizes assumptions whereby the rise of neoliberal governance implies an active search of performance assessment practices geared at quality evaluation of the public sector. It argues that, as of the 1980s, the British State was not articulating quality in terms of performance, but rather, in terms of the State's savings in public expenditure. These insights allow to contextualise university-level developments. The second part of the chapter in fact explores if and how the turn to efficiency in central government played out in the university policies between the 1970s and 1981, the latter being the year of the infamous "selectivity cuts" which defunded the

university sector in an unprecedented way. A second assumption of NPM scholars is thus historically problematised, namely that the rise of evaluation in the university sector be explained as an inevitable adaptation to the State's managerial policies. As I discuss, whilst the "Evaluative State" (Neave, 1988, 1998) was still in its infancy at the turn of the 1980s, VCs already had an array of managerial research at hand, as well as an in-depth understanding of the application of the PPBS model in higher education. This difference in managerial expertise, I suggest, is partially accounted for by the different networks of the CVCP to those of national governments. Further evidence supporting this argument is that the State's articulation of efficiency was driven by the priority of implementing expenditure-cuts in the public sector, whilst the CVCP's interests for using PPBS principles concerned a proper re-articulation of internal university governance. These interests were significantly shaped by knowledge on the way that managerial dynamics could help to influence decision-making processes in favour of top levels and enhance their ability to control resources and their internal distribution. Budgeting and performance assessment were in fact proposed by the committee, for they could assist the construction of such new power configurations. Section 4.2. explores the encounter of State and university approaches to management practices in the first RAE, wanted by Thatcher in 1986 to initiate selective research funding. If on one hand it was the result of a political thrust, the 1986 RAE was significantly influenced by the CVCP's own indicators, devised in collaboration with the then funding council, the UGC. Section 4.3. turns to the 1990s, a period in which the abolishment of the traditional binary line also influenced the development of research evaluation criteria. As is argued, the establishment of a new funding council – the HEFCE, (Higher Education Funding Council for England)– represents the State's attempt to centralize its power over the university sector , rather than to limit it. The state intervened by tracing the boundaries of action for evaluation, in this way positioning itself vis-à-vis evaluation arrangements which had developed *outside state walls* during the mide-1970s and early 1980s. This argument not only poses significant pressure on suggestions that evaluation criteria have an intrinsically competitive logic, but also enlightens the understanding as to how the CVCP reacted to threats to its traditionally autonomous role and its exercise of influence over universities. Now struggling to directly impact evaluation criteria as it had done in the mid-1980s, the CVCP tried to develop new strategies for evaluation. Keen to preserve the elitism of prestigious research universities, in the post-binary period the committee also argued for a differentiation between research-intensive universities, reproposing those binary divisions between vocational and academic higher education which the State sought to merge.


## 4.1. The State and Efficiency

Management consultancy and managerial practices were piloted in the British State since the mid-1960s. PPBS was experimented in the British PA during the 1970s but was soon side-lined as a result of political resistances to it and technical complexities to its implementation. Still, the idea that management could assist politics still developed, with different traits during Wilson's and Thatcher's governments. Each attempt to use management practices left a trace which successive governments either built on or discarded. Traces of a PPBS-inspired framework are most evident in the Program Analysis and Review (PAR) framework set-off during Heath's government (1970-1974). This reform movement attempted to use performance assessment to push civil servant staff to align ministerial policy priorities to political expenditure goals. However, significant failures revealed the intrinsically political struggles which the use of PPBS-inspired practices fuelled. Thatcher rejected this framework, and her leadership rearticulated the scope of management in central government from one of performance assessment to one of cost-benefit analysis. As is shown, this formed the core of her "value-for-money" policies in the public sector, including the university one.

### 4.1.1. Reforms in central government

NPM literatures argue that public sectors gradually adapted to the State's reformed administrative apparatus (Ferlie, Hartley, & Martin, 2003), and in so doing they indicate a specific, top-down directionality in the way that NPM developed, namely from the State down to public sectors. As the State's policymaking practices became more efficient, so did it reform its policies, with the result that public sector reforms were increasingly quality-oriented. In this view, the 1980s are the conventional time-period of reference. However, a first important wave of management-oriented reforms for central government sparked off earlier than the rise of Thatcher's government, which came to power following an initial spread of consultancies and management ideas in the Public administration. Whilst it is now common to consider politicians contracting consultancies for policy guidance, it is perhaps less common to historically consider how this came about and what diverse political motives have tied governments to management practices and consultancies. Different outcomes of reforms carried out during the "consultancy movement" were in fact highly dependent on internal resistance or alignment to the new practices which were being proposed. As Saint-Martin notes in his elaboration on the consultancy movement in Britain (2000), analysing the relative success of management practices cannot in fact depend exclusively on the assumption, – evident in supporters of what he denotes the "consultocracy thesis" (Hood & Jackson, 1991) – that '[...]the State is [...] an empty shell through which [...] consulting interests can easily press their views concerning public management issues and get their preferences almost automatically translated into policies' (Saint-Martin, 2000, p.24). The rise and development of a managerial

imprint in Whitehall in fact demarcates a period of different reforms to the functioning and procedures of central government and many obstacles to their implementation.

Push for such reform was initially grounded in an attack of the civil service's elitist Oxbridge background. In an open critique to the conservatives who had governed up until the mid-1960s, Wilson's report "*Labour and the scientific revolution*"(1963) strongly attacked the conservative party's inability at modernizing policy making by directing it towards social and scientific relevant fields in a period of intense economic development and modernization. He ascribed this tendency to its elitism and capitalist inclination, advocating that a "science-based" government would improve economic planning and industrial development in Britain (Wilson, 1963). In line with arguments for renovation[61], Wilson pushed for more appreciation of the scientific and industrial professions within government and set up the "Fulton committee"[62] in 1966. This group was charged with formulating a new operation and structure for the bureau. It quickly tried to replace the civil service's "generalist" skills (Fry, 1969, 1990; Saint-Martin, 2000). In the attempt to "specialise" the civil service, the Fulton committee invited consultancies to the British State (Fulton, 1968) to undertake a review of British Administration through management consultants' framework, such as those of the AIC (Saint-Martin, 1998).

This consultancy was hired by the *management consultancy group*, the operative arm of the Fulton committee, determined that AIC's external judgments would bring a set of guidelines for reform. In this sense, consultants were brought in to confirm, 'objectively', the claims and necessities of politics. The role of external figures would allow to demonstrate that Wilson's attacks and of Fulton's call for transformation. stood on legitimate grounds which industrial experts shared. One of the Fulton committees' proposals was that departments' work be balanced out by consultancy advice, amongst others (Fulton, 1968). To implement them, Wilson decided to set up the Civil Service Department (CSD), which, having expressed its distrust of the Treasury, overtook part of its functions, such as managing civil service staff (Saint Martin, 2000) through recruitments. In attempting to recruit new civil servants through 'outside specialists', the CSD was composed of a majority of members coming from private industry (Saint-Martin, 2000). The developments at the CSD upset the m, which responded by trying to ensure that new departmental recruitments of consultancies first pass passed first through its approval (Saint-Martin, 2000). The Treasury thus opened its doors to external consultancies' advice, at one point inviting representatives of over 20 consultancies (Saint-Martin, 2000). In the attempt to further solidify a new class of civil servants, the CSD set up a training college  for developing managerial expertise. However, the CSD was

---

[61] In Wilson's words 'For the commanding heights of British industry to be controlled today by men whose only claim is their aristocratic connections or the power of inherited wealth or speculative finance is as irrelevant to the twentieth century as would be the continued purchase of commissions in the armed forces by lordly amateurs' (Labour Party, 1963, p.139)

[62] Interestingly, Fulton had been Vice-Chancellor of the University of Sussex between 1959 and 1967

relatively short-lived (Saint Martin, 2000) for soon after its establishment in the late 1960s the Wilson government lost power to Heath's, with the result that the political force leading the CSD was radically changing.

After Wilson, the political use of management practices shifted from that of controlling the recruitment of civil service staff to evaluating its policymaking competences. To this end, Heath's government built on the grounds set by Wilson. Heath's agenda of a "reorganization of government" (White paper, 1970) was in fact prepared whilst in opposition by Schreiber and Howell (Radcliffe, 1991), two researchers of the conservative party. They sought to develop a framework of '[…]systematic analysis from the centre which would bring together material from the departments to develop a *strategic analysis of government policy*' (Radcliffe, 1991, p.69). Rather than attacking the civil service's elitist background, Heath's reform narrative was based on an attack on its competences, such that the "reorganization of government" was indeed largely based on the use of new policymaking practices. In broad terms, Heath saw efficiency as both a rationalising exercise meant to simplify procedures and reduce them;  and as the assessment civil service staff's work specifically (Gray & Jenkins, 2002). Heath in this way developed the scope for management within government in two relevant aspects. Firstly, he reduced the number of political ministries to encourage a simplification of the bureaucracy. Secondly, he  introduced a template for assessing departmental policies through Program Analysis and Review (PAR from now on). Secondly, he tried to bring consultants closer to ministers (Saint-Martin, 2000) and have them assess their operation.

Determined to merge ministries through the "unification of powers," Heath tried to modify the processes through which departments attracted public resources for implementing political agendas, starting from the number of ministries and their departments. One of the crucial elements of his reform wave was the necessity to overcome a key effect of departmentalism, namely the weak coordination between departments in policymaking. Stressing the unification of functions served to create fewer but larger departments in ministries, merging the work of different civil service staff under a single title. This centralization and streamlining, it was believed, would ameliorate the efficiency of departments' work, and, somehow implicitly, align departments to broader policy goals, formulated in budgetary, expenditure-constraint terms. Given the bureaucracy's inter-ministerial conflict for resources (Saint-Martin, 2000), Heath's team thought that enlarging ministries in this way would reduce political conflict between them.

Next came the assessment of how staff concretely constructed policy programs. In 1971, the Program and Analysis Review framework was therefore introduced to make civil servants within the new and larger ministers-now responsible for larger sums of public resources- more *directly responsible* for the receipt and expenditure of public money and more inclined to coordinate their

activities rather than compete. This was ideated through the introduction of a standardized framework that staff would use as reference for reporting their activities, very much in a step-by-step manner. This was meant to set off practices of transparency and objectivity into policy making. PAR was thus intended to remedy the absence of an '[…] analytical instrument by which the impact and effectiveness of policies [particularly in relation to public expenditure] could be assessed and monitored' (Gray & Jenkins, 2002, p.131). The Public Analysis Review Committee (PARC) which largely developed PAR worked from the starting point that ministers' political objectives tended to be unrealistic when faced with the material and limited resources available. As a corollary, it was believed that - '[…] socially scientific diagnostic tool(s)' (Lewis, 2011, p.774) could limit this tendency , particularly in social policymaking, where it is most difficult to calculate potential public costs and benefits . "Relevance" became a new buzzword (Gray & Jenkins, 1982) demanding that policies be as coherent as possible with political priorities, be they articulated as general political achievements – "access to social services" – or as explicit expenditure goals – "expenditure at % on the investment of computers in universities"–.

A key reason for which PAR aimed to closely monitor policymaking in detail was to align it to the political agenda. PAR and its respective committees were therefore crucially concerned with implementing activity reporting, a step which could leave politicians the freedoms to elaborate agendas without being strictly tied to departmental concerns. PAR could limit departmental claims for resources at early stages, for example. Similarly to how PPBS introduced new decision-making templates, PAR '[…]was concerned both with the consideration of objectives and with an examination of the scope and need for changes of policy' (Gray & Jenkins, 1982, p. 435). Its structure indeed mimicked PPBS's to some degree (Lewis, 2011), this was most wanted by Lord Rothschild, who headed the Central Policy Review Staff (CPRS) –founded in 1971– and advocated for its widespread use. An example of PPBS's use is a CPRS report published in 1972, "Resource allocation: Social Affairs", which used the notion of resource constraint as a starting point for staff reporting activities. This approach would'[…] help Ministers to judge (a little more) rationally whether the allocation of resources to a given output is greater or less than it should be' (Lewis, 2011 p. 774), suggesting that as a result of reporting MPs could review their short and long-term strategies. Whereas departmental resources were conventionally calculated on incremental basis and thus taking increasing resource needs as a given, the use of resource constraint questioned those very needs. Departmental programs now, had to state intended goals, the amount of resources necessary, and the possible alternative means of achieving objectives, justifying each step as well as why the program eventually chosen was the best option among alternative scenarios(Saint-Martin, 2000). This mechanism of departmental review recalls the way that program budgeting emphasized an attentive quantification and consideration of multiple strategies as a decision-making method.

The CPRS committee used PPBS principles to keep hold of the resource reins through its judgments of programs' relevance (Gray & Jenkins, 1982). In this way, it took on an evaluator role of departmental programs for a period of time, providing departments with templates known as "skeleton charts" (Lewis, 2011). At the same time, the CPRS published few reports explaining the developments and outcomes of PAR, an aspect which did not help to remedy its gradual decline (Lewis, 2011). As some have argued, the strong division within the CPRS itself between those who supported and those who objected this new rationalist drive to policy design and analysis (Lewis, 2011) was strong, such that it eventually lost momentum four years following its establishment (Gray & Jenkins, 1982).

This happened for other two key reasons. Firstly, although the CPRS had an important evaluative function geared at monitoring departmental expenditure, it was not a fully autonomous and authoritative committee. The Treasury, also a member of the CPRS, often exerted its strong bargaining power through bilateral negotiations with departments over policy priorities and their expenditure levels, in this way circumventing the CPRS's role in the PAR framework (Grey & Jenkins,1982). A second limitation was ministerial political objection on the grounds that the attempts to replace traditional policymaking procedures through the complex PAR methods and rules was seen as neither practical and efficient nor desirable. Indeed, the framework was seen as a means which depoliticised their political functions, turning them into mere executors. Though ministers could decide program goals, they then had little leverage over the design of programs, which needed to be developed 'rationally.' Figures conventionally attached to the managerial turn, such as Margaret Thatcher –at the time Secretary of State for Education– and Keith Joseph – who later took on Thatcher's role during her government in 1981– actually strongly refused the very principles of PAR's *objectivist turn* in politics (Lewis, 2011). They claimed that PAR constrained political decisions in exchange for an efficiency that was not advantageous to them, for the policy assessments could also result in less state resources for ministers. Wildavsky's observation that '[…] theory which contains criteria for determining what ought to be in the budget is nothing less than a theory stating what the government ought to do' (Wildavsky, 1961, p. 184) captures these political reactions to PAR. One of Thatcher's criticisms reaffirmed the primacy of politics over generic rationalist frameworks such as PAR, considered too distant from the complexities of political reality. In her words: '[….] the theory that there should be an overall strategy in social policies […] requires departmental objectives to be overridden by those of the overall strategy if necessary' (Thatcher as cited in (Lewis, 2011, p.777). The PAR system was once more advocated in Wilson's later government as an instrument of budget control during the oil crisis in 1973, yet it was consistently attacked until 1979 and eventually abandoned until its abolishment shortly after Thatcher's election (Lewis, 2011). As has been convincingly argued, (Gray & Jenkins, 1982;

Greenwood et al., 1977; Lewis, 2011) the primary reason for the crumble of the PAR system and in particular its functioning within the CPRS is indeed ascribable to the fact that political networks and negotiations play an important role in shaping the attainment and use of state resources, and thus the political clash of interests that PAR brought about by seeking to reduce these dynamics was a significant obstacle. Greenwood, Hinings and Ranson (1976) show that political struggle is an inherent feature of the processes of articulation and use of the state budget. The budget is in fact divided among departments according to the political priorities of the political party in power. The authors suggest that it is thus unrealistic to assume that detailed analyses of budget allocations could somehow have watered down that political discretion central to the process of departmental planning and resource allocation. Wildavsky (1964) similarly emphasizes how much changes to the budgetary process necessarily imply changes to the distribution of powers within the political context, as such reform touched on both the sum of funds that departments receive and organization of responsibilities. The PAR system thus did not achieve enduring transformations in policymaking processes of the 1970s (Grey & Jenkins, 1982, Lewis, 2011, Saint-Martin, 2000, Radcliffe, 1991) and neither, as will be discussed, in those of the 1980s. Along PAR, Heath also formed a group of consultants advising directly to the ministers, defined as the "businessmen's team". Members came from consultancy agencies such as PA Consulting and Booz-Allen (Saint-Martin, 2000) and were entitled to develop ministers' policy approach and to an extent also policy content (Saint-Martin, 2000). The team in fact had direct access to decision-making centres, for it was recruited inside the party's policymaking machinery.

Differently to PAR, this framework persisted in Thatcher's government, which borrowed this model when the "efficiency unit" was established. Thatcher brought advisers even closer to politics, directly in the prime minister's office (Saint-Martin, 2000). Although it might appear counterintuitive given her objection to rationalist procedures in politics, this move actually strengthened her control over consultancy and simultaneously took consultants out of departmental decision-making processes. In this way, Thatcher was able to eliminate that dynamic of PAR whereby new procedures limited departments' autonomous functioning and downplayed ministerial roles. Within this re-empowerment strategy, ministers were now freed from the constraints of PAR procedures. Thatcher thus still refused consultancies' decision-making procedures for policymaking yet embraced their role as advisers on policy agendas or on understanding societal and economic dynamics. This brought management consultants in direct contact with the State agenda. Managerial advice spread also through the "Number 10 policy unit"[63], considered '[…] among the key organizational channels through which management consultants could make their voices heard within the inner circles of policy making' (Saint-Martin, 2000, p.97). This unit embodied the

---

[63] Curiously introduced in 1974 by Labour (Saint-Martin, 2000).

"presidentialist turn": closeness to the Prime Minister meant that if consultants had clear ideas on the State's goals, then they could 'politicise' civil service staff and act as **consensus-makers** (Weiss, 2008; Saint-Martin, 2000). With Thatcher, consultants acted in the prime minister's interests, not for objective notions of bureaucratic efficiency. Ministers could still review departmental programs and practices (Saint-Martin, 2000), but this was left to their initiative rather than being imposed onto them.

In this new context, it was the Financial Management Unit (FMU) established in 1982 – '[…] the economic liberal gospel as applied to the Civil Service' (Fry, 1988, p.5) – which most reflects Thatcher's perspective on what efficient policymaking meant: reducing public expenditure. The FMU hired consultants specifically for advice on expenditure cuts, starting from within central government where these starkly opposed by civil servants (Fry, 1988). The FMU's thrust was importantly supported by Rayner's "Scrutiny program," later led by Ibbs. Through Rayner and Ibbs, the FMU stressed efficiency in terms of cutting costs, within departments and policy programs, financial management was seen as useful for the State's financial sustainability. It was meant to avoid, using Rayner's own words, that costs of administering policies continue to be seen by civil servants as '[…] the candle-ends of public expenditure (Rayner, 1984, p.2 as cited in Fry, 1988). The consultants hired at the FMU were from the Consultancy Association, which had begun to work on governmental advising during consultancies' broader turn to exploiting the '[…] commercial potential of government and other parts of the public sector' (Saint-Martin, 2000,p. 101). Rayner – CEO of Marks and Spencer – was appointed by Thatcher to advise on financial efficiency and strengthened a "value for money approach" to policymaking, which conceptualised cost-cutting as a synonym of efficiency (Fry, 1988). The value for money approach was crucially different to Heath's performance assessment approach. The former stressed efficiency to push cost-cutting, whilst the latter insisted on relevance of goals and resource use assessments. Indeed, Thatcher did as much as possible to dismantle the previous logics[64] in favour of value for money.

The efficiency unit tried to diffuse cost-considerations both in the running government and in policy expenditure plans (Fry, 1988) with the aim of having all departments work towards the most cost-efficient (and expenditure-saving) policies (see Ibbs Report, 1988) – bringing the Treasury close to Thatcher and actively interested in the FMU. Departments were asked to act in a more cost-aware manner by carrying out cost-benefit analysis, and the unit also established budgetary control systems (Fry, 1988) to ensure that departments engaged in '[…] the best use of their resources, a critical scrutiny of output and value for money' (Fry, 1988). A further feature of Thatcher's cost-oriented understanding of efficiency is the fact that these management practices

---

[64] See a recent parliament report discussing key changes of this period in "Changes in the public service since 1967" at https://publications.parliament.uk/pa/ld199798/ldselect/ldpubsrv/055/psrep07.htm

were further encouraged when the Prime Minister's electoral promises of reducing the costs of the State machinery began to appear unfulfilled, and thus used as a means to remedy the situation (see Fry, 1988). Thatcher at first thus articulated management in financial terms, rather than in quality or performance terms as an instrument for supervising civil servants. Importantly, this explains why public sector efficiency and evaluation-oriented policies were not made through preliminary assessments of the quality of service-provision and without the recourse to specific models for understanding quality. Instead, the priority was to identify and cut superfluous expenditure from public sectors, pushing them to work with less. Policy formulation was thus centered around the primacy of cost savings and much less around finding a benchmark for guiding expenditure. The legitimation of cost-attentive strands of managers within the State actually postponed complex qualitative issues–and the qualitative implications of cost-cutting in the public sector – to a second moment. Evidence of this cost-focus is the significant wave of privatizations and redundancies set-off by Thatcher during those years[65] (Fry, 1988). Ibbs's 'Next Steps' initiative (1988) in fact followed this rationale and stimulated contractionary policies of value for money in the public sector (Shattock, 1987). In the university sector, the policies of the 1970s inherited this cost-cutting approach which made very scarce use of specific quality considerations. The DES did inherit the performance-assessment guidelines introduced by Heath, and its policies instead made use of existing instruments at its disposal, such as student numbers, to push universities into using fewer resources. The next section discusses university policies with a long-term perspective to show the *absence* of a managerial state in action and discuss the consequences of the actions of a neoliberal state in the sector.

### 4.1.2. Reforms to higher education: managerial or expenditure-led?

During the 1970s, governments frequently tried to gain control over the exceptionally autonomous university sector. Traditional funding structures were either modified or removed, such that by the 1980s the sector was more vulnerable to political interventions than a decade earlier (Shattock, 2012). By 1980, it no longer stood on stable State funding such as quinquennial grants and those for overseas students. These were crucial grounds which granted universities the ability to develop through public support. The 1970s changed much of this, they were marked by a series of incoherent political decisions which created widespread uncertainties on the sector's future (Kogan & Kogan, 1983). The DES's growing interference into the UGC's operation (Shattock, 2012, Kogan & Kogan, 1983) further deepened these uncertainties, because it threated the integrity of the universities' key reference body. The reasons for such drastic changes during the 1970s can be found in political attempts to regain financial control over the university sector in order to then de-

---

[65]Such as British Steel where the number of employees fell by 56 % between 1979 and 1982/83 (Abromeit, 1986) and the number of non-industrial civil servants was cut by 14% between 1979 and 1986 (Fry, 1988, p.4)

finance it, rather than to devolve governance powers further. To be precise, this became more evident in the 1980s. Indeed, the exceptionality of the English university sector lied in the fact that universities, in freely allocating their State grants internally, could rely on political resources yet were to a large extent shielded from political or other interferences unless they actively sought it (Silver, 2003). The only concrete criterion that governments of the 1970s and 1980s had at their disposal for steering universities was student numbers, the factor that governments most used to influence expenditure levels and instrumentalise the sector to political priorities. As I argue more specifically, governments had no concrete *quality* criterion cutting expenditure, as the build-up to the 1981 cuts and the cuts themselves show. To provoke thought on this argument, it is worth asking the question of why, if governments did possess such managerial instruments, these were not put into practice from the onset, as a managerial-evaluative state would be assumed to do. Following this argument, the section defines the policies of the 1970s all the way through to the mid-1980s as *expenditure-led*. This has significant repercussions for how the State's quality evaluation of the university sector can be analysed from a scholarly perspective. Driven by successive efforts to reduce public spending (Kogan & Kogan, 1983), governments intervened on universities' unit costs and student numbers. Efficiency was treated as a synonym of quality, largely conceptualised as the sector's ability to continue functioning with less funding at each round. Indeed, it is perhaps not a coincidence that politics often intervened on its own policies to remedy the discrepancies and chaos which these caused (see Pritchard, 1994; Giovinazzi, 2019). Similarly to developments in central government, the evaluation of universities concerned neoliberal education ministers (such as Keith Joseph) only for ensuring cost-efficiency. Policymakers scarcely elaborated what research and teaching quality meant for them. As I argue, **this explains why the first concrete articulations of research and teaching quality** – and respective methods to achieve it – **developed within university walls**. It was only following university-level developments in evaluation that State and State-devised funding councils (the UFC and the HEFCE) developed concrete guidelines on quality assessment. Meanwhile, university groups took clear positions on internal management, efficiency, and research performance. To fully capture the significance of these institutional-level transformations it is crucial to consider the defunding context of the 1970s and explore the consequences that expenditure-led policies had on universities. The policies were influenced by two key developments: firstly, the oil shock of 1973, which justified further cuts; secondly, the political struggle between the UGC and the DES over who should have the power to distribute state resources in the university sector.

The 1973 shock represents the beginning of a long period in which Robbins' optimistic expansionary spirit, expressed in his radical reforms for improving access to higher education through increases in enrolments and staff recruitments (Robbins, 1963), declined, leading to

growing disillusionment with government policies. In 1962, for example, minister of education David Eccles had increased expenditure on education to over £1,000 million (Kogan, 2018), and later the school leaving age was raised to 16. The optimism of the 1960s had been substantiated by the fact that politics was helping higher education to develop and to improve attainment levels, largely as a response to the social reconstruction efforts of the post-World War period. The British State had positively expanded its role in education[66], and as numerous new polytechnics ("technological universities") and universities ("plate-glass universities") were founded, they sustained the growth in student enrolments. Successive labour governments insisted on education as an indisputable good, yet governing education through the social demand principle turned out to be a more complex task than governments had expected. In the 1980s, 60% of 16-year-olds were still leaving education after compulsory schooling (Neave, 1985). Compared to other European countries, England's structurally higher retention rates[67] were left unaddressed by the funding policies of the 1970s and early 1980s. Between 1963 and 1979 –previously to the infamous 1981 cuts– governments cut, *on average,* £100 million from HE (Shattock, 1982). By 1981, the university sector had witnessed almost a 10% reduction in funding per student (Kogan & Kogan, 1983, p.26). These cuts varied, yet they generally reduced the costs of HE. ***The decade of the 1970s therefore reversed the positive State-higher education of the 1960s***, when the sector lived '[…] the most secure planning and funding framework at any time in its history' (Shattock, 2012, p.123).

The UGC – the universities' "shield" against political instrumentalization (Shattock, 2012) – faced increasing obstacles in defending universities against these cuts (Shattock, 2012; Shattock & Berdahl 1982). The efficiency turn in higher education was initiated by the DES in 1970, when it tried to override the UGC's role as a "buffer" between the State and the universities (Berdahl & Shattock, 1984; Eustace, 1982; Palfreyman & Tapper, 2014). Since 1918, this mostly consisted to a limited function of distributing State funds to universities following a simple student number policy (more students, more funds). This formula came under pressure at the turn of the 1970s, when the DES began criticising universities' excessive freedom in internally allocating State funds. The efficient use of State money was raised as a priority which also justify the DES's increasing interference on the UGC's formula and the broader scope for its existence, given that the DES claimed that it could do this itself (Berdahl & Shattock, 1984). Thatcher's 1972 White paper "A framework for Expansion" is an interesting into how the DES' put the student number formula

---

[66] Silver notes that 'By 1946, in the UK, over 50% of university resources were parliamentary grants. In 1952 State grants were six times greater than in 1938, radically different figures compared to the pre-war period when universities funded themselves primarily through the financing of lay members, industrialists, and philanthropies' (2005).

[67] Neave suggests (1985) that this was largely because the shift to massification in secondary schooling was not sufficiently radical to reform the elitist curricula, with obvious repercussions on the ability to continue into higher education.

under strain by making cost-considerations. Thatcher argued that universities could achieve efficiency gains if student numbers increased in the 1972-77 quinquennium and suggested that this growth would be matched by grant provisions. However, this projection was based on a specific type of growth, where it was more desirable that "cheaper" arts-based disciplines grow proportionally more than science-based ones. Indeed, this would have been less costly (Sizer, 1987). Accordingly, efficiency gains could be achieved through tighter student-staff ratios. This type of increase in student numbers would have allowed unit costs (i.e., costs per student) to fall, one of Thatcher's explicit objectives: '[…] if expansion is to continue as indicated in the decade ahead, unit costs cannot be allowed to go on rising and scope must be found for economies of scale.' (White paper 1972, §125). This attempt to introduce cost efficiency principles to the conventional student number formula, made the expansionist projections unrealistic. The projected and then actual rising student numbers signalled that State grants would continue to grow, something which universities crucially depended on. How then could the efficiency-oriented growth advocated play out, and the type of growth influenced? It was not, and in fact rising student enrolments were not matched by funding. Although the Treasury had exerted expenditure containment pressures on the DES (Shattock, 2012), the latter did not attempt to sustain growth as it had announced. By the mid-1970s, universities' positive dependency on State funding since the post-war period turned into a negative one. When the crisis hit, Thatcher did not modify the White paper's exaggerated growth projections[68] nor did she engage in negotiations with the Treasury to provide universities with financial support for stimulating growth in student numbers and facing the impacts of inflation[69] on salaries and university costs. These were, however, provisions which both the UGC and the universities expected to receive, hoping to live the stagflation period within a more financially protected environment. When inflationary pressures reached 10% in 1973 (Shattock, 2012), the government went in the opposite direction, halving a planned supplementary expenditure for the quinquennium which was meant to cover for inflation(Shattock, 2012). As the UGC reported,

> '[…] the additional sums thus made available fell very substantially short of the increase in salaries, wages and prices registered during the year […] steps reported to the committee included restrictions (in some cases an absolute prohibition) in the filling of posts falling vacant among teaching and non-teaching staff' (UGC, 1975,§18).

---

[68] According to Sizer ( 1989) government was clumsy in presuming that the age participation rate would increase from 15% to 22%, and thus the projections were exaggerated even without the then existing context of financial crisis.
[69] Oil prices doubled while wholesale prices increased by 28%.

This situation was accompanied by the DES's failure at safeguarding the quinquennial system, whose erosion was promptly considered and articulated in the 1973 Public Expenditure White paper (Shattock,2012). The abolishment was formalised in 1976 (Neave, 1985), and government modified its funding methodology from that of *volume planning* to one of *cash limits* (Shattock 2012). This meant that universities were now asked to operate with a funding ceiling, '[…] institutions dependent on public money could no longer expect protection from the State if things went wrong' (Shattock, 2012, p.124), nor hope in salary increases (Sizer, 1987). The abolishment of quinquennial funding gave the State -and the DES- major leverage over the sector, as grant sums could now be modified frequently through student numbers without having to adhere to pre-set quinquennial plans. The government's tightening of the resource belt was likely welcomed as an unexpected yet effective way for pushing universities into cost-cutting. The DES in fact seconded the abolishment of quinquennial funding, and began to replace the concern on numbers with a concern on costs (Berdahl & Shattock, 1984). Following 1976 student number projections –and grant provisions – were consistently downsized until the 1981 cuts. These constantly declining projections hit universities more hardly than they would have in a quinquennial funding context, now universities had to readjust their (declining) allocations on a yearly basis.

In universities, this period was marked by increasingly tailored attempts to plan activities according to the changing context. Still free to internally allocate resources, universities had to think about yearly strategies rather than thinking through quinquennial grants and allocation plans. This led many HEIs to plan on how '[…] to concentrate on short term economy campaign as a substitute' (Shattock 1982, p.199). Some established one-off review bodies (Shattock, 1982) which devised possible strategies for internally managing the cuts. Universities began in this way to try and plan activities with the assumption that current and future costs would need to fall. This burden had significant weight for all universities which were expanding, especially for very young universities which had just began reaching the student enrolment targets agreed on an expansionary logic with previous governments upon their founding (to give an example, Sussex university had to reach a 3,000 student enrolment target (Lockwood, 1981).This context reflected the political instrumentalisation of use of student numbers for the broader goal of cutting State expenditure. Indeed,

> '[…]the system of competitive tendering by individual establishments to provide student places is an especially interesting example of *negotiated autonomy in action* [...] so long as the State underwrote financial requirements of higher education [...] and so long as it was prepared to recognise intermediary bodies as standing at the limits of its own intervention, the State itself guaranteed academic autonomy [it] *removes the burden of responsibility from*

*the State and places it onto the individual institution itself'* (Neave 1990, pp.113-114, emphasis added).

The State could claim to be non-interfering with universities' internal matters as it dismantled the sector's traditional axis. However, actions were clearly impacting universities' ability to self-manage resources, because there were less of them, and the means of doing so, namely in a more precarious way. The safeguarding of autonomy was to an extent more rhetorical than practical (Pritchard 1994) because expenditure control through student numbers appeared as a centralizing and allocative device. This situation symbolizes the controversies that emerge when defining the retreating neoliberal state as a synonym for its reduced interference. In this case, as the State stepped away from its financial responsibilities in higher education, its use of student number targets effectively turned it into a more *dirigiste* agent: demanding targets whilst doing little to sustain their achievement, starving the beast to centralize its previous power.

The tensions created by these manipulations augmented in 1979 upon the election of Thatcher, who openly attacked the sector (Anderson, 1986) and announced the political intention to reduce expenditure as in other public sectors. This began with the immediate abolishment of yet another element which for a while had proved to be the only remaining grant certainty: the provision of grants for overseas students. Its removal was particularly advocated by Mark Carlisle, secretary of State for education (1979-1981).The Conservative government eventually reduced Labour's previous grant projections for 1979-80 by £9 million, and its choice to increase overseas student fees by 20% ended up reducing grants by a further £5.5million (Sizer, 1989). Although the Labour government had set student number targets for home students, in fact, it had left overseas targets undefined, pushing universities to try and attract the latter in order to receive overseas grants and make up for the losses incurred by declining numbers of home student numbers. (Kogan & Kogan, 1983). Scholarly research of the 1980s extensively elaborated on the contradictory content and modalities of the Conservative policies in HE (Kogan & Kogan, 1983; Shattock, 1982; Sizer, 1989). So far, the new funding arrangements introduced in the mid-1970s showed a strong priority to achieve a rapid contraction of the sector, treating it no differently to other public sectors (Shattock, 2012). This neoliberal spirit arguably left space to a managerial evaluative one at the turn of the 1980s, specifically in the 1981 selectivity cuts. However, the build-up to this point and the cuts themselves appears more in continuity with the 1970s than is often argued. Similarly to the 1970s, politics continued to prioritize (an incoherent use of) student numbers to steer the university sector into efficiency. There were little political manifestations of pre-meditated ways for implementing a 'quality framework' for higher education. Performance assessment and quality considerations were indeed not present in the 1981 cuts, concerned primarily with dismantling the State's financial 'obligations' to the higher education sector.

**The 1981 cuts**

The 1981 selectivity cuts were a first attempt, through DES's increasing pressures on the UGC, to politically implement conditional funding. They were preceded by a period of communications (mostly circular letters) in which the UGC prepared universities to the sector's contraction. However, these communications were marked contradicting messages which added to the complexities caused by the policies of the 1970s. Tensions between the State and HE intensified throughout the 1980s, as government continued to determinately advance an efficiency mantra whilst universities tried to come to grips with the implications. In October 1979, a circular letter Stated future hypothetical funding scenarios, opening up the series of discussion meetings between January 1980 and June 1980 (Kogan & Kogan, 1983; Sizer, 1988). The final cuts were announced in circular letter in May (circular letter 8/81), and the final grant letters communicating universities' allocations in July 1981. Considered together and in some detail, these three key moments sent conflicting messages, making government policy appear confusing and in certain instances improbable – to the point that at times universities didn't even react to them (Sizer, 1989). Afterall, universities could only partially influence a policy that hadn't been sufficiently discussed between the DES and the UGC in the first place (Sizer, 1989). The communications represent a governance failure in implanting *quality-oriented selective cuts* in two main respects. Firstly, they distorted government policy to the point that they created further uncertainty within universities. Secondly, they evidently showed that the UGC had no detailed quality selectivity criteria to advance, making the cuts an outright attack to the university sector rather than quality-justified decision. These points are discussed below.

Between 1979 and the spring of 1981, universities had severe difficulty in envisioning a possible future scenario. Firstly, confusion lie in the fact that the consultations with universities only included a restricted, prestigious part of their staff, Vice-chancellors. This meant that '[…]many senior academics were hardly aware that it [process of negotiations] was happening' (Kogan & Kogan, 1983, p.45), and cuts were to an extent un-consensually discussed mostly by Vice-Chancellors and the UGC. Secondly, the circulars created a 'false sense of security'(Sizer, 1989) for they downplayed the real dimension of future cuts. In fact, although the cuts eventually reduced university income between 11 and 15% in the years from 1980-81 and 1983-84, (CVCP 1981 as cited in Kogan & Kogan,1983) the circular which presented three possible funding scenarios to universities in October 1979 did not project reductions in grants going beyond a 5% figure (Sizer, 1989). As Sizer has shown (Sizer, 1989) even the 5% figure was defined in broad terms, as '[…] a position in which the financial situation could be worse, for some or for all universities' (circular letter 16/79, 15th October 1979) . It was only in 1982 (following the final

grant letter) that the UGC stated the real dimensions of grant reductions lying behind the three hypotheses which had previously been communicated to universities (Sizer, 1989)[70]. Out of the three projections which it had ambiguously stated, the UGC ended up implementing a fourth and much more drastic policy. Furthermore, a growth hypothesis was also included, confusing universities by asking them to plan for a scenario which government clearly had no intention of implementing. The same circular letter in fact indicated that total recurrent grants would remain at the 1979-80 levels – known as the "level funding" policy– for a few years to come, hence the unrealistic proposition of eventual increases. Furthermore, its determination to cut future grants was explicated in the announcement that overseas students would directly cover fees. Amongst British administrators, for example, this backup option created the shared feeling that '[…] severe cuts would *not be made* in higher education' (Shattock & Rigby 1983, p.22, emphasis added). Universities were thus being asked to plan for underscored cutback scenarios and unrealistic growth ones. The "dialogue meetings" held between January 1979 and June 1980 (Sizer, 1989) were also carried out along these contradictory axes. Following the cuts Edwards Parkes –the UGC's chairman – eventually admitted that there had been distorted communication between the UGC and the Secretary of State. It appears in fact that Mark Carlisle '[…] began to consider the lower options' (Parkes as cited in Sizer, 1989, p.647) towards the summer of 1980 (specifically on 23rd July 1981, as suggested by Sizer, 1989), a year after he had communicated level funding. To summarise, suffice it to say that:

> '[…] the sheer waste of skilled manpower involved in requiring universities to generate these guesses on hypothetical grant levels while ministers were deciding grant levels of an entirely different order was considerable' (Kogan & Kogan, 1983,p. 46) .

These failures at governing the sector through selectivity have been attributed to the political determination of implementing cuts rapidly and inexpensively. The Conservative government aimed to rapidly downsize the higher education sector without incurring major "dismantlement costs" (Kogan & Kogan, 1983). Since cutting through redundancies – the major cost – would have implied high redundancy payments, government searched alternatives until the very last minute. In fact,

---

[70] The circular letter 16/79 set out the three hypotheses for funding home students as follows 'Allocation (a) is included not in any expectation that total resources will enable these figures to be made good generally, but because at individual universities the Committee may consider such an allocation desirable; (b) represents the continuation of the 1979-80 figure modified only in respect of overseas students; (c) represents a position in which the financial situation could be worse, for some or for all universities'(as cited in Sizer, 1989, p.646).In 1982, this changed to the following grant reduction percentages: '"hypothesis (a)" represented a 2% increase in grant, "hypothesis (b)" level funding, "hypothesis (c)" a 5 per cent reduction in grant' (UGC, 1982, Appendix B,§2 as cited in Sizer, 1989, p.646).

such that the selectivity methods presented in the final grant allocations of July 1981 were elaborated starting from March 1981, only 4 months earlier (Kogan & Kogan, 1983). In December 1980, government had given a preview by asking for reductions in universities' intakes of home students, and in March 1981 the more drastic -and realistic- dimension of the cuts was presented in the following way:

> '[…] by 1983-84 institutional expenditure (net of fee income) will be rather more than 8 per cent below the level planned […] the detailed implications of the plans for the university and non-university sectors are under discussion with the UGC and the local authority associations' (White paper, March 1981, as cited in Kogan & Kogan, 1983).

In other words, the methods for implementing the cuts were decided *following* the decision to implement cuts. This lateness and unpreparedness for radically unprecedented 'selectivity' cuts in the sector reflects the lack of planning done by government as well as its scarce ability to consider alternative criteria, such as socio-economic ones (Kogan & Kogan, 1983), for selectivity. The UGC and the DES were also increasingly struggling to collaborate, leading to the UGC's more submissive behaviour (Palfreyman & Tapper, 2014) where it carried out the cuts without advancing major scrutiny nor impediments to its application[71].

The broad criterion eventually considered was that of limiting new student enrolments through targets and ceilings, making student numbers and costs fall. Within this general decrease, there was a subject component. Student enrolments in subjects of "national economic interest" –the physical sciences, technology, mathematics, computer science, business studies and biology – could continue to increase at constant levels, whereas government intended to make them fall above average for the arts and social sciences (Kogan & Kogan, 1983). The decision to maintain the unit costs stable had implied that, in a context of cuts, student numbers could not be allowed to rise excessively as this would have significantly further decreased universities' income per student, already down by 10% since 1973. Grants for each institution were thus calculated by multiplying the unit costs by the number of students in each type of subject:

> '[…] the letters sent on 1 July 1981 (circular 10/81) were very different from anything received before […] telling every university in the country that it must 'substantially' or 'significantly' reduce student numbers in specific subject areas […]it also required

---

[71] The UGC did not –as was expected of a body considered the universities' ally in negotiations with government (Shattock, 2012) –do much to impede or obstacle the expenditure cuts. As Kogan & Kogan have suggested, the UGC '[…] could have refused to act until the government had showed its hand on compensation for tenure […] the DES could not be resisted […] but could it not have been compelled to face an explicit revolt and the resulting political row?' (1983, p.107).

universities to make cuts in areas that were recruiting well' (Kogan & Kogan, 1983, pp.52-53).

The grant letters also made exceptional mistakes for such a delicate policy, such as suggesting that some universities close non-existing departments and failing to explain what the indications to 'substantially' or 'significantly' reduce student numbers in certain areas translated into in quantitative terms (Kogan & Kogan, 1983). These aspects highlight an improvised and expenditure-led *ratio* of the 1981 cuts, quite in line with Thatcher's willingness to delegitimate higher education and its workers (Ryan,1998). Carlisle's and Thatcher's actions tried to defund the sector by steering it towards lower expenditure targets, rather than implementing selective methods based on quality criteria. The UGC's 1981 cuts clearly reflect a peculiar lack of familiarity with university evaluation criteria even though they were politically narrated as quality-oriented selectivity cuts. This brings to the surface an interestingly exceptional case when read from a NPM framework. Besides the decision to impede falling unit costs, no evaluation of the sector's capacities or specific achievements was made to guide the cuts. If the UGC or the DES had had, during those years, any form of managerial model even sufficient to make cost calculations for implementing a selective criterion of some sort, it would have been more opportunistic and beneficial to use it, if not to merely to avoid the political embarrassment which the mistakes made brought. There was, evidently, no *professional* criterion guiding the DES and the UGC in 1981, whilst the political intent to reduce funding was instead evident. Following the stark criticisms raised to the cuts, the State learned that future cuts would require more legitimate and solid grounds than the shaky justifications of a value-for-money approach to quality and the consequences of the economic crisis. As policies continuedly strained the university sector into financial stringency, they fuelled a managerial engine which had already begun firing within universities, particularly amongst university administrators and vice-chancellors.

## 4.2. Planning for change: managerial models of university governance

Thatcher's expenditure-led embodiment of efficiency-oriented policies makes the notion of a fully-fledged managerial orientation of the state more questionable than scholars suggest (Palfreyman & Tapper,2014), making the sector's adaptation to managerialism a historically problematic perspective (Deem et al., 2007). The result is that the emergence of quality as a criterion guiding the State's use of public resources is intrinsically tied to that of expenditure savings. This brings to the fore similarities and differences between cost-efficient management models and quality evaluation ones, making the conflation of the two as sides of the same coin problematic. Although cost-efficiency and quality evaluation may generally share principles, the frameworks used to introduce

quality through efficiency by government and those used within universities had two diverse articulations, turning the rhetoric of value for money into too general for historically understanding the rise of university evaluation specifically.

Policies geared at obtaining efficiency from higher education clashed with the university sector's own search for efficiency. By the time Thatcher advocated value-for-money in higher education, in fact, managerial frameworks along PPBS lines were already being drafted within universities. The 1981 cuts were a big challenge, reducing overall resources in many universities – such as at East Anglia where they fell by 12.2% (Kogan & Kogan, 1983). This context created a new challenge, that of internally managing the cuts (Shattock, 2012; Kogan & Kogan, 1983). The grant letters were an unprecedented form of negative political interference in the university sector, made more dramatic by the fact that they left university agents with crucial decisions regarding where to cut and how, for '[…] expenditure reductions still left the initiative to find ways of reducing unit costs to the individual establishment' (Neave, 1985, p.358). The constraints posed by the 1981 put vice-chancellors in a position of practically deciding how to cut internally an uneasy task. As I argue, however, Vice-chancellors turned this arduous challenge into an opportunity for themselves. Specifically, the CVCP quickly understood that the confusions and anger created by politicians could turn the committee into a more trustworthy body. In the autumn of 1981, that same autonomy which had left universities free to allocate resources during expansion had turned into a burden: whereas earlier university heads could have been appraised for their recruitment and development choices, now they were handed down with the responsibility - in the name of autonomy - of *implementing* the cuts. As Guy Neave (1985) has put it, this'[...] was a rare example of academics putting their house in order […] the traditional relationship based on State non-intervention still pertained.' (p.358). The urgency to deal with the cuts was felt by many.

Administrators were very surprised at the extent of the cuts and attempted to make sense of the new political climate by focusing on what could have been done differently on the higher education side. In a book authored by the Conference of University Administrators (CUA from now on) published in 1983 and headed by M.A. Higgins–then Registrar at Loughborough university– members compared the actions of British registrars and administrators with those of their US counterparts. They pushed for a change in the traditional relationship between HE leaders and politicians, and with the Treasury in particular. The CUA lamented that the British were too obedient to politics, and that as the 1981 cuts showed, they had proven unable to halt radically disadvantageous political decisions. For the CUA, British HEIs had much to learn from their American counterparts, starting from the diversification of their income sources, for British HE was excessively dependent on very limited amounts of public expenditure. They received 2.5% of total public expenditure and 23% of the DES's (Shattock & Rigby, 1983). In budgetary terms, the CUA

suggested that British universities begin to assimilate US universities' internal governance structure. Acquiring more bargaining power over allocation decisions required that HE representatives act *before* the publication of government's expenditure White papers. White papers were the results of political negotiations between the Treasury and departments, first, and between Treasury and cabinet at a second stage. HE administrators' ability to influence their contents was highly restrained by the fact that political actions and eventual criticisms to the decisions tended to occur at the second stage. The result was that –as the reactions to the cuts showed–objections were often directed at the DES following its already-made agreements on allocation sums, rather than being directed at the Treasury during the decision-making process. For the CUA, British university administrators' weak negotiation abilities in the political realm were in turn influenced by British universities' internal governance structure. Their collegial model, it argued, translated into lengthy decision-making processes in universities. Contrarily to US universities' pyramidal governance structure, British universities discussed and approved internal budgets through committees, and could not benefit from the executive model where individual officers had a degree of autonomy in taking decisions through restricted groups such as the board of regents. For the CUA, this internal organization allowed US universities to '[…] live with a more bumpy and uncertain economic climate' ( Shattock & Rigby, 1983), and endowed them with sufficient flexibility for taking quick internal decisions (such as the budget totals required) in moments of crisis. They could also rapidly intervene as lobbyists in political circles, as the presence of the American Council of Education (ACE) and the Association of American Universities (AAU) within State legislature exemplified. British Administrator thus cast their concerns on their weak political influence both within and outside universities as an organizational issue, suggesting that governance be re-organised away from a strictly collegial governance model. Accelerating internal decision-making by having administrators act within an executive governance model was considered a key step for more effective political negotiation of HEIs with both the Treasury and the DES (Shattock & Rigby, 1983).

The CUA was not alone in advocating for significant changes in the status quo. Universities' 'enforced retrenchment' (Shattock & Rigby, 1983) can in fact be defined as an ignition flame to a managerial revolution that was already taking place within universities. Indeed, the post-cuts context gave previously marginal advocators of a managerial turn an opportunity to speak up with a now more attentive audience willing to hear, an audience shocked by the "attack on higher education" (Kogan & Kogan, 1983), and fearful of future similar policies. By the mid-1980s, the Vice-chancellors grouped in the CVCP had gone from rejecting Shirley Williams' "thirteen points" for efficiency in universities to a *de facto* endorsement of university funding selectivity '[…] in our view a university which did not have some good research in most of the subjects taught would not

be worthy of the name' (CVCP,1984, as cited in Shattock, 2012 p. 174). Further, by the end of the 1980s the committee seemed keen to replace the UGC:

'[…] it is clear that a shift in power and control is in progress with the elimination of the 'buffer State' and a polarisation of interests […] it will fall to [the CVCP] to take on the historic and necessary role previously played by the UGC in 'speaking up' for universities and making sure their needs are appreciated by government' (Flowers Report,1988, as cited in Shattock, 2012).

This new stance on selectivity owed much to the cuts, though not exclusively. The uncertainties created by the abovementioned policies of the 1970s  (Kogan & Kogan, 1983) had importantly fuelled new opinions on how to change in the context of a tightening resource belt. The consensus amongst administrators and vice-chancellors, for example, silently began transforming as early as the mid-1970s, and was increasingly inclined towards the adoption of efficiency models. These were often discussed in research on educational governance, which dissipated through key research hubs such as London. From dealing with more "on the ground" issues, research on universities importantly turned to policy and its effects on university-level developments, and an array of policy proposals slowly emerged from such research streams. The diffusion of this new research facilitated a greater degree of universities' political leverage over government, which had proven to be largely deaf to the sector's requests. Researchers, vice-chancellors, and administrators first searched audiences outside of university walls, developing studies and alliances with industry (such as Cadbury, BP), foundations (such as the Leverhulme trust) and consultancies (such as McKinsey, Coopers & Lybrand, Peat Marwick, and Mitchell), seeking to find support and strengthen their views vis-à-vis government's. A central hub for the development and dissipation of a new wave of HE research was London, home of both the Institute of Education (IOE from now on) at UCL - today one of the largest education research bodies in the UK- and the Society for Research into Higher Education (SRHE from now on) were situated.

### British research on university governance

The SRHE emerged in 1965 following Robbins' recommendations, taking it upon itself to establish research in conjunction with trusts and foundations (Shattock, 2015). The society's research immediately following its establishment was based on 'on the ground' topics such as access, working conditions, number projections and expansions, curricula, and teacher trainings, but this changed in the mid-1970s when financial difficulties drove the society to make a turn. The

hostile financial climate of the 1970s had however complicated the society's reliance on stable funding as neither the DES's nor the SSRC's sufficed (Shattock, 2015). Following Gareth Williams' proposal to engage directly in HE policy research, the society secured new financial partnerships which allowed it to stay afloat. The partnership's research focused on more imminent issues: financial sustainability, administrative change, and the governance of HE in contexts of marketization pressures. Williams was aware of international developments in human capital theory and the debate surrounding it (Williams, 1982), and for him the merging of economists and educationalists was a significant success of HE research. Williams referred to the favourable political context of the 1960s as a case where 'the motorist whose car is running well has little reason to look under the bonnet' (Williams, 1982, p. 105). The 1980s, instead, was a decade in which economic and managerial knowledge would bring important lessons to educators hoping to untangle ways for dealing with political and economic constraints. He suggested that existing attempts to use management research in education could improve, for, as he put it, '[…] it is not possible to tell a convincing story about inputs and outputs without understanding the processes by which inputs are converted into outputs' (Williams, 1982, p.105).This thrust gave the SRHE new lifeblood. By 1980, it had secured a grant of £250,000 for research into the future of HE from the British Leverhulme trust[72] which was used to run the "Leverhulme's programme of study into the future of Higher Education". This series of seminars was held between May 1981 and May 1983 and discussed new research, which culminated in a wide array of the society's publications (Shattock, 2015). These were well-publicised and exerted significant influence in both the university sector and politics[73]. The society's decision to turn to research on HE governance structures therefore eventually opened new doors to financial support and political attention from the DES and education ministers, such that Heath himself is said to have consulted the society for policy advice (Shattock, 2015).

If for the SRHE research with the Leverhulme trust was a first go at the study of a new steam of knowledge, for the latter its support of the SRHE was yet another piece in the mosaic of its existing funding of management studies. During the triennium 1968-70 funds were directed at research on "manpower issues" and the trust invested 75% of its social sciences funds[74] in the development of business and management, industrial relations and accountancy research (Briggs, 1991).The prestigious Leverhulme funds were keenly taken on and used, facilitating and accelerating, amongst others, the construction of a comprehensive management system at Sussex

---

[72] The trust was founded in 1924 following the death of Sir Leverhulme, who expressed, in his testament, the will to direct part of his assets to research and education, and, specifically, to research on education (Briggs, 1991).
[73] As Shattock explains (2015), publications such as 'Excellence in diversity' (Berrill, 1983) cause both drama and curiosity, stimulating, for example, the research of the Reynolds committee into academic standards, leading into the Reynolds report (CVCP, 1986) as well as the establishment of the academic standards group.
[74] A total sum of £181, 260 (calculated from data provided by Briggs, 1991, p. 158)

University thanks to a fellowship in university accountancy awarded to John Fielden in 1968 (Briggs, 1991). Fielden used his fellowship to apply his consultancy experience acquired whilst at Peat, Marwick, Mitchell & Co to the higher education context. With Geoffrey Lockwood working at Sussex, he wrote vanguard book "Planning and management in universities: a study of British universities" (discussed in section 3.2.2.1.). As he specialised in Higher Education accountancy and set out many new principles in the process, Fielden became increasingly convinced of the benefits that analytical planning could bring to British universities' administration, so much that he advocated for a consideration of pros and cons of applying PPBS experiments from US HE also in the UK (Fielden, 1969). Fielden later expanded his research to topics of marginal costing in university resource management (Fielden, 1978). For him, the use of PPBS would permit closer scrutinization of how each extra pound of resources was being spent, as opposed to continue making average cost considerations which were less effective at relating economies of scale with performance assessment. Together, Lockwood and Fielden used Leverhulme funds to develop an experimental governance model at Sussex university which made used systems analysis principles (Lockwood, 1981). Considered the most managerially advanced university in the country for this reason, Sussex university also obtained the Leverhulme trust's first fellowships in university administration[75].

The Society for Research into Higher Education's (SRHE) new alliance with the Leverhulme trust was a key channel through which the society not only gathered new research, but also solidified its influence and created reciprocal opportunities. The society's work and extending force in turn augmented its prestige and that of its affiliated researchers. Through its book series and journals – amongst which the famous "Higher Education Quarterly"–the Society publicised the work of academics specialising in topics ranging from curricula development to accounting and management. Thanks to its partnerships with important publishers – such as the Open University Press -  many academics working with the SRHE later became renowned for their contributions to British HE research, such as Tony Becher (Becher et al 1977), Shattock (1983), Halsey & Trow's "The British Academics" (1971), Moodie and Eustace's "Power and Authority in British Universities" (1974) and Kogan and Becher's "Process and Structure in Higher Education" (1980). Importantly, the Society welcomed the work of "outsiders," namely staff working at the DES or at the UGC or at international circles. By integrating works coming also from outside the university walls, the society's search of both non-academic and academic publicised more technical views of British higher education. This resulted in the construction of an important research space where

---

[75] Additionally, a part of the Leverhulme's funds were explicitly destined to preliminary evaluations, over a five-year period, of newly established courses in business studies at the London Graduate School of Business Studies and at the Manchester Business School. (Briggs,1991).

managerial approaches gained consensus and acceptance within the broader field of higher education research.

One of the SRHE's management-oriented publications stands out as an example of the society's ability to diffuse emerging studies which were still at the margins of the wider debate on university governance. This is the case of Lockwood and Davies' "Universities: the Management challenge" (1985). Both researchers were working extensively on university management for dealing with an 'externally changing environment' using and organizational approach. Both Lockwood and Davies had, in fact, been already actively involved in international projects seeking to develop higher education management models (see sections 3.2.2.1. and 3.3.1). Lockwood's work on the application of PPBS principles in higher education had begun as early as 1969 (Lockwood, 1972) at the OECD and was then presented in the UK in 1973 (Lockwood & Fielden, 1973). Similarly, John Davies was a key figure in the CRE's vice-chancellor management trainings. In the co-authored book, the two argued that universities be conceptualised as organizations living inside an *open system* (Katz & Kahn, 1966), meaning that their coexistence with an externally changing environment relied importantly on their managerial ability at transforming external' forces – positive or negative – into constructive inputs and opportunities. In setting 'agendas for university management in the next decade' (Davies, 1985), they argued that vice-chancellors become familiar with universities' *life cycles* (see appendix C), namely their moments of birth, development and growth. This allowed to plan university activities according to institutional needs at different stages and ensure that they produce exhaustive '*returns to inputs'* (Lockwood & Davies, 1985). For the authors, this influenced universities' ability to signal their reputation to the external environment, composed of government but also of other universities, research bodies and non-governmental funders. Situated in a framework which extensively borrowed from the managerial experiments in higher education (Balderston & Weathersby,1972), the authors attempted to help UK vice-chancellors and administrators 'improve [universities'] manageability', by assimilating the trends in US universities, which paid vast attention to 'market research and institutional advancement' for sending efficiency signals to a wide array of stakeholders (Lockwood & Davies, 1985). Davies's views espoused well with Lockwood's (Lockwood, 1979), for he also considered the vice chancellor as the key managerial agent of transformation (Davies, 1980). As the language used by these researchers debated such administrative and managerial aspects, VCs had more material at hand for analysing their role within their university. importantly presenting a new challenge to the collegial conception of the university. Suggestions made included the following:

> '[…] universities are organizations which have *corporate responsibilities*, and which
> possess powers to manage the activities of their members in order to carry out those

responsibilities. Members of the academic staff collectively constitute the major element in the government of a university, but individually they are employees by contract […] in terms of organization and management the existence of a university creates a firm boundary. Faculties, departments and other units *are not autonomous units* within a guild structure, *they are inter-dependent parts of a unitary organization*' (Lockwood & Fielden, 1973, p. 20 as cited in Lockwood & Davies, 1985, emphasis added)

This organizational framing of the university along a new hierarchy of powers, found resonance in the propositions of the Jarratt Report, a strikingly *political* document in which the CVCP expressed its determination to become a key player in both the higher education sector and individual institutions. As the authors showed, conceptualising the university similarly to a firm, staff as employees rather than as autonomous units endowed with self-governance powers; and leaders as executives, created new spaces of action for VCs, who needed to have greater leverage than hitherto over decision-making processes if they were to assist universities' modernization. By the end of the 1980s the view that VCs act as chief executives monitoring the financial health of their universities was supported by suggestions, for example, that they become responsible for drafting institutional contingency plans to be implement in moments of grant reductions (Sizer, 1988).As both Lockwood's and Davies' previous international research began to show its fruits when discussed specifically for the UK context and within the British research environment, so did new governance notions dissipate and become increasingly shared among institutional representatives. This broader managerial movement in HE research deserves further research in scholarly studies, for it is likely that there existed also other research networks bringing their previous research to the table in the context of the 1980s. The research developments discussed so far indicate that administrators (the CUA) and groups of academics alike were importantly asking themselves questions on how to bring about concrete change in internal management. Certainly, Lockwood's own research found fertile ground in the CVCP's changing conceptions of university governance, which the committee explicated in the Jarratt Report. Borrowing significantly from Lockwood, Fielden, Davies and more broadly from the developing managerial movement in higher education research, the CVCP set out its own new efficiency guidelines which solidified a managerial re-articulation of university governance. The Jarratt report voluntarily marked a radical break from the status quo, shaping also the committee's later proposals on quality evaluation principles and practices.

### 4.2.1. The Jarratt report: systems analysis and managerial governance

The proposals published by Vice Chancellors and administrators in the Jarratt Report (CVCP, 1985a) advanced a more substantiated and *university-tailored definition of efficiency* compared to government's, which as we have seen articulated efficiency as a synonym of financial stringency at

the turn of the 1980s. Both the UGC and the DES still struggled to provide universities with clear indications on how to practically implement efficiency at the institutional level, (Cave et al., 1997; Shattock, 1984;) and this trend did not change after the contradictory selectivity cuts. In the build-up to the mid-1980s, when the Jarratt report was published, there was still little clarity over the government's concrete understanding of quality. As has been pointed out, '[…] government was […] interested […] in improving quality and efficiency in higher education', yet '[t]hese concepts were sufficiently formless as to be difficult to identify with particular policies except in very general terms.' (Shattock, 1984, p.205). These general terms often invoked financial efficiency as a common denominator. Keith Joseph's "concordat" in 1984 confirmed this approach when he asked universities to simply provide evidence of their engagement with financial management practices along the lines of Rayner's initiative.[76] This further demand of efficiency in universities did not refer to specific quality methods or assurance procedures. For the CVCP and an array of academics, articulating efficiency in relation to quality was a less arduous task than for government. Also the use of Performance Indicators (PIs) had, by the mid-1980s, already been extensively discussed, particularly by John Sizer who carried out analyses of their possible applications (Sizer, 1981), and later contributed to the writing of the Jarratt report (Tyles, 2008).

The study –organised by the then chairman of the CVCP, Maurice Shock – was carried out between 1984 and 1985 by a restricted committee of 13 members selected by Sir Alex Jarratt, then chancellor at the university of Birmingham. The committee which he formed was seen as that group with the necessary networks and expertise for making judgments on universities' managerial health. It had a mixed ('academic-free') membership made up of VCs, industrialists (CEOs of Cadbury Schweppes, Ford, and Plessey corporations) the prime minister's personal adviser on efficiency, Robin Ibbs, and a UGC representative, Peter Swinnerton-Dyer[77],who was rather critical of Edward Parkes' chairmanship of the committee. Geoffrey Lockwood, invited by the chairman, also participated. Ibbs's presence in the committee did not create tensions, for he had previous working ties with Jarratt. Together, members used their expertise lay out a set of proposals on university efficiency (see appendix D for the list of members). Above all, the report assessed the effectiveness of management structures, studying institutions' arrangements for monitoring and steering resource allocation towards cost-efficiency (CVCP, 1985a). In many ways, the Jarratt report represents universities' equivalent of the financial management initiative carried out along similar terms in other public sector institutions. Differently to other public institutions, however, the university

---

[76] 'Sir Keith Joseph told the House of Commons that the Government was to consider with the UGC and CVCP some further financial provision for the universities in 1987-88 and the following years but that the 'the government's willingness to make such provision depended crucially on evidence of real progress in implementing and building upon the changes that are needed (HC Debates Vol 89, Col 179' as cited in Cave et al., 1997, p.4)

[77] Swinnerton-Dyer was Chairman of the UGC until its demise and first president of its successor, the Universities Funding Council (UFC).

sector was granted the opportunity to carry out the efficiency studies *itself* through the steering committee, rather than through external bodies. This 'Rayner style' efficiency review had been discussed between government and university representatives starting from 1983 (Shattock, 1984) to later culminate in Joseph's 'concordat', which reflected his  to avoid being directly involved with efficiency through the DES and instead have representatives produce that hard evidence necessary to prove that universities were responding to political demands. This concordat shows how the sector's  – or more precisely the CVCP's – exceptional relationship with the State seemed to guarantee a degree of executive autonomy.

The participation of universities was rather limited – with six universities out of a total of forty-one in the sector included in the study – and did not follow a selective criterion for it was on a voluntary basis. Universities eventually assessed were Edinburgh, Essex, Loughborough, Nottingham, Sheffield and UCL. All were considered for the 'general study', whereas a restricted number were selected for 3 specific studies: Edinburgh and UCL participated in the 'financial management' study, Essex and Sheffield in the 'purchasing study' whilst Loughborough and Nottingham in the 'buildings' study. The Jarratt committee explicated its managerial purpose from the onset:

> '[…] the commissioned studies will not extend to issues of academic judgment nor be concerned with the academic and educational policies, practices or methods of the universities', and prioritised judgments to evaluate '[that] optimum value is obtained from the use of resources, that policy objectives are clear, and that accountabilities are clear and monitored' (CVCP, 1985a,p.6).

The report pushed for a major role of non-academic councils vis-à-vis senates by emphasizing the former's greater expertise with financial and administrative matters, in this way arguing that managerial agents make their way through at the cost of academic participation in institutional governance. For the committee, councils were crucial because:

> 'Vice-Chancellors have virtually no formal constitutional powers other than those which may be -delegated to them to act on behalf of Council, Senate or a committee […] It does, however, lie within the powers of Council to strengthen formally the executive role of the Vice-Chancellor'. (§3.59, p.26).

Although the study did not investigate strictly academic issues and used a rather well-brushed managerial language to define the "objective" scopes of its enquiry, the recommendations crucially touched on the organization of academic activities for they suggested to re-structure academic

staff's participation in university governance. The report largely undervalued academic staff, suggesting that academics (and, by extension, senates) were unfit for what universities most needed, 'strategic' planning (CVCP,1985a, §3.49). Underneath the committee's attempt to limit the influence of academic senates in qualitative and quantitative terms there was the concern over the growing influence that these had acquired during the expansion of the 1960s. To sustain this side-lining, the committee portrayed academics as reluctant to modernization :

> '[…] It does seem to us that the relative decline in the exercise of influence by Councils has increased the potential for Senates to resist change and to exercise *a natural conservatism*. Vice-Chancellors and university administrators have, in the past, been trained to believe that harmony between the two bodies should have a very high priority in a university' (§3.50, p.24)

This supported the message that academic senates were ill-equipped in modern technical aspects of governance, and simultaneously served as a basis for igniting discussions on practices for monitoring academic activities. By stressing the importance of planning systems and the urgency of implementing them to avoid further financial downturn, the committee justified more space of action for its technical and 'expert' approach to governance. The fact that academic senates were increasingly populated (CVCP 1985a, §3.48,), was identified as that factor which *reduced the efficiency of planning decisions*: '[…]large senates are not the best places in which to undertake planning and resource allocation' (CVCP,1985a, §3.50g,). Councils' non-intervention in academic affairs was thus  unjustified because it was incoherent with the division of executive role models in industry and in the corporate world (CVCP,1985a, §2.2). To enhance planning, the committee proposed the use of new, technical criteria for academic staff recruitment inviting universities to recruit on the basis of cost-benefit considerations (such institutional goals and student number targets). Academic tenure – "the tenure issue" (CVCP,1985a, § 2.7d) – was listed amongst those factors which '[…] complicate the managerial process' (CVCP,1985a, §2.7) because of academics' tendency to resist and obstruct change. The committee voluntarily provoked senates in these statements, it defined new tensions between councils and senates as '[…] necessary in the circumstances now facing universities and can be creative and beneficial in the long term. That can only happen if Councils assert themselves' (CVCP,1985a, §3.50)[78].

---

[78] To give another example, 'The special powers vested in the Senate stem from the inherent nature of teaching and research and enable it to protect academic freedom. However, in times of financial constraint the potential for friction between the Council and Senate is increased; and the changing circumstances of universities since 1981 are now leading to a questioning of the working relationship which prevailed throughout the period of expansion.' (CVCP, 1985a, §2.2., p. 8).

A reduced scope for academic senates in university governance ran thus parallel to new and different powers for council and expert members. The distribution of powers proposed rested on an organizational understanding of the university *as a system* where staff is gathered in "units" which can be coordinated from above for the organization's functioning. Following these lines, the committee proposed that universities set up planning committees ideally directed by a VC and appointed by the council. This new committee would use management practices to '[…] integrate academic, non-academic, financial, and physical planning' (CVCP, 1985a, §3.43b). The idea was crucially borrowed from Lockwood & Fielden's work (Lockwood & Fielden, 1973), which had advocated a new model of executive governance vested in participative terms. They had similarly argued that '[…] the Vice-Chancellor should be either personally responsible for planning or should appoint someone with that responsibility' (Lockwood & Fielden, 1973). The Jarratt committee rearticulated this into a 'planning and resources committee' in which Vice-chancellors could surround themselves with a restricted number of lay and academic members, asked to '[…] pursue the corporate interest of the university and not to represent sectional interests' (CVCP, 1985a, §3.43b). This 'top' planning structure served to fortify a broader new structure of the university as a whole. The latter was ideally composed of multiple planning committees overlooked by the 'top' planning and resources committee. The merging of academic and non-academic planning was proposed in order to widen the scope for budgetary concerns into the organisation of academic activities, effectively subjugating the autonomy of the academic senate to that of council's experts.

The fusion of academic and non-academic planning could be set off through universities' budgets. Specifically, the committee proposed modifications to processes of resource planning and use. Universities could no longer afford to spend public resources through informal or generic processes, hence they needed to come to grips with the inevitable turn to accountability. The abolishment of quinquennial funding had eroded most hopes of future incremental resources and ignited the practice of yearly financial planning in universities. This was to an extent an opportunity for aspiring university managers, who could step into this new context with concrete solutions to a still largely unfamiliar institutional practice. The idea to introduce the notions of budget constraint and budget devolution went in this direction. This step implied more than just the end of incremental resource-planning. Budget constraints would work to ensure certain expenditure thresholds. Budget devolution introduced a key innovation, as in this dynamic the university would function through *multiple* budgets. Devolution of budgets represented a way to introduce budget constraints across the entire university precisely by breaking down a university's large, single budget into many smaller ones, each to be planned and spent meticulously. This dynamic increased the scope for resource planning and for planners. The committee made sure to explicate this novel organizational structure by referring to budget centres (CVCP, 1985a). This structure emulated to

that of 'spending units'[79] (Lockwood & Fielden, 1973). The idea was that institutions would function through the logic of a budget constraint, with each unit – or level– of the university (the faculty, the department, and the administration) being held accountable for its own budget. Multiple budgets ideally belonged to university departments, which would set up budget centres. This 'fragmentation' did not eliminate the university's overall budget. However, the latter was now conceived as an aggregate of multiple ones. By *redefining departments as budget centres*, a planning and resource committee would become entitled to exert the new powers being advocated and to demand that new initiatives be taken on for the scope of efficiency and budget control, such as expenditure reporting. This would function to 'integrate academic and non-academic planning' as departmental budget plans also needed to work with budgeting considerations, whereas traditionally this function had remained outside departments. Placing budgetary responsibilities on academic departments in this way was in many respects a process through the CVCP constructed new practices of accountability. The reorganization of internal powers thus did not limit itself to a mere articulation of councils' roles but relied importantly on new interdependency relations constructed through novel procedures, recalling Dahler Larsen's suggestion that the construction of evaluation often depends importantly on the o*rganizationalization* of management structures (2019). Heads of departments in the CVCP's proposals turned from being responsible of departmental activities to being responsible for the *expenditure* of their department specifically, just as Fielden and Lockwood envisioned the development of deans' roles (Lockwood & Fielden, 1973).The committee thus also proposed new modalities for electing heads of departments so that non-academics could also compete for the position (CVCP, 1985a). Introducing budget devolution thus allowed to pursue efficiency at the cost of transforming academic activities into interchangeable objects. Academic skills were no longer crucial for running a department. Further, the report's concerns on accountability and efficiency did not consider how such modalities restricted the ability to carry out research and teaching. Academic activities were first and foremost cost-incurring activities which could be readjusted according to budgetary concerns. The pursuit of budgeting thus lay in the containment of the institution's overall budget, which in the process entailed the second innovative procedure of measuring staff performance through cost-analysis.

These aspects clarify how the committee tried to justify the rise of a managerial power acting towards efficiency and placing the assessment of academics at the centre of such a process. Budget devolution has been interpreted as a feature of NPM's decentralizing push, (Deem et al., 2007), as that aspect granting new freedoms to university groups. In this perspective, it is broadly

---

[79] '[…] the structure of the budget should consist of a grid of spending items and spending units. The items would consist of objects of expenditure (e.g., academic staff salaries, support staff costs, telephones, consumables) though eventually they might develop into programmes based upon the purposes of expenditure' (Lockwood & Fielden, 1974, Appendix 6/3).

seen as one of the many pieces of evidence of NPM logics playing out in higher education. However, it was more complex than that, vested with multiple meanings and implications which tend to be ignored by academic literature. The way in which budget devolution was articulated by the committee, in fact, treated it as a centralizing device –particularly in planning – rather than as a decentralizing mechanism geared at delegating powers as the NPM framework implies. Despite the non-academic scope stated in the report, the Jarratt committee was well aware that influencing resource allocation along cost-efficiency axes necessarily implied intervening and interfering in academic matters, if only for the simple reason that they are universities' central and most cost-incurring activities. This emerges clearly in how the committee insisted on the relationship between new financial 'autonomies and accountability:

> 'We stress that in our view universities are first and foremost corporate enterprises to which *subsidiary units and individual academics are responsible and accountable*. Failure to recognise this will weaken the institution and undermine its long-term vitality.' (CVCP, 1985a, §3.41).

Further, the committee emphasized how such subjection to assessments was to be considered universities' new organizational principle:

> '[…] Unless a university pays rigorous and systematic attention to the quality of its resource allocation and monitoring procedures it is certain to be wasting its resources, especially the time and energies of its academic staff, and not giving the best value for money.' (CVCP, 1985a, §3.3.8b).

Comparing this last standpoint with Lockwood and Fielden's (1973), it emerges that Lockwood's role in the committee importantly imprinted the Committee's orientation. Universities could not avoid a full integration of the management model being proposed, just as Lockwood and Fielden made clear when conceptualising the university institution as a system:

> '[…] A university is a system; the understanding and management and its activities cannot be properly achieved without recognition of that fact. The use of the systems analysis techniques can be maximised only when such techniques form an integral part of a total system for the management and planning of an institution (Lockwood 1973, p.97).

The report diluted the systems analysis language to crystalize this same view, and to concretely work towards its application it initiated a new stream of work for the CVCP on performance indicators (PIs). Through the Jarratt Report, the committee had introduced, for the first time and in a comprehensive way, its political stance on managerial governance in the university sector. Yet there was still a lot of work to be done, for academics as well as many administrators and VCs were still uncertain as to what exactly PIs were, how they were meant to function and what exactly their use entailed. It was between 1986 and 1990, as managerial notions were discussed for research evaluation purposes, that they would begin to have a better grasp of the implications of performance assessment, still somewhat secondary in the report to inexpert readers. Exploiting the political narrative on the accountability of the public sector was allowing allowed systems analysis concepts to develop and introduce and elaborate on the suggestion that VCs and administrators were the most suited agents for managing universities' responses to their turbulent surrounding context. In this way, I argue the steering committee made a pragmatic use of its context as a means to legitimate a new type of influence both at the sector level, acting as a group, and within institutions, acting as individual managers. Hence, efficiency became an opportunity for vice-chancellors much more than a subjugating force that could not be escaped. It was used to advocate that executive and managerial figures deserved more say in university governance than that which they had acquired in the collegial model. The degree of their innovative articulation of efficiency cannot thus be stressed enough. In their demands of efficiency, politicians had not requested such a specific articulation, and perhaps neither expected that which eventually emerged from university walls.

## 4.2.2. Institutionalising new practices

Although through the Jarratt report the CVCP created the basis for a new system of resource management within universities, it was well aware that it needed to go beyond a mere appraisal of planners' and managers' expertise. Academic staff was well aware that the status quo was increasingly vulnerable, yet this did not mean that a dismantlement of collegial governance would automatically occur. Further, academic resistance was likely to emerge following the delegitimating provocations which the academic body was receiving from within (Trow, 1992). The CVCP therefore knew that concrete evidence of its claims was an important step, making the road to the construction of a system of 'policies, practices and procedures to improve efficiency' (CVCP, 1985b, p.39) a rather bumpy one. If it sought to implement principles of efficiency and performance assessment, the committee needed to show what these meant and what their use implied. These concepts had been around for a while (Sizer, 1989), yet there were very few cases of a concrete and substantiated use in universities. The CVCP tried to build on the Jarratt report's recommendations

and encourage this in second report, the National Data Study (CVCP, 1985b)[80].Specifically, the report touched on topics of performance assessment and its relationships to a proposed national selectivity funding framework based on quality evaluation. The thrust of this study was supported by the UGC which contributed to the report. The CVCP was happy to use its support, albeit with some reservations. In 1985, the committee spurred grants committee to re-organize its functioning if the two were to seriously collaborate. In pointing out the UGC's weaknesses, the committee proposed a 'review of the structure of sub-groups and their relationship to the committee itself.' (CVCP 1985b, §4.21). The UGC's clumsy management of the 1981 cuts, further, supported the view that a 'technical' turn in the committee's membership and functioning was needed, for '[…] the environment calls for a greater capacity for analytical thinking in the secretariat than hitherto' (CVCP1985b, §4.22). These proposals were an active attempt on behalf of the CVCP to exert political leverage over the State in a moment of debate over universities' governance by bringing the UGC on its side. This was rather successful, leading to grounds for a productive collaboration between the two. In part, this is because the CVCP was aware that Swinnerton dyer – by 1984 the new chairman of the UGC – was keen to reverse the UGC's damaged reputation (Shattock, 2012). The CVCP's suggestions also responded to its own preoccupation of detaching itself from the influence of the DES and have more space for autonomous action in universities, concerned that the DES could try to take control of the reins of the sector. The fact that the Jarratt committee explicitly wanted to build a robust presence vis-à-vis the DES emerges in its suggestion that:

> '[…] the need to strengthen the [UGC] secretariat in this way will raise questions about how appropriate it will still be for it [the UGC] to remain so closely tied to DES given the sensitivities about the need for some distance between the DES and the UGC.' (CVCP,1985b, § 4.22).

Swinnerton-Dyer's own reversal became evident upon his autonomous decision to initiate research on a national selectivity exercise for research funding, where '[…] neither the DES nor the secretary of State personally were at all involved in the decision and neither were consulted beforehand' (Shattock 2012, p. 175). The national data study thus created important arguments in favour of a more proactive and assertive UGC which could guide universities and evaluate them with their approval, rather than through uncertain criteria or targets. The UGC and the CVCP saw selective funding as the 'node' through which their ambitions could converge and develop, helping them to act as first-steppers in the construction of evaluation criteria. This emerges clearly from an attentive analysis of the National data study.

---

[80] Conducted as an integral part of the Jarratt report.

Particularly central was the role reserved to the UGC for coordinating universities through objectives. The study argued that it become more expert in taking allocative and selective funding decisions, by focusing on objective-setting: '[the UGC ought to] include some indication of how the university's performance can be judged against its aims' (CVCP, 1985b, § 4.9). The UGC's previous failure at transparency was raised as a further motive for having a more coordinated relationship between the UGC and the university sector, a coordination that would in future be based on shared objectives and criteria. The document explicitly asked the UGC to act more decisively on the criteria that it saw fit for allocating funds. Differently to the Jarratt report's focus on institutional resource management, the UGC's criteria would ideally be sector-wide, and this obviously concerned universities. The study posed questions such as whether the UGC preferred a holistic method of making institutional-level judgments or if it was more inclined towards subject-based criteria and funding (ibid § 4.15). Tied to this, the UGC was asked to find criteria for allocating portions of teaching and research activities (§ 4.10). These considerations added a second dimension to efficiency compared to that of the Jarratt report, now defined also as *the coherence between sectorial objectives and results obtained at the institutional level*. Thus, efficiency was not defined merely as a function of (the quality of) internal planning structures, but also as the adherence to new criteria which the UGC would devise. Behind these considerations lay thus the innovative manipulation of the political discourse of "value for money", used to insist that universities' achievements were no longer considered as legitimate as in the past; and that this was something which universities could not object to or modify, but rather had to accept by adhering to the CVCP's articulation of efficiency and accountability. This was an opportunistic move on behalf of the two bodies because they insisted on external constrains but were well aware that no external and politically-demanded criteria existed yet, reason why they were the ones seeking to devise them. Thus, politically accountability was being indeed requested, however not through rigid nor yet evident benchmarks for these did not yet exist for the sector. Whereas universities' validation – and to an extent their institutional legitimacy –previously concerned their degree-awarding powers, the CVCP and the UGC were now beginning to actively put this under strain by asking universities to adhere to new criteria for proving their worth.

Proposing a national evaluation framework that exposed universities by showing their internal and external assessment results was no simple task. There needed to be a degree of horizontal consensus on shared objectives and evaluation criteria. For example, universities could not simply be asked to agree on new objectives when charters were themselves an autonomous expression of institutional missions. From the CVCP's perspective, the planning and resource committees proposed could not function and assess staff with insufficient information on its activities (such as time spent on teaching and research, costs incurred, student staff ratios, etc).

Similarly, the UGC could not suddenly claim that an institution was inefficient and justify resource allocations without 'hard evidence' inefficiency nor justify the amount of resources allocated to it without 'hard evidence' to back up its claims. In other words, both committees were aware of the crucial importance of obtaining 'input and output' data from the sector which they sought to more directly govern. This information would allow for a wide variety of scopes, the most important of which was the setting of *both institutional and sectorial benchmarks* supporting evaluation: '[…] to draw up plans, and to evaluate performance against them, requires information about achievements' (CVCP, 1985b, §3.6). Such information could not be forcibly demanded of academics, but, rather, had to be gathered through consensus for, crucially, it had to be provided directly by them. To confront this issue, the steering committee made a "pragmatic use" (Konings,2010) of the sector's 'external environment'. The committee articulated data and information on academics' research and teaching activities as a quality-signalling device for 'data users' such as the UGC and the DES, responsible for funding. The stress on the uncertainties caused by the turbulent political climate was functional to the construction of new criteria in the higher education sector, built through the manipulation of information on its activities. Although universities lamented the 'external context,' the national data study suggested, they did not seem to be doing enough management work to construct '[…] a planned strategy to maximise future effectiveness' (CVCP,1985b, § 3.4,). Planning through information was key so that universities could decipher '[…] how the balance between competing aims [research and teaching] should be struck' (CVCP, 1985b, §3.5). Like the Jarratt report, the national data study did not see an alternative to planning. Here too, the university's success with its 'external environment' was brought back to a managerial question .

Engaged in spreading an evaluation culture, the national data study introduced preliminary external evaluation criteria. By building on the principles of system analysis, the study invited universities to compare their institutional goals –to be defined through the new decisional dynamics proposed in the Jarratt report – with these and broader objectives. The idea was that both universities and the UGC could improve goal definition and engage in a dynamic of constant negotiation between the institutional and the national priorities. The national data study gave a go at defining national research quality criteria as follows:

'[…] it is possible to gather some simple picture of research output (e.g., publications) achieved, presumably, by using working hours that remained after teaching commitments had been met […] each department could be required to produce a brief annual report (4-5 pages) accounting for its expenditure (including the use of staff time); this would need to include measures of outputs and of input in order to assess its activities against Stated intentions. The report should include various pre-specified performance indicators. *Much of*

*the management information required for monitoring and assessing performance at this departmental level would be the same as that required for the resource allocation process for the institution. A few universities are already considering this.*' (CVCP, 1985b, §3.24, emphasis added).

In line with the traditional structure of PIs (Cuenin, 1987), the national data study advocated that universities could start reporting research efficiency by constructing an input-output relationship based on staff working hours and publications. There had been previous debate over the usefulness of these quantitative indicators (Sizer, 1981) precisely because their articulation of quality in cost terms was seen as problematic for thinking about qualitative processes. Inputs and outputs in research and teaching activities are hardly material and quantifiable.

Still, one of the possible reasons why cost-focused indicators persisted in projections of future quality criteria is that the committees' approach was backed by a stream of researchers, including Jarratt himself (Jarratt, 1986; Moodie, 1986; Berrill et al., 1983; Becher et al.,1977), and especially John Sizer, who later entered the UGC (see Sizer,1981,1987a,1987b,1988,1989). From their perspective, politicians' determination to dismantle the public sector was a further reason for a fast turn to managerial practices. This became increasingly evident by 1987, when the UGC and the CVCP committees continued to harmoniously build on their initial work (CVCP,1985b). Following the publication of the Jarratt Report and the National data study, overlapping networks often rotating around the CVCP devised a series of PIs (Cave et al., 1997). *As I argue, the National Data Study and the Jarratt Report introduced the first performance indicators for quality assessment in British universities.* Whilst the CVCP recognized the differences between PIs and quality evaluation ones –the latter more oriented at analysing research specifically– and the fact that PIs could have distorting effects, these were still considered an important step to make. With the support of members which had already worked on PIs in universities, such as John Sizer, the CVCP and the UGC began constructing them. As has been shown in detail, from 1985 onwards the development of internal and external PIs moved parallelly (Cave et al, 1997), and, crucially, from *within* the university sector. This occurred largely because the PIs emerged from the joint work of the two committees, with the CVCP concerned with developing PIs for internal management and the UGC for sector-wide evaluation. Although it has been argued that the push for PIs came strictly from the 'external' political sphere, the intensive work of the CVCP/UGC group shows that the pressures coming from within the sector were significant. These pressures carried an extra weight compared to state ones, because they were very specific in articulating concrete new practices to be implemented. To an extent, and as the next section discusses, the CVCP/UGC working group

gradually became extraneous to institutional level discordances around evaluation. The group's reports throughout the years gradually evolved into proper demands on universities: that they standardize procedures, measure, report and adhere. This perspective on the CVCP's advocacy of a managerial turn in universities allows to consider how evaluation developed instrumentally to the institutionalisation of the committee; rather than functionally as an adaptation to demands of the neoliberal polity. In pushing for institutional change through its proposed indicators, the CVCP tried to exert significant influence over the question 'how should universities be governed?'. The diffusion of proposed practices could, in the medium and long run, bring to consensus over executive governance. What has been traced so far is an analysis of the rising managerial class within higher education, and what follows is an insight of how it constructed evaluation benchmarks through a specific managerial framework which had its roots in PPBS. A reason for considering this argument can be found in the indicators themselves. As has been already suggested (Cave et al, 1997) and as will be further elaborated, the PIs published in 1986 were largely the result of the use of data which was readily available to the UGC and the CVCP. When keeping in mind the recommendations of the Jarratt Report and the National Data Study, it is evident that the CVCP/UGC working group fused the internal management data provided by the CVCP and sector data provided by the UGC. This is largely because *there was no proper and specific 'external' criterion imposed nor requested by politicians*, who left the committee with wide margins of action, in its recall to "value for money." Importantly, it is from a joint CVCP/UGC working group (CVCP/UGC, 1986, 1987a, 1987b, 1988) that the first reflections on research quality assessment were made, which borrowed heavily from the cost-focus of the CVCP's earlier proposals (CVCP 1985a,1985b).This collaboration created a simultaneous development of internal and external PIs between 1985 and 1989. These were the first PIs to be developed and then implemented in the sector, a transformation which can be explained through two key factors. Firstly, the CVCP/UGC group's membership and its interests, which increasingly turned the CVCP/UGC group into a sector's key point of reference for evaluation matters. Secondly, the management information which it readily had available, which raised heated criticisms on the interpretation which was made of research quality.

The restricted group emerged from specific difficulties which members of the two committees faced. As mentioned, the CVCP's managerial shift had support from vice-chancellors but still faced evident challenges at steering universities from within and see planning implemented as proposed in the Jarratt report. Combined with the UGC's difficulties at regaining a solid reputation following its blown attempt at selective funding, this made the creation and the CVCP/UGC working group an attempt at legitimating both committees. The working group was first chaired by CVCP member Mark Richmond (see appendix E for the members of the 1986 joint

working group), and then by E.S. Page. A key UGC representative was John Sizer, key advocate of HE PIs (1979,1981). The rest of the members did not have a background in strictly educational matters, and the CVCP opted for a "technical" membership made up of planning officers, a registrar, and a vice-chancellor (see appendix F for the full list of the working group). Interestingly, Sussex continued to have a special place in university management, with its planning officer in the committee. Their insights in the two assessment 'spheres'– internal and external – were thus combined through the membership. This group distanced itself from government and in particular from the DES. Aware that the CVCP/UGC group was devising efficiency criteria following Keith Joseph's concordat with the committee (Shattock, 2012, Kogan & Kogan, 1983), the DES in fact kept a relatively loose grip over the university sector. Besides promising future increases in funds, the famous 'concordat' between Keith Joseph and the CVCP[81] (Kogan & Kogan, 1983; Shattock, 2012;) had established a relationship whereby if universities could prove to be working on efficiency, the State would maintain a limited interference in their resource management. *These aspects enabled Thatcher's government to keep track of developments in efficiency indicators without having to be directly concerned in their construction.* When working towards the use of the first efficiency criteria in the mid-1980s, the State was mostly concerned with '[…] withdrawing from the murky pain of overwhelming detail, the better to take refuge in the clear commanding heights of effective strategic 'profiling'' (Neave, 1988, p.12). As such, the CVCP's lists of performance indicators acted as signaling devices to the DES, communicating that universities were working on efficiency, providing sufficient evidence of developments toward quality evaluation as far as the DES was concerned.

The data readily available to the CVCP/UGC group as of 1986 mostly concerned university costs, partially as a consequence of Jarratt's insistence on financial management (Sizer, 1987b). Additionally, the CVCP had gathered an array of internal management information through the Academic Audit Unit. This unit represented an internal auditing mechanism, where AAU members– in collaboration with external consultants – monitored the quality of universities' management and quality-assurance systems by checking for the use of efficiency-oriented cost-indicators (Williams,1992). Upon the establishment of the AAU, it was believed that gathering internal cost and efficiency information would have, in a sort of performative way, further stimulated its use, demand and gathering. The efforts at devising performance indicators so rapidly thus aimed to speed up existing procedures wanted by the working group.

---

[81] 'Sir Keith Joseph told the House of Commons that the Government was to consider with the UGC and CVCP some further financial provision for the universities in 1987-88 and the following years but that the 'the government's willingness to make such provision depended crucially on evidence of real progress in implementing and building upon the changes that are needed (HC Debates Vol 89, Col 179' as cited in Cave et al., 1997, p.4)

The CVCP/UGC's first statement was published in 1986 (CVCP/UGC, 1986) and contained preliminary financial management (Cave et al., 1997), mostly universities' cost-efficiency in service provision (library, careers, premises etc.). The indicators calculated costs per FTE student. With regards to indicators of research performance, the situation was more complex due to the issue of conveying quality (and qualitative judgments) through quantitative benchmarks which had initially been designed to judge financial – budgetary matters. Still, the first statement presented some research and non-research PIs. The research ones were (CVCP/UGC, 1986):

- Cost per FTE student
- **Research income per FTE staff**
- Equipment costs
- Contribution to postgraduate and professional training
- Costs per graduate
- Equipment costs per FT academic staff
- Submission rates for research degrees per FT postgraduate students (time between registration and submission)
- **Number of research and sponsored students**
- Citations (number of times that academic staff is cited in work by others)
- Peer review
- Membership of research councils (n. of department members who are on a main or sub-committee of a Research council)

As Cave et al (1997) have noted, the group's second statement reproposed similar indicators, but the ones eventually considered for immediate implementation were research income and number research students. The working group's second statement (CVCP/UGC, 1987a) solidified a cost-focused framework which defined performance as the use of resources at a micro-level rather than modifying it in light of research on bibliometric and research-specific indicators. Indeed, the working group postponed their integration in the list of PIs (CVCP/UGC, 1987a, § 4.2) multiple times until 1995. The key concern was thus to elaborate on the relationship between costs – the "inputs" – and numbers of researchers –the "outputs"–. New research-specific statistics were gathered around this, such that none of those published in 1987 (CVCP/UGC, 1987b) used citation numbers, numbers of publications or other (see appendix G for the full list of PIs). The working group was aware that '[…] the varying structures of universities make it difficult to obtain full

consistency in the definition of cost centres.' (CVCP/UGC, 1987a, §6.6.) and decided to analyse cost-centres' expenditure only in relation to resources coming from 'general funds' (i.e., State funds) as opposed to evaluating how universities spent funds coming from research grants or contracts ('special funds'). This made the search for comprehensive research quality indicators more general and reflected the fact that the working group's key priority was to gather data on expenditure. True, the working group had not been set up with the specific aim of gathering research quality indicators, which in part explains the scarce attention to more elaborated ones. Yet, it seemed to pose itself this task, with '[…] the activities of the working group […] directed primarily to the major functions of teaching and research' (CVCP/UGC, 1986,§3, §19), suggesting that its existent data was itself utilizable for research (and teaching) performance assessment. The committee was aware that this position on quality was to an extent reductive and in a way also unrealistic, yet still gave a go at elaborating cost-oriented research quality indicators. In so doing, it conflated cost-performance with the quality of research. The overwhelming attention to cost-efficiency as a valid criterion for universities' quality did not seem to present, in fact, an issue for the committee. Page made it clear that he was concerned with devising indicators of *productivity*. In a simplistic vein, he made little distinction between quality and cost indicators, for in his view PIs were first and foremost '[…] all those numerical data which are useful *in managing a university*, assessing its operations, costs, and performance' (Page, 1987 as cited in Cave et al., 1997, p.52). Who, then, could begin to make use of the information being published? Though invoking future research for comparative purposes, the working group proposed that its PIs be used by national bodies as a guidance on funding allocations (CVCP/UGC, 1986, §20). This was an ambitious step which sought to enhance its approach, made more ambitious by the suggestion that its rather restricted list of PIs could also be used for internal management

'[…] the working group envisages the development of some indicators primarily for use within institutions, while others may have particular value for inter-university analysis; and yet more may have their main use for comparisons across the binary line' (CVCP/UGC, 1986,§13).

This position suggested that PIs were flexible and could interchangeably be used by planning officers for monitoring staff productivity , by departments and cost-centres for monitoring resource use, and by policymakers for institutional comparisons. Seen from VCs' perspective, the diffusion of indicators both within and outside universities could facilitate managerial control over resources and staff. Constructing indicators and emphasizing their usefulness was an important step to make in the attempt to diffuse them for validating the approach lying behind the indicators.

**Cost-Centres**

Cost-centres, which have so far only been mentioned, are another key innovation of the CVCP/UGC working group, one which supported its construction of PIs. They mark indeed an important transformation in how university activities would be assessed for years to come. They allowed new developments on data gathering, on one side, and on the group's elaboration of information in terms of rankings, on the other. The usefulness of PIs in fact did not just depend on their diffusion, but also on their ability to convey very specific information on expenditure instead of extrapolating only institution-wide data, for cost-assessment could be better made with more detailed information. As discussed, the working group's key innovation lie in the construction of micro-level information geared at supporting budgeting and planning practices advocated in the Jarratt report. In proposing that information on activities be collected by cost-centres, the group stepped in and tried to influence the traditional way in which disciplines were categorised. Broadly speaking, cost-centres were treated as a synonym of departments, which traditionally separated different disciplines and grouped similar ones. At the same time, they represent a new way for categorising departments. The working group's first and second statements (CVCP&UGC, 1986, 1987b) listed 37 cost-centres (see appendix H) which did indeed reflect disciplinary divisions but not more specific sub-divisions within departments. In fact, through cost-centres the committee grouped research information according to a unit of analysis which was smaller than the department (for example, social sciences), but bigger than the specific sub areas within disciplines (for example economic sociology). In the words of the working group:

'[…] A cost centre brings together aspects of the university's work having a common character and cost pattern, to provide an entity of which comparisons between institutions can reasonably be based. In particular, all academic teaching and research will be assigned to one of 37 academic departmental cost centres, to which student load and academic staff will also be assigned.' (CVCP/UGC, 1987a,§6.5)

Thus, this categorisation of cost-centres was mad possible indeed by defining them from the onset as units with 'similar expenditure patterns' and less attention to how disciplinary (dis)similarities played a part in the qualitative dimension of cost-centres. Cost-centres were therefore constructed to convey, in simple and comparable terms, different disciplines' costs - the use of resources for carrying out research and teaching –. The way that the CVCP/UGC working group categorised costs/expenditures (e.g., research income per FTE academic staff) reflected the information

required by cost centres. This would provide input information which could be related to other data for constructing indicators of performance. Additionally, each type of information (such as research income per FTE academic staff) could -individually or in the form of a PI - be compared between cost-centres, making their structure therefore key for conveying performance and efficiency. To give an example, research income per FTE academic staff could be related with expenditure per FTE academic staff to convey its research productivity in "general engineering" (cost-centre 19) at university A. Similarly, research income per FTE staff in general engineering could be compared with "pharmacy's" (cost-centre 6) at university A to assess which of the two cost-centres received more research income and thus, potentially, which cost-centre was more deserving of research funds. The same information could be compared for a single cost-centre in two different universities: so, research income per FTE staff in pharmacy at university A could be compared with that a university B, conveying in this way differences of performance in the same disciplines/cost-centre. For the working group, high cost-effectiveness translated into the productivity in achieving university "outputs": full time students and researchers. This did facilitate the use of indicators for institutional management purposes. Further, it had a broader implication at the sector-level. From the UGC's perspective, the cost-centre structure was coherent with the view that institutions' average costs were a suitable measure for comparing efficiency and quality *between* universities (Cave et al, 1997). To this end, cost-centre units of analysis were indeed of use. Each cost centre's average costs could be calculated and used to compare universities to each other on the basis of their expenditure performance. In other words, the average cost of each cost-centre could also be compared. Finally, a national average aggregate of cost-centres could allow to evaluate how single universities performed either above or below it, providing the UGC with a much more detailed picture of unit costs.[82] Gathering cost-information at a cost-centre level as the Jarratt Report had suggested thus opened a wide array of evaluation paths which rested on managerial decisions to be made at the top levels of the institution or of the sector on how to elaborate and use the data being gathered..

As research and teaching activities were analysed from this managerial imprint, they were increasingly defined as *cost-incurring* activities rather than *formative* ones. Thus, research and teaching *performance* assessment predominated over *quality* assessment. As Cave et al (1997) suggest, '[…] the working group's approach implied […] that lower costs of higher education were desirable and there was a danger of ignoring PIs concerned with the quality of the output' (p.53) The two statements discussed thus represent a go at a re-articulation of quality in terms of

---

[82] According to Cave et al (1997), unit costs were a predominant factor for policymaking '[…] if the funding body is concerned only with minimising unit costs, it may adopt the same policy in either case – of concentrating resources on low unit cost institutions. If it is concerned with other objectives as well – for example, ensuring a regional balance […] – it will be more inclined to concern itself with the explanation of higher unit costs.' (p.122).

performance, for the cost-PIs presented *were defined as useful for considering and calculating research and teaching quality* (CVCP&UGC, 1986, 1987). Potentially, cost-considerations could have developed – had they not been heavily criticised by the academic community – as carrying equal weighting in research quality evaluation as other indicators such as citation indexes. It is for these reasons that the PIs presented in 1986 and 1987 reflect an exercise of consensus-building as well as instruments of a broader political struggle to slowly steer universities into a managerial framing of quality where VCs and UGC members would ideally appropriate the indicators and control their use for resource allocation. The CVCP's previous advocacy (CVCP,1985a) for institutional planning was indeed recast in the CVCP/UGC's statement that'[…] sound management information is necessary for the effective running and financing of any large and complex organization' (CVCP/UGC, 1986,§3). Cost-focused PIs were treated as objective measures of performance to be experimented in universities. However, experiments were often objected for a wide share of the academic body felt that 'the PIS were open to the criticism that […] the way in which the data are to be interpreted ought to be obvious' (Cave et al, 1997, p.53). This lack of transparency over the implications of PIs was no easily negligible detail, which as we will see did in fact pose significant new challenges to the working group. As the group limited its definition of the use of PIs to management purposes (CVCP/UGC 1986), it did not help make academic community feel that their concerns on PIs' objectivity and usefulness were being addressed and understood, sparking the first widespread debates on research PIs (Phillimore, 1989, Bence and Oppenheim, 2005, Johnes & Taylor, 1992). Indeed –and rather paradoxically when considering current objections to quantitative evaluation methods– the academic community saw bibliometric and research-specific indicators as more objective than the working group's PIs. Academics felt 'safer' being assessed through quantitative yet research-tailored and research-focused indicators than through managerial ones. (Cave et al., 1997). This interestingly shows that quantification of research quality *per se* in abstract terms may not always be the overriding issue when discussing research evaluation. What is more pressing and more worth being debated in detail is the content of quantifications and the agents proposing such content. These aspects reveal a lot about the way that quantitative criteria develop for they unveil information about the contextual factors behind a push for quantification. As of the mid-1980s in England, quantification in general terms was a secondary issue to the *content* of the then existing indicators. Bibliometric criteria such as numbers of citations, or peer-review judgments based on reputation or journal publications appeared much fairer and objective to researchers than cost-efficiency PIs (Phillimore, 1989). They did not reject quantification insofar as it was geared at research concerns. An emerging obstacle for the academic community in the late 1980s was thus that of being able to demonstrate that there was, a methodological flaw in the CVCP/UGC's PIs., and that more debate

was needed first and foremost, on the type of information that universities should provide.

### 4.2.3. The first Research Assessment Exercise

The construction of these first PIs played a role in shaping the weight of cost-oriented research performance within broader set of research quality criteria developed in the mid and late 1980s. In fact, the UGC eventually used a selection of the PIs developed with the CVCP in 1986 (CVCP&UGC,1986) when it formulated the first Research Assessment Exercise. For this reason, the work of the group can be said to have exerted significant influence on the first national evaluation exercise for English academics, and– though with very problematic aspects – also in forming the grounding basis of successive RAEs' criteria (Cave et al., 1997). The continuity between the group's PIs and criteria of the 1986 RAE cannot be easily analysed in first-hand, for the criteria used were not publicly published and were simply communicated through UGC's circular letters to universities (Phillimore, 1989) , a very debatable choice which re-sparked criticisms over the UGC's competence. General summaries, however, were provided by journalists and academics of the time (Cave et al., 1997). Broadly, the 1986 RAE has been described as having proceeded in the following way:

> '[…] cost centres within universities were evaluated by UGC subcommittees of "experts" in each discipline, who then classified the research performance of the component departments of each cost centre as outstanding, above average, average, or below average' (Phillimore, 1989, p.260)

The UGC's evaluation meshed objective and subjective types of judgments. Criticisms, however, highlighted that the exercise was very problematic because different types of publications were not weighted, making the judgments appear as abstracted generalisations of significant differences in workloads (for example between writing a monograph and a journal article). Overall, 5 key aspects were evaluated for each cost centre (Phillimore, 1989, p. 260) 2 research aspects were evaluated subjectively:

- – Publications: 5 publications per cost-centre (from 1981-1986 period)
- – Research performance assessed through peer review.

Whereas three aspects were evaluated objectively (per cost-centre):

- – Number of research grants, studentships and "new blood" lectureships
- – Income from industry
- – Fellowships and prizes to faculty.

The UGC acknowledged some issues of its 1986 RAE, announcing that it would correct them in future RAEs by integrating more elaborated peer-review and bibliometric indicators. Academics were however well aware of the risks of the cost-centre legacy and tried to show them. On a methodological level, negative reactions to the first RAE emphasized that '[…] no allowance appeared to be made for the difference in the size of departments or cost centres', such that some suggested how, paradoxically, a "per capita performance" measurement would have been more reflective than assessing cost centres' total output in terms of research grants, for example (Phillimore, 1989). A very critical issue was also that the sub-groups of the UGC used different scales for judgments, with some using a 3-point scale and others a 4- or 5-point scale (Phillimore, 1989, p. 262). Academics criticised the lack of a coherent, common definition of research performance based on the issue of using cost-centres as a unit of analysis. An academic and research-based assessment of research quality was, according to critics, being outweighed by quantitative aspects such as research grants, which '[…] favoured high-cost areas of research' (Phillimore, 1989, p. 262). These income-focused aspects, meant to reflect the quality of research, were based on the idea that good research would attract more money, a reasoning that could apply in broad terms yet requires very attentive elaboration of the sizes, lifetimes and influence of different departments and disciplines in the sector. Some cost-centres, for example, may attract more resources than other simply thanks to their larger size and their wider networks, or through specific lobbying. The working groups' PIs were thus appearing more vulnerable than expected. Although the 1986 PI of "research income was broken down into three aspects (the objective indicators), this did not in fact suffice to achieve academic consensus over evaluation criteria. The work of the CVCP/UGC group halted in 1988 upon the demise of the UGC. The CVCP, however, continued it work with national funding bodies, collaborating with the UFC between 1989 and 1992. Specifically, the two continued to devise PIs and the CVCP initiated a survey on research publications in 1991 (JPIWG, 1993).

## 4.3. A contested sector-wide evaluation framework

If the first RAE was relatively controlled by the CVCP and cost-biased indicators were quite central, this began to change as the context of the 1990s changed and the State began to act more decisively in evaluation matters. Although it now appears natural to think of research evaluation as regular, national exercise, this was not the case at the end of the 1980s. It was not yet clear how the selectivity experiments of 1986 and 1989, if at all, would develop into constant evaluation able to drastically reshape universities' ability to sustain research and teaching funding. Furthermore, it was hoped that future assessments would remedy previous mistakes and not develop into a top-down

punitive exercise which reduced quality to a case of cost-efficiency coupled with wealth-generation. In part, these doubts were addressed by the State, which in 1990 elaborated its vision of a future structure and functioning of the sector. The 1991 White paper "Higher Education, A New Framework" (DES, 1991) confirmed the willingness to continue national research assessment exercises in the future. Successive frameworks, however, would place more emphasis on the relationship between research quality assessment and research funding. Though the abolishment of the binary line intended to somehow unify universities and polytechnics, this goal created new tensions between their institutional representatives, because it became evident that evaluation could turn into a source of competition between universities and ex polytechnics. Between 1991 and 1997, contestations between the two concerned both quality audit and research quality evaluation criteria. Though this research predominantly focuses on the latter, a brief elaboration of the different positions taken on quality audit following 1991 are important for understanding successes and obstacles to the CVCP's work on evaluation. As is later discussed, a key issue for the committee was managing to institutionalise university-tailored auditing procedures in a very different sector. If on one hand the White paper showed the State's intention to continue making use of representatives' views, on the other it also emphasized that this would occur through an ideally collaborative dynamic between former polytechnic representatives and university ones. For the CVCP, influencing future decisions on evaluation would not be as simple as had occurred through collaboration with the UGC in the mid-1980s. Developments in evaluation methods are therefore to be considered as ruptures than as a progressive evolution of evaluation towards constant improvement, they were importantly shaped by the forces at play and how these interacted. If new criteria emerging in the 1990s can be considered an improvement of those of the 1980s, from an agency-centred approach focused on changing power relationships this view is rebalanced in no marginal way. Modifications to previous quality criteria in fact reflect the new configurations of power emerging with the establishment of a new funding council, the HEFCE. The content of the criteria, in other words, depends importantly on the groups of agents constructing it.

Still concerned with value for money, the State announced that it would abolish the binary line which had historically divided polytechnics and universities into vocational and academic-research oriented institutions, respectively. The binary line had also guaranteed different types of political presence within the two sectors, with polytechnics being largely scrutinized and universities being largely autonomous primarily thanks to their degree-awarding powers and to the presence of the UGC, which mediated university interests with the State (Shattock, 2012). The proposals made in 1991 (DES, 1991) and then concretized in 1992 in the "Further and Higher Education Act" (DfE, 1992), gave the British State a chance to shift the burden of research funding selectivity to the newly proposed funding councils – later denoted "seats of power" (Neave, 1988) –

acting in a radically new quality assessment framework. By abolishing the binary line, the White paper also abolished the Universities Funding Council (UFC) and the Polytechnics and Colleges Funding Council (PCFC), both set up following the Education Reform act in 1988. The idea was to create a single funding council which would regulate all universities in the sector (later commonly denoted as "pre-92" and "post-92" universities to distinguish between universities and incorporated polytechnics). As far as England is concerned, the establishment of the HEFCE in 1992 meant that state grants for research and teaching would be allocated by this new body. If the White paper simplified the funding procedures for the State by establishing regional rather than sectorial funding councils (DES, 1991), quality assurance and the ways in which funding councils would make quality judgments for funding were articulated in a vaguer way. This is because the State sought to maintain a degree of control through the new councils and agencies (Neave, 1988, 1998), yet it was also aware of the criticisms that could emerge if decided to establish a central quality assurance structure responsible for all aspects of assessment. It therefore made distinctions between quality *audit* and quality *assessment*: HEIs' internal quality audit mechanisms for monitoring internal resource use teaching quality would continue to be carried out at the institutional level. It stated that '[…] any doubts about the effectiveness of self-regulation are more than offset by the self-interest which institutions will have in demonstrating that internal quality controls continue to be rigorous' (DES, 1991, p. 26). Preferring not to take the riskier path of interfering with autonomous internal management the Secretary of State for Education Carlisle then articulated quality assurance as the result of such internal processes, so that the '[…] the quality of what is actually provided' (DES, 1991, p.28) would be assessed by external bodies. Quality assessment thus rested on a complex articulation of how these two levels of internal quality assurance and external judgment of the outputs eventually achieved would be coordinated and communicate different types of judgments. In terms of research quality assessment methos, the state presented a simplistic view of evaluation, which was supposedly a key pillar of the White paper.

The White paper announced that it would maintain the funding structure as a dual support system, where research funds were provided by both the State (through the funding councils) and the Research councils. However, it vaguely articulated how it would act in new provisions on external research quality assessment. The scarce references made to quality criteria were still preliminary, research assessment would thus develop '[…] through quantifiable outcomes' and that '[…] performance indicators […] will have a role to play in the future.' (DES, 1991, p.29). It appeared that the State was inclined to maintain the RAE structure – and criteria – relatively unchanged. It did not see this as its own concern as much as that of the funding councils. The document did not really provide additional guidance on criteria to the existing set and confirmed that "general" research would continue to be funded in a "single channel." Major novelties in

research evaluation concerned polytechnics more than universities, for the latter were already receiving funding through the dual support system. As I argue, the State was not concerned with the methods of research quality criteria at this stage, and the 1992 Further and Higher Education Act did little to develop it. The priority of the two key papers (DES,1991; DfE, 1992) with regards to research selectivity was just to *delegate this function* to the Funding Councils and their "Quality Assessment Committees" which would'[…] secure that provision is made for assessing the quality of education provided in institutions whose activities they provide […] financial support for' (DfE,1992, §70 comma 1a). With this in mind, institutions' reactions to the 1991 White paper ought to be distinguished between those concerning the changes to quality audit, and those concerning quality assessment (i.e., selectivity) criteria. Later academic studies have often generalised – culpable the fact that eventually audit and assessment were merged in the QAA in 1997(Brown, 2004) – the stark difference between quality audit and quality assessment into one single category (Palfreyman & Tapper,2014; McGettigan,2013; Slaughter & Rhoades, 2004), generalising transformations occurring into a single process of managerialisation or neoliberalisation of HE. Instead, capturing the State's managerial imprint during its first reforms on research quality assessment methods needs to be approached with greater historical accuracy. The changes brought about in 1991 and 1992 were relevant for evaluation intended in broad terms, but rather marginal in terms research quality assessment. This importantly speaks about its specific role in the latter. Delegation on the aspect of methodologies to the HEFCE turned responsibilities for evaluation onto a variety of "interest groups" that were not at all strictly from the political-bureaucratic sphere. When the HEFCE was set up, its research quality direction was still to be decided. Whilst the merging of and quality may present a centralizing act on behalf of the State, the same cannot be said research quality assessment responsibilities. If on one hand members of HEFCE's quality assessment units– as those of other regions' FC – would be mostly recruited from HMI staff employed in education (§ 82), others were to be recruited from the academic world. This reveals that an interaction –wanted by the State –between political-bureaucratic and university occurred during the construction of evaluation for the new sector Furthermore, the White paper suggested that quality assessment units could set up committees of institutional representatives who would '[….] advise them on the operations of the units' (DES, 1991, p.29). This concretely meant that university Vice-Chancellors could continue to play a role in quality assessment as they had done with the 1986 and 1989 RAE. Still, polytechnics' probable development into universities meant sharing the cake with more members coming from different backgrounds and possibly diverging interests to those of the pre-92 committee. This bargaining process soon set off as university vice-chancellors influence over evaluation was threatened by two key factors.

On one hand the State would have a greater role. Whilst the legislator acknowledged the funding council's statutory independence, it explicitly granted itself more powers in influencing the funding council's operation and steer it when considered necessary. This effectively went in the direction of turning the traditional symbol of the funding council as a buffer body into a formal feature. Ten years into its establishment, the HEFCE was widely criticised by the academic community for having turned into a government arm (Taggart, 2004) operating without questioning the DES's (or DfES's) funding framework and financial memorandum. The fact that members of HEFCE's board of council were often directly appointed by the Secretary of State for education also validated views of political centralization over research evaluation. Historically, the first annual reports of the HEFCE often seconded government policy and worked to achieve efficiency savings (Brown, 2004) as was being politically demanded.

Secondly, the funding council had new leverage too. As it became *de facto* responsible for funding selectivity, it could decide to limit the use of institutional representatives' advice and opt for alternative methodologies to those of previous RAEs. From the perspective of the CVCP, the HEFCE's leverage implicated a possible end to its active engagement in devising research performance indicators, as well as the process of institutionalisation set off since 1984 and continued with the UGC and UFC. This also threatened its managerial aspirations in universities. The CVCP was well aware that funding councils would have to integrate a much larger bulk and diversity of institutions when making quality considerations.

Thus, interweaving close ties with the HEFCE may not have sufficed for advancing its own criteria, as the committee needed to also bargain over the content of research performance criteria with the directors of ex-polytechnics, who had their own methods and quality priorities to advance. In these aspects, the White paper could end the CVCP's determined managerial movement which had benefited from an aura of elitism. This latter point allows to better conceptualise power struggles in the new university context of the 1980s. Institutional representatives would indeed seek to obtain a degree of control over the HEFCE. The potential loss of control over evaluation concerned mostly the CVCP, which as we have seen was well central in previous RAEs. Maintaining a hold over research quality assessment criteria was thus crucial for avoiding that the HEFCE's establishment develop into a fully top-down relationship where the CVCP's previous efforts could end in vain. By 1992 the governance of higher education appeared *compartmentalised.* As the State increasingly *juridified* higher education and assigned responsibilities for monitoring the sector's functions (teaching, research, management) to an array of intermediary bodies, it created new *spaces of contestation* over how different aspects of the sector could be governed. In this way, the governance of quality assessment was developing into a "broad operating space" (Neave, 1998) marked by negotiations, a grey area in which quality evaluation responsibilities were

never fully centralized by political control. The elaboration of performance indicators was to an extent more "centralized" in the hands of the UGC/CVCP previously to 1991 (UFC, 1991). The establishment of the HEFCE changed this, as resource allocation decisions would be made through the advice of both a new quality audit body (the proposed HEQC) and a research quality committee. This was a wider battleground compared to the narrower CVCP/UGC working group. The end of the binary line opened the doors to a struggle between traditions and ideas on higher education, whish, as we will see, met, clashed, and changed in time. Two key moments of contestation in English Higher education following the abolishment of the binary line can be identified. They represent concrete struggles which the CVCP and its work on indicators faced. The contestations between 1991 and 1997 involved both quality audit and research quality assessment criteria and procedures. Though the focus of this research is on the latter, the CVCP's actions towards quality audit following 1991 are insightful for they reveal that the committee tried to keep internal management criteria that it had developed since 1985 intact in the context of the emerging Higher Education Quality Council (HEQC), the audit body proposed. This effort had significant effects on universities' bureaucratization, it helps to understand how these were now filled with multiple demands, criteria, and procedures in policy, assessment, and audit. This overwhelming inflow of changes allows to think about how they turned into the victims of a struggle over the control of quality audit. A single audit body meant that polytechnic directors and university vice-chancellors could work together to devise a system of internal auditing which could suit both types of institutions. Had they managed in this intent, they could have created the first-ever institutionally-led national quality audit body in British higher education.

For polytechnics, the possibility of turning into universities offered important advantages. Firstly, polytechnics could decide to rearticulate their functions and objectives. Along with a university status, in fact, they could foresee access to more research funds, given that they traditionally received less than universities. As universities, polytechnics could also get rid of the external validation requirements of their courses and degrees, a function historically carried out historically by the CNAA. This could importantly empower them. Unlike universities, polytechnics had been used to external quality assessment because this was a perquisite of their functioning. The supervision of her Majesty's inspectorate on teaching and learning had made polytechnics very familiar with the notion of accountability to the "external environment" that universities were instead still quite unfamiliar with. Eventually, 31 polytechnics acquired university status (Brown, 2004).

Universities were the most threatened institutions. On one hand the abolishment of the binary line meant that a university head may have had the chance to 'eat up' a polytechnic by supervising its transition, in the process take over its decision-making responsibilities (Education

Reform Act,1988). For the CVCP, the proposal of a single funding council was a downgrading of universities' prestige (Brown, 2004). Evaluating the teaching and research in polytechnics and that in universities through the same criteria translated into a sort of underestimation of universities, particularly the oldest ,because they considered themselves as prestigious precisely because of their high-quality, non-vocational research which drove teaching (Brown, 2004). Competing for research funds with polytechnics was a de-legitimation of a *traditiona*l prestige, an aspect which the CVCP openly opposed from the onset. Further, more tensions were also fuelled by the fact that the Committee of Directors of Polytechnics (CDP) and the CVCP were treated differently by the government officials. The latter '[...] was addressed annually at its conference by the Secretary of State, while the CDP had to be satisfied with a more junior minister.' (Shattock, 2012, p. 69). All these factors made the university representatives feel entitled to remain detached from polytechnics when the binary line was abolished.

### Contestations over audit

These different sentiments were concretely exposed when universities and polytechnics were asked to collaborate at the HEQC. For government, a unitary auditing body was much more convenient, as it could simplify the decision-making process of the funding councils (DES, 1991) and also rationalise resources in the long run. It therefore pushed the two to work together on internal auditing indicators, but this was no simple task. Designed to support the HEFCE, the HEQC was intended to act as an institutionally-devised yet external auditor of the sector, feeding its conclusions to HEFCE which would consider them if it saw necessary. For the polytechnics represented by the CDP the establishment of the HEQC could put the CNAA behind and allow to actively participate in the construction of standards which it saw more suitable. For the CVCP, the rise of the HEQC was more complex to balance out. It was a success because the fact that the HEQC largely emulated the functions and procedures of the AAU (Brown, 2004) brought prestige to the CVCP, the AAU's founding parent[83]. Precisely for this reason, however, the HEQC also presented an obstacle, as the CDP  was keen and legitimated to propose its modifications to the AAU model. For the CVCP it worked well the way it was, not least because it empowered VCs in

---

[83] The origins of the AAU trace back to the academic standards group in 1983, first chaired by Reynolds, author of the famous Reynolds report (1986) which had set out key guidelines of internal quality audit through a code of practice for monitoring standards. The group worked on standards rather informally until 1988 when the newly established UFC pressured universities to act more decisively on internal auditing. AAU reports sought to make universities vigilant. For example, quality reports were made available to universities' students, who would in their way assess the strengths and weaknesses of their university: '...the intention was that students should be influenced by knowledge of quality assurance machinery and draw towards those HEIs which did well in the Audit (CVCP/AAU, 1992). The AAU's formal establishment in 1990 is more recalled than its longer lineage, marked by the CVCP. It is for this reason worth presenting Peter William's -first chair of the AAU -  observation pointing once more to the active engagement of the committee in devising quality criteria: '[…]the [Reynolds] report can fairly be said to have started the widespread effective discussion about quality and standards in British universities.' (Williams, 1992, p.143)

universities. The management board of the AAU was chaired by the CVCP secretary and consisted of a majority of VCs (eight, with other four independent members from the non-university sectors). Through the subscriptions that it charged to universities (on a student number basis), the CVCP had gathered important sums for the AAU's operation and for itself. That of 1991-2, for example, totalled £450,000, excluding costs incurred by universities for setting up and maintaining quality assurance systems (Williams, 1992). These aspects strengthened the CVCP's view that universities were already well-equipped with internal quality auditing, leading it to act confrontationally with its polytechnic counterparts. As Perellon's research (2001) has highlighted, '[...] some members of the CVCP [...] thought the HEQC ought to do exactly what CVCP told it to do and nothing else' (interviewee, 2.V.e.9, as cited in Perellon, 2001, p. 111). This struggle over the HEQC's procedures and modalities led to a proper class between the sector's two institutional representative committees (Brown,2004), turning the construction of university-devised audit criteria into a debate over their control as much as that over their content. The resulting hybridity of the HEQC has indeed been defined as fatal for universities (Perellon, 2001). The struggle over control of auditing led them to miss a concrete chance at having a long-lasting ,institutionally-led and credible quality body in which the greater involvement of academics could have stimulated qualitative rather than political debates on quality audit; and resulted in a less criticised framework. The HEQC's loss of credibility okayed an important role in its demise. Discussions continued over the establishment of the HEQC's successor, the QAA (Brown, 2004) from as early as 1995, with the CVCP continuing to favour the AAU and self-auditing procedures. The committee, at one point, made a new and hazardous proposal, which seen in historical perspective makes much more sense. It suggested that quality assessment be incorporated by institutionally-led auditing (Perellon, 2001; Brown, 2004). In sustaining this, the CVCP once more raised the primacy of cost--efficiency – to be achieved through institutional management – in research quality assessment procedures. This undermined the stark differences between the two types of evaluation. After rejecting the CVCP's proposal for the QAA (whose functions remained undefined until late 1996 (Brown, 2004) – the HEFCE stepped in and pressed the secretary of State to speak out to maintain quality assessment and audit separate (Perellon, 2001). Specifically, the HEFCE warned that merging two distinct quality evaluation approaches (one based on judging quality mechanisms in place, the other on research quality) could distort in the meaning of quality itself: institutions that performed well in quality audits could also perform poorly at quality assessment of research and teaching; or vice versa: high quality research can be produced in a poorly performing managerial university (Perellon, 2001). When the QAA was established, HEFCE retained its quality assessment responsibilities.

**Contestations over quality**

The debate over the creation of a single quality audit body intersected with parallel contestations over research quality assessment criteria. The HEFCE's first RAE was the one carried out in 1996. Up until then, the CVCP maintained an important degree of influence over the RAEs. If the CVCP/UGC working group had led the 1986 and 1989 RAEs, it is also true that the UFC consulted extensively with the CVCP for guidance on the 1992 RAE. Although the establishment of the UFC was intended to represent a rupture, the UGC still left a legacy following its closing down in 1989. Swinnerton-Dyer's chairmanship moved to UGC to the UFC (Shattock, 2012), and the UFC follows some steps of its predecessor's path. In 1988, for example, the CVCP/UGC working group (CVCP/UGC, 1990) proposed to set up a sub-group on research indicators (SCRI) (Cave et al., 1997) to work on the technical development of quality PIs specifically. Here, '[…] the coming together of technical advances and policy requirements for the development of research PIs was especially strong' (Cave et al., 1997, p.157). The SCRI was composed of bibliometric experts as well as members from the CVCP/UGC working group. The SCRI – and thus also the CVCP – entered in this way the UFC in the early 1990s to collaborate on the 1992 RAE. The SCRI tried to keep account of previous criticisms to the RAEs. Comparing the 1989 and the 1992 RAEs shows the relative successes and failures of the CVCP's proposed quality assesment methods. Contestations over quality assessment thus occurred within the UFC and the committees working with it on the development of new PIs. Though still actively involved in shaping the research quality framework, by 1990 the CVCP's influence also began to dilute as issues raided on the previous RAEs needed to be corrected in future ones.

The criticisms over the strong cost-bias of the 1986 RAE led the CVCP to focus more on assessing research outputs. In practical turns it initiated a publications survey in 1988 (Cave et al., 1997) for the first time since its work on research PIs. The committee gathered a list of disciplines' different output types. The result was published in the UFC's RAE guidelines circular (5/92,), which showed which publication types were selected from the range proposed by the CVCP. Eventually, the UFC used 12 out of the CVCP's selected 20, and asked universities to submit research outputs accordingly.

Between 1989 and 1992, cost centres also changed. "Units of Assessment" replaced the CVCP's previous category, both formally and in practice. Indeed, whereas the 1989 RAE had 37 cost-centres of reference, the 1992 RAE presented 72 Units of Assessment (UFC, 1992, pp. 6-13). This new structure was a novelty for both universities and polytechnics. It was fairer than the 37 cost centres because it kept more account of disciplinary varieties, yet still advantaged the former (Brown, 2004). This is because polytechnics needed to refer to categories devised for universities in order to be evaluated. In managing to keep the cost-centre structure intact and having the majority of its publication survey results used in the 1992 RAE, the CVCP had a degree of success in

shaping quality assessment. Specifically, shifting its attention to publications did indeed guarantee a sear at the discussion table. However, the committee's later proposals for the 1996 RAE found less support. This became evident in 1993, when the HEFCE rejected the committee's research publication PIs, published following a survey on publications which it conducted in 1991 (CVCP,1993).

### 4.3.1. The Higher Education Funding Council for England and research quality evaluation

These proposals were discussed amongst other factors in a report of the research sub-group set up to devise new research quality indicators, the Joint Performance Indicators Working Group (JPIWG). Its role was to advise the funding council as the Further and Higher Education Act delineated (DfE,1992). As the 1986, 1989 and 1992 RAEs formed a varied and bulky amount of information for future RAE-developers to work with, the HEFCE equally had an increasing array of considerations to make. The research sub-group made new steps on PIs which left the CVCP's cost-focused ones behind, now considered almost obsolete. Now a much larger network (see appendix I for the list of members) compared to the one of the restricted CVCP/UGC, the committee mixed more opinions, bringing to light new approaches to research quality evaluation compared to those of the mid-1980s. The sub-group was made up of an array of members, much more varied than previous network of PI developers of the CVCP/UGC. The 1993 report was published shortly after the 1992 RAE yet before the HEFCE's own RAE in 1996, and three key aspects are worth pointing out: its criticism of the CVCP/UGC's previous indicators, its considerations on the challenges of evaluating diverse disciplines through publication indicators, the fact that it made clear that universities' data gathering exercises would increase for both the scope of research evaluation and for still-undefined other scopes.

The sub-group readily emphasized that the methods of the 1992 RAE marked a stark distancing from the PIs proposed by the UGC and the CVCP in 1986, and most crucially from the UFC's in 1989 (JPIWG,1993, §28-29). In pointing this out, the sub-group expressed its doubts on the CVCP's 1991 PIs specifically (JPIWG, 1993,§29) and proposed a different set of possible indicators for the 1996 RAE. Table 1 below show the two groups' different indicators. The CVCP eventually corrected the mistake of the indicator " number of publications per £1 of UFC research grant", which was officially published but was meant to be £1000 of income. In a document first published, calculations conveyed that, for example, that 0.09 publications were produced in England for every £1000 of research income of cost-centre 1, clinical medicine (CVCP,1993). This meant that, on average, each publication of English Clinical medicine cost-centres cost the State eleven thousand pounds. This was considered a rather high cost given that the committee itself advocated that it was working for achieving major cost-efficiency in universities, and wanted to transmit this in its indicators. The publication indicators list was therefore relatively unsuccessful

both because of the credibility lost in the mistake of calculating publication quality according to £1 of UFC money and because the initial data calculations eventually made did not really convey as much cost-efficiency as was hoped.

The heavily cost-oriented PIs appeared to reflect the attempt to signal that 'value-for-money' was being actively sought in universities as the legislator had asked. The indicators proposed by the CVCP were meticulously searching to obtain information on the costs-of research rather than on its quality, and quality was evidently assimilated to productivity. Though in fact these indicators now included the "number of publications"–differently to the 1986 and 1989– they still related this to very specific sums of income received for carrying out research. Relating the number of publications to concrete sums of money (£1000 of external income and £1 of UFC grants) was in all effects a quantitative judgment of research quality. In this respect, the sub-group pointed out that calculating research output based on the UFC research grants as a denominator was misleading, for many universities tended funded their research through their teaching funds. The sub-group's '[…] view of the inappropriateness of the indicators published by CVCP for 1991' (JPIWG, 1993, §29) was reinforced by its own specific view of how publications as an indicator of research quality had to be used. This view criticised how the CVCP's PIs did not '[…] differentiate between different types of publications', and it didn't account for '[…] the wide variety of mediums of output and the different values ascribed to them in different disciplines.' (JPIWG, 1993, §29). The sub-group expressed the danger of these quantitative-based research quality PIs, suggesting that they would encourage'[…] institutional managers [to] pressurise academic staff to publish' (§32). This was in line with the new indicators proposed and the report's broader belief that quantitative data – particularly of the cost-oriented kind– was an insufficient instrument for ascribing research quality because it imposed an equal scale for the sciences, the social sciences, the humanities, and the arts (§21-23).

*Table 1. PIs proposed by the CVCP and the JPIWG research sub-group*

| CVCP's PIs | Sub-group's PIs |
|---|---|
| – number of publications per £1000 of external research income | – number of articles published in academic journals |
| – number of publications per academic member of staff | – number of refereed conference papers given |
| – number of publications per £1 of UFC research grant | – value of research grants and contracts awarded, in particular by the Research Councils |
| | – numbers of research studentships |

The sub-group discussed the challenges of evaluating diverse disciplines through publication indicators in a rather articulated way. First, it elaborated on the fact that though a single scale was

difficult to construct, it somehow needed to be drawn out: it would have been equally controversial –as critics had pointed out in 1986– to evaluate research publications of different disciplines and fund universities by asking panels to use different scales. This fuelled a debate over how the issue could be solved through PIs that focused on research *output types.* Some 'publication type' indicators could be tailored to the social and others to the hard sciences, for example. The group presented 8 publication indicators (plus four which were considered of 'limited application as performance indicators': editorships, other conference contributions, articles for professional journals, other publications) in its list of recommended "institutional performance indicators" (§55-64). The publications indicators reflected the attempt to unveil how disciplinary differences manifested themselves on types of research outputs. To give some examples, the indicator 'number of articles published in academic journals' (§59) was considered an important research output type for all disciplines, the 'major medium of output'. In contrast, many of the indicators listed (such as number of academic reviews and number of edited books), were seen as more common in the social sciences, arts, and humanities. Interestingly, none of the publication types were seen as being specifically common for the hard sciences. Many of these considerations were the direct result of a regression-analysis study carried out by J. Taylor [84]– member of the sub-group. The sub-group had asked him to '[…] undertake a study of the *statistical relationship between the research ratings awarded to each unit of assessment and variables constructed from thee 1992 RAE database.'* (JPIWG, 1993, §36, emphasis added). This was done to capture '[…] the variables that explained the data most efficiently' (JPIWG, 1993, §36).

The insights from the study, and the subsequent description of indicators, could assist panels in making quality judgments, for these could convey that the 72 units of assessments had different combinations of research output types *as well as* quantities. However – and as future research might elucidate – the discussion on the diversity of publication types ended with the mere recognition that they differed across disciplines, for the sub-group's report did not elaborate further on how the observations being made would translate in terms of research assessment carried out by the panels. Reference was made to the fact that future panels would keep account of the type of publication, but no suggestion of assigning a weight to each type was made, for example. Perhaps this is due to the fact that the sub-group saw the 'number of publications in academic journals' as the predominant and hence most useful indicator for RAE evaluation. Finally, the sub-group discussed the governance of university data. The report aimed to use the 1992 RAE data to construct indicators ' also for other purposes'(JPIWG,1993).The sub-group was well aware that the indicators' audience was not limited to the expert panels, but also to new stakeholders in HE and, of course, 'institutional managers'. It acknowledged existing  concerns on the confidentiality of the data that it was asking

---

[84] See Cave et al., 1997 for a debate on Taylor's study.

from academics (§70-73). Indeed, the academic body was concerned of the uses that could be made of all the sensitive information being gathered Indeed, the users of this data remained largely undefined as a wide array of interested university and non-university groups (JPIWG,1993). Still, the sub-group turned down these concerns refusing to set rigid boundaries on data use and users. Echoing the national data survey's discussion of how 'data users' interested to open universities' books needed sufficient and well-gathered information,(CVCP, 1985b), it made it its priority to ensure that universities comply to submitting information without having certainties over its use. This is of no marginal significance, because it shows that RAEs were often accompanied by increasing and always more-specific demands of data from the academic body. This can be considered important in the wider debate on staff's increasing exposure to scrutiny and judgment from the non-academic community, and not just from the evaluating experts, but also from any interest group. These include agents from within the sector (institutional representatives, trade unions, etc.) and outside it (industry, firms, consultancies). The benefits and disadvantages of publicly available detailed information in general and in universities specifically can hardly be balanced out in general, theoretical terms. Accountability and transparency are now catchphrases, which carry access and manipulation as two sides of the same coin. On one hand the dissipation of universities' information on staff quantities, types, salaries etc. is important to academics as an instrument of comparison and evidence when substantiating their claims. This was the case for academic criticisms of pre-1996 RAEs, which managed to influence later PIs. On the other, what is peculiar about developments in research data gathered between the mid-1980s and the mid-1990s is that it was heavily oriented on obtaining information to calculate productivity. The sub-group for example showed that an improvement in the 1992 RAE was the decision to ask research submissions only from research active staff, narrowing down the burdens and the evaluated subjects. At the same time, the A+B+C+D categories used still included disputable definition of 'research active staff'[85]. Opening the books of staff information to such a degree was a debatable decision for the time, particularly considering that the report insisted on systemic data gathering whilst recognizing the dangers[86], eventually concluding that […] therefore it is recommended that the entire database, except personal details […] is made available electronically on request to institutions, Research Councils et al, other government bodies and other bona fide users' (JPIWG,1993, §,73). In many respects, the data-extensive approach of the sub-group formed

---

[85] A & B: '[…] academic departmental staff (teaching, teaching and research, or research only) who had a contract with the institution and received a salary from it irrespective of source of funds'. C+ D was: 'staff who make an independent contribution to the research of a department but who did not have a contract with the institution or receive a salary from it.' (JPWIG, 1993, §45).
[86] '[…] there would be some who would misinterpret information contained on the database by using it incongruously or by subjecting it to inappropriate analyses […] but perhaps of more importance is the value of information contained within the database for institutional managers, academics and other individuals and bodies (JPIWG, 1993, §71,72).

important steps for open-ended data governance, rendering the academic body vulnerable to manipulation. Academic staff did not have an easy way out, nor was included in discussions over the use of its data, aspects which represent the endurance of a broader mistrust in the academic community initiated in the mid-1970s.

To conclude, the research sub-group's report provides insights for analysing the genesis and initial developments of research quality evaluation more broadly. Firstly, the report shows that the end of the binary line created a re-distribution of powers, as the network developing research quality PIs changed from being a CVCP/UGC and CVCP/UFC group to that of the JPIWG which comprised polytechnic representatives, university ones, professors, and funding council members. This re-configuration largely diluted the influence of previous PIs on the construction of post-1992 RAE research quality PIs. As has been shown, the sub-group instantly questioned the CVCP's PI list, and its approval of the conversion of 37 cost centres into 72 units of assessment in the 1992 RAE was yet another step towards the re-definition of the CVCP-proposed structure for carrying out evaluation. With regards to cost-centres, the sub-group stated that:

'[…] should indicators be calculated for cost-centres, institutions that made a number of submissions to unit of assessment within the same cost centre would, in some circumstances, receive virtually meaningless information as differences in performance between the various submissions would not be apparent' (JPIWG, 1993, §31).

This crucially attacked the core of the CVCP's lists of indicators, such as those presented in 1986 and 1987 which proposed to calculate performance indicators at a cost-centre level. The unit of assessment can be interpreted as a more 'disciplinary-friendly' category compared to the more abstracted cost-centres. The new network of PI-developers tried to overcome the managerial distortions created by past performance indicators, marginalising the CVCP's managerial approach to research quality evaluation by highlighting its inconsistencies at a disciplinary evaluation level. Secondly, the report of the sub-group tells us a bit more of the how the first reactions to the abolishment of the binary line played out within the HEFCE's research sub-group, which was in all respects a direct advisor for the RAEs (Cave et al., 1997). What can be gathered is that 1992 RAE would be taken up as a good structure to work with, and that the drastic increase of data submitted by universities after 1989 for the 1992 RAE provided a new pool to work with for a long while, and with which considerable comparisons with past data and RAEs could be made. It also seems that the sub-group somehow tried to create new channels for the use of university data on research, in this way also streamlining the data that would be used for RAEs. By suggesting that other users

make use of data such as the CVCP's surveys or PIs, the sub-group implicitly reserved to itself the use of certain types of data and rejected the use of other.

## 4.4. Conclusions

This chapter has shown that the development of evaluation criteria in British universities has occurred in a multifaceted and highly political process (Cave et al.,1997). The relative influence of diverse actors has thus been constantly analysed in relation to that of other groups. This chapter has focused on the influence of the CVCP within the university sector vis-à-vis that of the State policies, first, and that of other university groups, after. As State policies aggressively defunded the sector, the CVCP found pathways to articulate performance assessment. Throughout the 1970s and into the 1980s, the government juridified higher education (Neave, 1988, 1998) by intervening on grant allocations because cost-cutting was its interpretation of efficiency. Further, Thatcher's government soon realised that to control to cut funds to the sector, it first had to control it, for it had been traditionally set up to function in autonomously allocating State grants. Thus, student numbers became the criterion for trying to steer the sector into cost-cutting. The fact that for the State, efficiency gains translated into policies aimed at relieving the State's financial burden on universities is further exemplified by the fact that the Jarratt committee was left free to carry out the sector's equivalent of a 'financial management initiative' as long as it could demonstrate its determination to implement cost-considerations within universities. As has been discussed in detail, the Jarratt Committee was quick to turn this apparent constraint into an opportunity to strengthen its role both at the sector level and within universities. Instead of considering the Jarratt report as the first of many instances of universities' adaptation to a corporate logic (Deem et al., 2007), this chapter has argued that vice-chancellors' advocacy for the PPBS-inspired managerial model incorporated the committee's specific interests, articulated innovatively through an exploitation of the financially stringent climate heralded by Thatcher. The committee introduced performance indicators when the State hadn't yet begun devising them for higher education. This vanguard position was facilitated by its network with Geoffrey Lockwood, who brought research on PPBS in university governance into the committee and helped to support an explicit attack on academic staff in governing bodies. The committee in fact placed the VC figure at the centre of the executive model when it introduced concepts of planning such as budgetary devolution, flexibility, and cost-efficiency. The Jarratt Report is thus a political document through which the CVCP tried to legitimate the use of systems analysis methods in universities to subordinate academic staff to the roles and priorities set by VCs and aspiring university managers and planners. As performance became rearticulated into a structural necessity to be defined by top-hierarchies, internal performance indicators for university planning were conflated with the CVCP's first research

quality indicators. This allowed the committee to *de facto* pilot research quality evaluation in the English university sector, and to do so through a managerial imprint. A key novelty was a new organizational unit for facilitating performance evaluation, that of cost-centres. Borrowed from the principles of budget centres in universities (Lockwood and Fielden, 1973), this innovation found its application in the first two RAEs as well as within universities, which adjusted to provide data on activities according to cost-centres. This problematizes the notion that evaluation of universities emerged through a top-down process guided by the State with a managerial imprint (Deem et al., 2007) or a marketizing one (Tapper & Palfreyman, 2014; McGettigan, 2013). The later establishment of funding councils with evaluation functions appears more as an attempt to appropriate the university sector's innovations on research quality evaluation rather than that of defining them *ex novo*.

# Chapter V

# The Rise of Evaluation in Italian Universities: 1989 – 2010

Italy is generally considered a latecomer to NPM and university evaluation compared to England. Still, university agents often tried to keep up with international developments and push for national changes. This chapter explores these actions, analysing the rise and initial development of evaluation in Italian universities. The chapter begins with governments' attempts to introduce efficiency in the public administration and in higher education, exploring the reforms and assessing relative successes and obstacles to their implementation. The section briefly identifies how these reforms influenced policymaking in higher education at the turn of the 1990s. As is discussed, one of their key pillars was the definition of universities as autonomous institutions, an aspect which actually turned universities' financial autonomy into a new bargaining instrument for both the State and university agents. The section ends with a discussion of the developments on the mid-1990s, when a governmental body charged with overseeing the university sector, the Osservatorio, presented its preliminary experiments with a new funding formula to re-organize the funding arrangements of the 1950s still in place in the 1990s. Secondly, the chapter turns the analysis to the university sector. Specifically, it discusses how such policy changes were perceived by the CRUI by exploring the actions it took to modify emerging developments that it disagreed with, such as minister Ruberti's 1989 law on  university autonomy. Driven by the determination to tailor autonomy, the CRUI actively attempted to counterbalance political requests with those of the academic community. As I show, this step engaged VCs significantly in a variety of aspects around autonomy, including university responsibilities in internal governance and evaluation procedures for improving the sector's lagging graduation levels, and brought the committee to equip itself with new skills and knowledge for retaining financial and organizational autonomy from the State. As I show, these efforts drove significant innovations on internal evaluation procedures and methods, marking the CRUI's relative advantage over the State in its ability to draw out a comprehensive framework of evaluation. Finally, section 5.3 explores the committee's turn to research quality indicators and its successes and failures at institutionalising them in the sector in a context where state-devised regulatory bodies such as the Osservatorio and the CIVR were beginning to exert their own influence over selective funding and research quality evaluation, respectively. The university sector's trajectory is thus marked, similarly to the British one, by a series of specific contestations over different groups' evaluation indicators. As I suggest, the CRUI's managerial approach to university governance did not make use of cost-efficiency indicators as the CVCP did. Rather, the CRUI's managerial understanding of governance emerges from its initial attempts to turn evaluation

into an instrument of governing bodies, first, and its later attempts to devise a framework centred around its role in monitoring universities' research.

## 5.1. The State tries to catch up through reform

The reforms of the Italian PA geared at implementing efficiency allow to broadly contextualise the arrival of managerial practices in policymaking. The priority of the reforms of the early 1990s was the simplification of administrative procedures, a process often hindered by the volatility of Italian governments and the fact that complex performance management practices were often rearticulated as processes of cost rationalisation. Acknowledging this allows to better grasp the way that early efficiency reforms for the HE sector also focused on reducing the state's financial burden rather than on setting up evaluation procedures.

### 5.1.1. A long series of attempts to simplify administration

Italy's trajectory in implementing efficiency within the State between 1979 and 2001 was marked by a series of consecutive failures. These have been associated to Italian PA's legalistic tradition, with a constant recourse to law-making as the key device for regulating the socio-economic sphere, and to its excessively frequent change of governments (Ongaro, 2006; Ongaro & Vallotti, 2008). The State's machinery was first reformed in the period1979 - 2001, with different methods advanced throughout. Key instances are the Giannini report in 1979 and the administrative law of 1993 which introduced internal evaluation units in the PA. Along with the Bassanini reforms passed between 1997[87] and 1999, this period marks Italy's turn to NPM-oriented governance.

The first attempts at a systematic administrative reform were set off in 1979 with minister Massimo Severo Giannini (in charge during both of the successive Cossiga governments, thus between 4 August 1979 and 28 September 1980). His non-paid ("ministro senza portafoglio") political role sparked the debate of reform of Italian Public Administration. His broad aim was to improve the relationship between the PA and citizens by simplifying the State's bureaucratic machinery (Tosatti, 2012). Giannini, professor of administrative law, articulated his proposals in the "Giannini Report," published in 1979. Here, key issues of the 'business State' were emphasized, elaborating on delegification, efficiency, and the measurement of productivity of the civil service, presenting them as essential features of a modernized PA. Giannini explained how the State was facing an 'organizational drama' (§1.3) because it had the responsibility of providing public services and regulating them more efficiently than private enterprises. According to Giannini the excessively frequent recourse to legislation often disclaimed an efficient functioning of State-provision, reason for which the State of the 1980s had to ensure that administrative procedures did

---

[87] See the Bassanini law of 1997, number, 59. In art. 21 of the law, it was explained that the schools' autonomy being introduces was part of a broader process of administrative reform of the sector aimed at granting more autonomies and re-organizing education at all levels.

not overwhelm society and develop into hindering devices. The State thus had to acknowledge how public governance was failing in order to improve it, starting from an analysis of its machinery. The decline of governance could be reversed through the use of efficiency indicators, which would '[…] adapt to changing political directions' (Giannini, 1979 p.292, § 2.4) and ensure continuous improvement without depending on political principles. The technical-objective character of indicators was their strength in combatting sudden political changes often caused by changing governments. A focus on administrative productivity would also reduce the high procedural hidden costs of the PA (Giannini, 1979), as efficiency indicators could allow to identify the most productive policies more readily and implement them. Giannini insisted on comparing the State to private enterprises to make his point, showing how administrative complexity caused the State to delay its provision of services, three times that employed by private counterparts (Giannini, 1979;Tosatti,2012).  He emphasized that hesitation to the use of indicators needed to be overcome through their experimentation and tailoring, as the '[…] possible imperfection of results cannot be seen as a factor exempting public powers from the need of applying indicators as this disadvantages the PA compared to the private sector' (Giannini 1979, § 2.3). The multiplication of an expanding State's functions and offices risked becoming further obstacles to an efficient State if these continued to be governed through the traditional recourse to laws and decrees.

This genuine concern led Giannini to initiate an enquiry of the State's bureaucracy, and specifically of its staff, through the FORMEZ consultancy, which assessed 300 offices in 16 ministries (Ferraresi, 1982)[88]. The enquiry's work is well summarised by its view that in the PA '[…] the structure precedes the functions, and prevails over it' (Ferraresi, 1982, p.70). It confirmed Giannini's preoccupation with the legislator's excessive demands of rules and procedures to be firstly understood and analysed and then followed by civil service staff. This was seen as the key hinderance to staff productivity, '[…] the principles of legality are often prepended to efficiency ones' (Ferraresi, 1982, p.60). The productivity enquiry was thus seen as crucial for responding to these constraints. It supported Giannini's emphasis of a practical approach to ministerial coordination, measuring staff productivity in terms of the *effective vs potential* number of work hours required for carrying out daily administrative tasks. This enquiry, based on interviews and surveys, concluded that low administrative productivity (between 40 and 50%) was the result of time dispersed or wasted in debunking numerous legislative procedures to be followed for policy-formulation (Ferraresi, 1982). The issues raised then are strikingly recent, as are their proposed solutions which still populate current public debates on the Italian PA: decreasing the number of ministers to avoid duplicating functions, cutting labour costs – the State's highest expenditure – and overcoming the Napoleonic model through an organizational re-structuring which mimics the

---

[88] The enquiry was carried out between 1980 and 1982 (Ferraresi, 1982)

Anglo-Saxon departmental model. Generally, productivity indicators were seen as devices that would help to rapidly and practically modify civil servants' functions as well as the coordination between offices because they would provide a means for comparing productivity between ministries without needing to recourse to legislations for enacting change. *In other words, at this stage indicators were primarily conceptualised in their regulative capacity rather than in their performance-assessment one*. Despite the far-reaching proposals, the PA was in fact not reformed as was hoped. In 1980, "organization and method" offices were established in the PA to initiate the formation of technically skilled civil servants. This was made compulsory by a council of ministers' resolution, and the offices were asked to study and identify key productivity indicators. Offices would then ideally advise on specific policy cases. However, these offices never acquired the desired central role (Tosatti, 2012). The change of government in October 1980 and the subsequent replacement of Giannini's role as ministry of public functions further marginalised the proposals made.

As the "Tangentopoli" scandal[89] emerged and the "clean hands" investigations developed, the State's then leading parties, the DC, and the PSI, crumbled and lost their electoral support. The decline of the left not only resulted in a shift towards moderate and right-wing politics, but also to the establishment of the first (of many) non-parliamentary and "technical-transitional" government, that led by Carlo Azeglio Ciampi, ex-governor of the Bank of Italy. The stark delegitimation of politics played an important part in shaping Ciampi's reforms, which made transparency and political responsibility new imperatives. The reforms to the PA advanced at the start of the 1990s, however, still faced the persistence of the "administrative law paradigm" (Capano, 2003). The latter is still frequently identified (Capano, 2003; Vallotti, 2000; Del Vecchio, 2002, Ongaro & Vallotti, 2008) as the major impediment to a concrete de-institutionalisation of Italy's legalistic and juridical PA and policymaking, also when managerial-oriented political streams exist. [1] It has been argued that management reform was even more difficult to carry out at the regional levels, where budget and performance reforms did not at all play out as expected (Ongaro & Vallotti, 2008).

One such managerial stream emerged in 1993 through decree number 29[90] which established "internal control nucleuses" (known as SECIN, "Sezioni di Controllo Interno") in every ministry. Interestingly, this law was a live emulation of the U.S.A's Government Performance Results Act[91] (GPRA),which introduced performance assessment procedures in federal agencies following the PPBS imprint and was passed in the same year.  As in the USA, for Italy this was not the first

---

[89] This scandal first emerged in 1992 and revealed that there was a high degree of bribery involved in Italian politics, where politicians guaranteed favours to businesses in exchange for large sums of money. The scandal initially involved Milan's group of the Socialist Party, specifically Mario Chiesa, and as investigations continued it emerged that all political parties were involved in this system, bringing a tremendous decline in electoral consensus.
[90] https://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario
[91] Voted on January 3rd, 1993, and inherited by the Clinton administration from George H. W. Bush's.

attempt to introduce PPBS in State corridors. Indeed, the GPRA's roots traced back to the diffusion of systems analysis and engineering in the 1970s (Archibugi, 2012), finding fertile ground in a marginal body of Italian 'planolgists' (Archibugi,2012). These researchers -–coming from sociological, urbanistics and engineering backgrounds – had tried to unsuccessfully institutionalise (Archibugi, 2012) studies on the transformation of the welfare State by advocating for the application of systems programming to economic policy and reporting their analysis of the US experience to the Italian government (Archibugi, 2012). Although presented to government (Archibugi, 2012), reports of this kind found inattentive audiences, reason for which as of 2012 Archibugi himself – author of "*A report on the introduction of budget programming in Italy*" (1970) – criticised politicians' dismay at the scarce implementation of managerial control procedures in government suggesting that their scarce past attention was the primary cause (Archibugi,2008) of managerial gaps. Archibugi sustained that the context in 2012 was almost identical to that of 1993, arguing that the 1993 law could have provided a definitive transformation yet failed because politicians limited themselves to the mere enunciation of strategic planning (Archibugi, 2012).

Differently to the USA, in fact, Italian legislators had not introduced guidelines on how the civil service was meant to develop strategic plans, nor had they required, as did the GPRA, that civil servants produce annual "performance plans" summarising departmental tasks. In other words, SECINs were introduced but there were still excessive ambiguities on how their functioning. No prescriptions were made of *ex-post* reporting exercises which are crucial for performance assessment and forward planning in the terms of program budgeting reform. This context of ambiguity led to a vicious circle marked by a constant '[…] debate over who had to lead these control functions, '[…] a debate over power matters, *more italico*' (Archibugi, 2006, p.6.). The SECINs introduced management and strategic controls yet failed to produce objectives and indicators, there were therefore no '[…] technical requisites necessary for exerting, – *conditio sine qua non* – the 'management controls' which the laws of the 1990s repeatedly advocated' (Archibugi, 2006,p.6) . The 1993 law did create a central organ which could guide change, nor did it provide standard practices for ministers to follow; preferring instead to leave the full responsibility of developing evaluation and management performance procedures to the ministers. This situation was further hindered by the fact that civil servants had no kind of familiarity with these concepts and could not easily be transformed into efficient policy performance analysts even with relatively expert ministers. The failures of the Giannini report made this even more of an obstacle, for there was no existent  rooting on strategic planning and neither of its *ex-ante* and *ex-post* dimensions. As such, '[…] SECINS on their own, without a cooperation or a political guide coming from expenditure bodies, simply could not work' (Archibugi, 2012, p.5), and it did not. The Italian PA required, from a strictly performance budgeting framework, an *ad hoc* expert body which

would introduce strategic programming and concretely train civil servants, a key step for '[…] transforming each PA manager into a strategic-programming oriented operator. [Strategic programming] is not just a practice to be introduced, but also a mentality and a way to understand one's role and mission, a new modus operandi of the PA' (Archibugi, 2006, p.).

From a system-analyst's perspective (Archibugi, 2006), Italian legislators thus very scarcely filled managerial gaps in the laws which they passed, also in those of the late 1990s. In 1999 (law n.286), SECINs were further regulated through the additional demand that they make yearly reports. A technical-scientific committee was also established to introduce strategic planning in ministers and remedy previously missing technical guidelines on the principles and modalities. Minister Bassanini – ex professor of constitutional law – tried to introduce a more radically managerial framework than those previous ones. (Ongaro & Vallotti,2008). His role as minister of the public function in Prodi's first government (between May 1996 and October 1998) left a key legacy in Italian PA, namely an array of decrees and laws named after him: the "Bassanini laws". Whilst it would be outside the scope of this thesis to discuss in detail the complex adjustments made through numerous decrees passed in these years, the Bassanini laws allow to identify what the political understanding of managerial reform was. In terms of equipping the PA with new managerial practices oriented at performance reporting to be monitored, these laws remained largely general because they introduced processes of simplification of procedures rather than processes typical of performance oriented policy analysis, with the result that the PA witnessed only a shy approach to *measurement* and *calculability*. In other words, Italy was not exempted from NPM's "betrayed promises" (Perry et al, 2009) of achieving performance assessment at a political-economic level. As of 2009, the *de facto* use of performance controls in Italian ministries was scarce and close to absent, so much that in 2014 the very meaning of performance management was reviewed (Rebora et al., 2016) through yet another legislative decree, n.190/2014. Whilst attributable to the abovementioned *administrative law paradigm* (Capano, 2003), this failure also reflects that the broader structural complexity and technicality (Bouckaert & Peters, 2002) of management control instruments which different state PAs have faced as a difficult challenge and been unable to concretely implement. This is key for understanding how much the Italian path towards NPM is characterized by a generalised articulation of such complex PPBS-inspired procedures. Rather, two diverse views of efficiency and productivity existed in the 1997-2001 reform period.

One view acknowledges that introducing the PPBS-inspired performance mechanisms is not merely a case of measuring productivity does not simply come about by scrutinizing staff's work. This view recognizes that implementing a systems-analysis inspired planning and programming framework consists, crucially, in reshaping *the entire accounting complex* (Rebora et al., 2016) of

an organisation[92]: accounting models, standards, templates, skills and professional profiles. Accordingly, performance management controls can only really succeed through a concrete homogenization of new procedures across the entire PA and, crucially, also across bodies which ought to regulate it civil servants' work. At a central level, it is crucial that the budget itself be re-designed according to programming principles around objective-setting and budgeting. This means that little use can be made of a generic re-design of the State's budgetary planning if civil servants are then unequipped with different programming and planning procedures or, as in the UK, engage in a specific articulation of policy analysis. One example of this in Italy is establishment of "centres of administrative responsibility" in 1997 (law n. 94 of 1997). These worked more as "financial accountability" organs (Archibugi, 2006). They did indeed adopt new procedures for reporting resource use, however they were not part of a coordinated re-design of the state's resource-planning framework. The result was that the centres were indeed trying to modify their approach but then did not find any concrete effect of this because, as Bassanini himself acknowledged (Bassanini, 2010), the central state level had been unable to re-define the State budget through missions and programs (Bassanini, 2010). This central re-design was more formal than practical, and also later remedies such as those in 2009 (Rebora et al., 2016) failed to accomplish a full organizational re-structuring of the state budget, the programs within it, and the middle and lower levels' reporting of ministerial programs accordingly. The result was that performance assessment itself was politically redefined as an accountability exercise. Introducing evaluation procedures was thus a major unsuccess, even though Italy actually anticipated other European countries' reforms (such as France's LOLF reform) in terms of gradually abandoning incremental state planning towards program planning.

Bassanini's reforms reflect the second (more common) view of management practices which treats management as a synonym of procedural and organizational simplification, and thus as a practice pursuing efficiency for this reason. From this perspective, elements alternative to abovementioned ones can suffice to depict an optimistic picture of administrative reforms: These include the reduction of civil service staff costs, where '[…] by the end of the year 2000 Italy had 2 million fewer civil servants than France and the UK' (Bassanini, 2010); the merger of ministries the partial transfer of administrative responsibilities to the regional and local levels. Italy's thrust to enter the single currency importantly turned alignment to these processes into the PA's priorities. These were indeed key features of the Bassanini laws which from a simplification and rationalisation perspective were indeed successful (OECD, 2001; Bassanini, 2010). These aspects make Italy a critical case as far as the *de facto* implementation of management practices in the State machinery is concerned, a process which was aggravated by the country's extreme political

---

[92] Arguably, it is problematic to consider the PA as an organisation in the first place.

turnover (Ongaro & Vallotti, 2008) marked by 21 different governments over 22 years (March 1979 – Apr 2000).

*5.1.2. University autonomy and expenditure control*

Between 1989 and the mid-1990s Italian universities were encapsulated in the administrative reform period discussed above. This broad timeframe can be summarised as the beginning of devolved university governance, not void of its contradictions. The State shifted major responsibilities to university governing bodies, the primary one being financial management and the organization of staff's research and teaching activities. Throughout this period, a new funding structure for the sector was proposed and evaluation procedures were first articulated The aim driving the key efficiency reform in 1993 was intended to sustain a comprehensive re-shaping of the relationship between the State and the university sector by radically modifying the public funding structure. However, the fact that universities first needed to be devoid of the financial comforts brought about by a fully State-managed institutional governance made this a challenging task for the State. Indeed, policymakers instrumentalized the notion of university autonomy in order to achieve efficiency through the new funding structure they had in mind. This emerges, as discussed below, in the different meanings that university autonomy acquired during different governments: between 1989 and 1993 it was meant to introduce new and unprecedented self-governance freedoms, and between 1993 and 1996 it was meant to represent a bargain for more cost-responsible behaviours in universities. Within this process, the introduction of evaluation appeared more as an instrument to ensure financial accountability feedback in a new state-HE relationship rather than a means to monitor the quality of universities' teaching and research. Firstly, law n. 168[93] of 1989 introduced pillar principles for the University which in Italy still did not exist. These rotated around the central principle of autonomy, now granted in organizational, teaching, financial and research terms. For the first time, universities also had powers to set their own charters, opening the doors to academic self-governance. Regardless of its ambitious prescriptions, the law was quickly superseded by what is also in academic literature considered a more significant law, namely law n. 537 of 1993[94]. This law has replaced the former's reputation for having *de facto* granted university autonomy. The reasons behind the gaps of the 1989 law will be elaborated in detail from academics' perspectives in part 5.2, which explores university agents' understanding of their changing context. First, it is worth exploring such a changing context, and consider how and why the 1993 law was so innovative for universities by studying it from different governments' perspective and investigating its underlying driving motives. This is crucial for fulfilling the aim of this thesis, namely problematising the emergence of university evaluation as a top-down process. In many respects, university efficiency

---

[93] https://www.gazzettaufficiale.it/eli/id/1989/05/11/089G0202/sg
[94] https://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario

was the daughter of autonomy and accountability demands. As I argue, when initiating NPM reform in the HE sector politicians did not articulate evaluation as exhaustively as scholars assume (Pinto, 2012, 2019). In fact, law 537/1993 had a strongly administrative character, it was first and foremost concerned with preparing the grounds for devolving governance to institutions in order to allow the State relieve its financial obligations to the sector. Within this goal, evaluating efficiency was part of universities' new financial powers and understood as a product of autonomy. Evaluation was part of autonomy in broad terms and did not refer to the assessment of the quality of teaching and research activities specifically. Looking at the Amato and Ciampi governments' documents and research shows why this is the case, and reveals once more that the concept of quality in NPM paradigm is highly vulnerable to the agents articulating it.

### Why a new law on autonomy?

For implementing devolved governance, the State first needed to find ways to radically re-structure its obligations towards the university sector. Introducing autonomy allowed to make this path because it created both the consensual and practical grounds for constructing a new funding channel. A new approach for analysing the sector's productivity crystallized first of all in the government's technical committee for public spending ("Commissione tecnica per la spesa pubblica", TCPS from now on) and amongst some key proponents of fiscal federalism and expenditure control; Giuseppe Catalano, Piero Giarda, Paolo Silvestri and G. Brosio. Their policy advice for university policy into the 1990s (Brosio, 1993; Catalano & Silvestri, 1992; Giarda, 1993) introduced two novel concepts for the sector: working with lump-sum budget constraints and the notion of a "system of cost-analysis"(Giarda,1993). Piero Giarda had a background in economics and specifically in fiscal federalism, which he developed importantly whilst at Harvard and Princeton between 1962 and 1968, where he studied US economic governance and brought some insights back to Italy. Convinced of the benefits of a fiscal system, Giarda's views were practically developed during his political roles as president the TCPS (from 1986 to 1995) and as vice-secretary of the Treasury (from 1995 to 2001). In terms of Italy's budgetary policy, Giarda argued for a public service funding systems based on devolution and regional incentives, he was convinced that regions needed greater autonomy and responsibilities and that this would help the State to eliminate major resource wastes (Giarda, 1966).[95] This view extended to the public funding of the university sector, which in his view necessitated an efficiency-driven shift away from a State centralized management model He thus suggested that the State consider two novel approaches: one

---

[95] See Esposito, 2019, for a critical study of the development of fiscal federalism in and its distortionary effects on Italy's southern regions.

for streamlining the State's funding channels to the sector, and, related to this, a channel for implementing academic accountability (Giarda, 1993).

In stark contrast to English universities which were free to administer funds, Italian universities obtained public resources through multiple and pre-established channels, each for different type of internal university expenditure: staff, estates, students, equipment (amongst others). This system worked well during the sector's post-war expansion, for it directly provided universities with resources and allowed the state to constantly intervene for sustaining growth. At the same time, this system rested on a large number of bureaucratic procedures which often tended to delay the concrete use of resources, leading to resource waste and thus in these ways often hindering universities' positive development (Giarda, 1993). As a matter of fact, the State injected significant sums of money in the sector following the post-war period, yet dropout rates were very high in the country well into the 1990s (Vaira, 2011). In light of these Giarda argued that the State would soon witness an even wider gap between resources assigned and universities' results (graduation rates as a percentage of students enrolled) if it did not act to modify the sector's funding structure. This inefficiency was in many respects both allocative at the State level, with the state struggling to find the right resource channels, and at an institutional level, with universities unable to steer spending for increasing attainment levels.

As such, Giarda's proposal rested importantly on how the State could incentivise universities to improve their resource management: the centralized and fully public structure seemed to be too reassuring for universities, who arguably treated public funds as '[…] free goods without costs, goods which can be appropriated, given that the financial cost, first, and the costs for the socio-economic system, secondly, fall on subjects that are not using these goods.' (Giarda, 1993). A change in universities' behaviour from that of mere beneficiaries to that of responsible managers was thus necessary. According to Giarda, funding devolution could initiate this transformation:

> '[…] I propose to drastically reduce the funding channels for universities to a single (or maximum two) 'fund for the funding of the university' which – distributed between the single universities – will allow each to pay the entire cost of their 'inputs' (factors of production), used in the production process. The resources for the cost of professors, researchers, non-teaching staff, functioning costs, libraries, and access would be managed in *full autonomy* (accounting and administrative autonomy) as is already the case in private universities (Giarda, 1993, p.83, emphasis added)

Single block grants could thus shift funding responsibilities to universities and replace traditional funding channels, radically reducing them to a single source of public resources. This would in turn

bring about virtuous internal management behaviours, as universities would be faced with concrete and unprecedented decisions on distribution. It is with this rationale that government's experts such as Giarda argued for university autonomy. They supported the principles of autonomy prescriptions made in 1989 (the introduction of self-governance freedoms in universities) yet added key practical elements to tailor autonomy so that it could concretely also spark a new relationship between the state and the sector. Most importantly, they tried to explore how financial autonomy could be used to shift responsibilities of resource mismanagement to the universities. Universities' new managerial responsibility represented a challenge that they would have to take on in exchange for greater freedoms

Within this new framework, the state could ask push universities into strategies that would in the long run reduce the State's financing burden on the sector. Indeed, Giarda's proposal also comprised consideration on costs. Surely, the State would continue to fund the sector through the single channel. However, a new chain of responsibilities would emerge from this structure, for the idea was that those receiving education funds needed to participate in sustaining its costs, and these groups were the universities and the students. Indeed, universities were encouraged to consider the importance of student fees as a way to diversify university income (Giarda, 1993). This was part of a broader conception whereby universities' financial autonomy would ideally develop greater '[…] awareness of the relationship between results and costs' (Giarda, 1993, p.84), beneficial to the State because '[…] incremental funding […] can no longer be justified on sufficiently rational nor convincing basis (Giarda, 1993, p.84)'. As such, institutions' financial autonomy could justify the abandonment of incremental funding on the basis of institutional financial sustainability, with universities in future expected to demonstrate cost-efficiency and responsible allocation. These proposals reflect why the 1993 government intervened on earlier political prescriptions on university autonomy and, importantly, how managerial capacity was articulated as functional to the new political will of reducing HE costs in the long run. Furthermore, fiscal contraction had already been sustained for a while as a response to Italy's growing debt/GDP ratio, accumulated throughout the 1980s (Felice, 2015). Whilst it would have been rather contradictory for the State to justify lower HE funding in a context of still high increases in student enrolments, this narrative on responsibilities offered a suitable alternative justification of the same objective. It is important to point out that devolved governance in this way gave a new leverage to the State for the reasons mentioned above, and thus that it is very difficult to generalise as a retreat of state responsibilities. If of course on one hand the state could adopt this strategy to effectively achieve cost-reductions in the long run, it simultaneously widened the space for conditioning the sector to new demands even without a direct central control.

**Why evaluation?**

In a TCPS report on the yearly budget law, Catalano and Silvestri (1992) explored these issues and suggested pathways for reducing the State's funding burden Following an elaborated study of HE costs the authors concluded – similarly to Giarda – that HE was inefficient at producing graduates regardless of rising State investments which accounted for 85% of public universities' total income (Catalano & Silvestri, 1992). Their proposals, however, rested on a different type of analysis to Giarda's. Specifically, they suggested that the State encourage systemic cost-benefit analyses in universities, for this strategy would allow to keep track of their productivity. Their report was a first go in this attempt. It did not present productivity indicators, but rather a state-of-the-art worth being discussed because it highlighted two key issues: efficiency and equity.

Inefficiency was seen as the result of unproductive university behaviours, on one hand, and of the funding and distribution structure, on the other. In general terms, the situation was worrying because although the State was investing an average of 9 million lire per enrolled student, graduate success rates were constantly decreasing since 1970. The study showed that at the end of the 1980s only 36 out of 100 enrolled students actually completed their degree course, representing the lowest achievement rates in Europe (Catalano & Silvestri, 1992), meaning that high student drop-out rates of the mid-960s (Vaira, 2011) were still the central concern well into the 1990s. The 36% success rate level was worrying for two reasons: it marked a fall from that of the 1970s (55%), and it showed that students' average stay at university was of 7 years (Vaira, 2011, see p.59). Importantly, the authors accompanied this macro analysis to a micro-level analysis which investigated the outcomes of State funding within institutions. They did this by analysing how costs varied within and between universities, making thus "requirement" and "attainment" comparisons (Catalano & Silvestri, 1992).This attentive peeking into the practical outcomes of the sector's funding structure allowed to decipher how much of its inefficiencies were actually the State's responsibility. Resources per student varied enormously between disciplines in the same university, for example, as well as between universities in general. Similar trends were found in teaching loads, which varied within and between universities but especially *between disciplines*: to give an example, a science teacher at Salerno had 173 students in stark contrast to one in Trento, with 37 (Catalano & Silvestri, 1992).As the study explained, this was clear evidence that universities needed major organizational freedoms. The State was importantly responsible because it decided the majority of faculty-level resource allocations. Although universities could in theory autonomously decide 50% of resource distribution at the faculty level, in practice they actually had much less leverage because the State's central coordination of faculty staff numbers conditioned how funding could practically be distributed in faculties. It was this dynamic which accounted significantly for the cost-variations presented in the study, and not just the universities' irresponsibility or inefficiency. These structural

weaknesses could for the authors be remedied by leaving universities the ability to internally manage their funds, in the hope that future behaviours would aim to re-balance student staff rations between faculties and thus in the aggregate ameliorate the divergences between universities. If universities did little to remedy the low graduation levels and dropout rates, thus, it was also in part because they could do little to improve student-staff ratios or influence costs if they could not influence staff numbers autonomously. Rebalancing Giarda's outright attack on universities' managerial abilities, this study supported for different reasons the concrete political implementation of institutional financial and organizational autonomies.

Secondly, equity was another structural problem, and the report discussed student fees as part of the solution. Taxpayers' money functioned in a regressive way, with the State equally covering university costs for students of different social backgrounds acting in this way as a "reverse Robin Hood" as the authors put it (Catalano & Silvestri, 1992). This funding structure dated back to the 1950s when the post-war arrangements had decided to fully sustain enrolments, so was arguably no longer suitable in a society now composed of a better-off student generation. Thus the proposal that the share of fees in university incomes increase in order to help the redistributive process. This proposal went along with the suggestion of pushing resource-management developments from below and thus '[…]modify the funding structure of the university system and the modality of internal governance' (Catalano & Silvestri, 1992, p. 25). This confirmed the view that State and institutional level transformations depended on each other and needed to change together. This approach found effective integration in the 537/1993 law proposal advanced shortly later in 1993 upon the initiative of Cassese, Ciampi and Barucci.

To re-define the State's HE funding channel, law 537/1993 eventually followed abovementioned arguments, introducing a simplified public funding channel, and, importantly, new arrangements for evaluation. Specifically, the State granted financial and organizational autonomy and simultaneously demanded that universities signal accountability through the establishment and work of internal evaluation units ("Nuclei di Valutazione Interna," IEUs from now on). In the law, politicians thus tried to construct a new relationship between State funding and universities' internal governance through the relationship between new autonomies and evaluation. IEUs would signal '[…] the correct management of resources, productivity of research and teaching, and the impartiality and well-functioning of administration' (law 537/1993, Art.5 comma 22). These features were in line also with the government's broader attempts to remedy the PA's inefficiencies, emulating the structure of SECINs introduced in the university sector. IEUs would thus become concrete assessors of universities' financial activities, and this was the specific declination given to evaluation in universities at this stage, one which was rather in line with the concerns raised by Giarda and others on how to best re-structure State funding through new institutional autonomies.

Indeed, the law did not communicate examples of evaluation criteria besides cost-benefit ones (law 537/1993), to be practically defined by university governing bodies: '[…] nucleuses determine the parameters of reference for control also through indications from general governing bodies, to whom they report at least annually' (Art 5, comma 22). As of 1993, then, political attempts to define the quality of teaching and research and the means through which it ought to be evaluated had this administrative accountability framing. This was largely because the State intended to steer universities towards resource waste reduction, achieve a fall in average unit costs, and diversify universities' income sources (Giarda, 1993;Catalano & Silvestri, 1992).

The law introduced the single lump sum budget, the "Fund for Ordinary Funding," commonly abbreviated as the 'FFO'. For clarity, it is important to point out that the FFO was not selective from the start. Although this was the intent, research funding selectivity actually began only following the first 2001-2003 VTR (Reale, 2008) carried out in 2004. This partially occurred because of the delays in setting up a body which would formulate the FFO's selectivity formulas, the Osservatorio (with Giuseppe Catalano as one of its 7 members).In this context, evaluation remained functional to the readjustment of the sector's costs. The ideal goal was to have State funds divided between a predominant share of "ordinary funding" ("quota base") to be distributed as a lump-sum, and a minoritarian share of a "re-equilibrium" part to be distributed on the basis of at the time still undevised quality-oriented selectivity criteria. As Banfi & Viesti have noted (2015), these prescriptions were not as straight forward in their implementation. Firstly, up until the Osservatorio's establishment only the ordinary funding pool was used, with very low percentages of selective re-equilibrium. "Historical funding" (i.e., funding through an incremental-approach) was thus implemented, with universities receiving funding depending on levels received in 1993. Thus, although the introduction of the new fund was meant to reduce inequalities in unit costs across the sector, it actually perpetuated them by '[…] advantaging further those universities which had obtained more funding and/or those which had higher staff costs for reasons of size (Banfi & Viesti, 2015, p.6). Secondly, the design of funding was in some respects contradictory. As I argue, it represents the dominance of an expenditure-saving approach over a quality-definition one. Incentives criteria – the law's "results-based-funding" – were in fact conflated with structural funding necessities, forming a new funding pool which confused the quality of activities with the costs of running a university, creating a *fusion of two aspects: objectives of the sector and instruments for their achievement* (Banfi & Viesti, 2015)[96]. These contradictions were made more complex by the fact that the results-oriented criteria for funding were still being developed in 1998, as substantiated by the Osservatorio's own report of that year (Osservatorio, 1998). Teaching

---

[96] See appendix J for a summary of the composition of the ordinary and premial parts of the FFO as set out by the Osservatorio in 1998

criteria were eventually constructed on the basis of an "equivalent student" measure, whose technical construction was not thoroughly explained in public reports[97]. With regards to research criteria, in 1998 the Osservatorio stated that '[…] there are still no adequate indicators for use' and delineated broad indications of its future work on research quality, starting with '[…] a data gathering exercise and a systemic evaluation of research activity to reach ratings. This will allow to fund this part of the fund[98] according to effective, rather than potential, research activity' (Osservatorio, 1998, p.6).

As the next section investigates, the sector's reform thrusts between the late 1980s and the passing of the 1993 law created widespread confusion on many fronts (see Banfi & Viesti, 2015). Combined with a peculiar interpretation of rewards, autonomy began to appear as a concession with strings attached, strings which however were still being knit – and still scarcely so 5 years after the passing of law 537/1993. In universities, autonomy could create new spaces for new initiatives in the internal management of State resources, which could be supported by the qualitative judgments of internal evaluation units. However, universities had to bargain their new autonomies with a body which was struggling to trace clear boundaries between State control and university powers in future funding and evaluation procedures. This left university agents facing new opportunities and not just new constraints. Indeed, this degree of uncertainty over future prescriptions on quality was also exploited by VCs as a means to step into the definition of quality and university governance following some years of debates, within the sector, on these matters.

## 5.2. The CRUI and its active minority

The Ministry for Universities and Research tried from the onset to formulate a comprehensive mission for the university sector and the country's research. University autonomy had been thoroughly discussed ever since the 1960s but never actually implemented (Corradi, 1998). In light of disillusions and failures, the new Ministry established in 1989 was really a chance to finally bring about concrete change once and for all. Political failures at giving university policy the due attention were increasingly evident by the end of the 1980s, when legislation on autonomy could no longer be postponed. As soon as this legislation passed, however, it appeared that it, too, failed to capture the interests of different university components. A redefinition of procedures and objectives was needed.

Stark criticisms were quickly raised to the new Ministry's first Minister, Antonio Ruberti, on the law which he actively advanced, law number 168 passed in 1989[99]. Ruberti was well aware

---

[97]The report can be found at http://www.miur.it/UserFiles/672.pdf
[98] i.e. the re-equilibrium part
[99] For the full text of the law see https://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario

of the predictable resistances which the 'autonomy law' faced (Talamanca, 2012[100]), but was keen to establish a new, independent ministry for higher education. He opposed an atomization of sectors and insisted for an organic coordination of the sector. The law was thus an instrument for ensuring the very existence of the ministry specifically for universities and research which could detach itself from the State and formulate tailored policies. Upon the passing of law 169,however the introduction of teaching, organizational and financial autonomy did not appear to translate into autonomy for university groups. The introduction of autonomy from the centre and the following attempts to impose it as a driver of organizational change sparked contrasting views on the net opportunities and limits which autonomy could bring about.

The CNR's researchers reacted by re-affirming their autonomy in contractual terms: they saw the institution of a minister combining university and non-university research as potentially hazardous because it would have meant that non-university researchers would have been subjected to the same labour contracts as university researchers and academic staff. In their view it was therefore wrong to unite national research and universities in the same ministry and policies as though they had the same structures and functioning. Their view of 'autonomy' was that non-university research had to remain detached from university research on a contractual level.

Between 1988 and 1993 university autonomy acquired diverse declinations because it was such a radical novelty for the sector. Sceptical reactions can partially be explained by the fact that Ruberti's law tried to reshape university powers without concretely defining their boundaries. Whilst the later 1993 framed autonomy in financial and administrative terms, for example, law 168/89 established first and foremost statutory autonomy. These two formal elements could be seen as being inevitably tied to the latter (especially from today's perspective),for there can be no freedom of setting a governance statute and define governance powers without a freedom to decide who those powers ought to be reserved to and what each ought to do.

Ruberti's proposals affirmed that university governance was in the hands of academic senates, which now had the opportunity of setting their own rules and missions. This also opened the doors to the autonomy of seeking alternative financial income and integrating external members into the governance bodies, an aspect refuted by the student population. Worried of the risks of de-democratization of decision-making processes and the increasing convergence with a privatized HE model, they reacted by affirming their objections to the law. Students' role in decision-making was in fact being side-lined in both law 168/1989 and law 341/1990, limited to a mere consulting role. This changed following a radical and national surge of students' collectives *(collettivi)* between December 1989 and April 1990. Starting from university occupations in Palermo, students gathered and formed assemblies in faculties all over Italy, hitting the major cities: Rome, Naples, Genova,

---

[100] Accessed on 20/10/2020 https://www.roars.it/online/la-vera-storia-dellautonomia-universitaria/

Torino. Students drafted criticisms and new proposals to the laws as well as to the operation of internal governing bodies. The strength of students' unity as well as their success at forming a communication network won the student body a victory which they can still enjoy today. In fact, Ruberti eventually ceded to student requests by passing an important amendment, namely the strengthening of the National Students' Council (CUN) as a key policy advisor whose views need to be taken into consideration before being centrally approved by the ministry.

Different groups of university agents thus posed challenges to the radical changes being advanced. Furthermore, the perceived limits of the 1989 reform were a crucial spark also for Vice-chancellors' own articulation of bottom-up governance. VCs were key leaders in a new debate seeking to formulate more specific definitions of autonomy. As I show in what follows, universities' turn to managerial governance and evaluation emerged importantly from this debate. VCs first comprehensively discussed autonomy and evaluation between 1989 and 1991 and began to actively construct its foundations between 1991 and 1993. The fact that these developments occurred previously to law 537/1993 and hence also previously to the rise of IEUs hints that the network emerging from the reactions to the Ruberti reform managed to exert significant influence over the future of evaluation principles and methods.

### 5.2.1. Autonomy and evaluation, two sides of the same coin

#### 1989 – 1991: Luigi Berlinguer and the Ruberti Reforms

As mentioned, the Conference of the Rectors of Italian Universities (CRUI) also objected Ruberti's reform for yet other reasons to those of researchers and students. From VCs' perspectives, the law could open up possibilities for becoming more influential in university governance. The autonomy to draft statutes, for example, meant that vice-chancellors could work more closely on the direction and development of their institutions. The 1980s closed with the Ruberti law as the '[…] first structural, comprehensive, and non-emergential reform to higher education' (Capano, 1998, as cited in Vaira, 2001). Still, many academics and VCs shared the opinion that the 168/1989 law was incomplete because of '[…] internal ambiguity regarding the scope of power transferred to the universities, […] the absence of a specific regulation circumscribing the room for action left to Universities, […] the resistance it met from the administrative bureaucracy and professors to its implementation (Reale & Potì, 2009)'. This opinion on how autonomy played out in terms of internal power balances was sparked by the fact that, paradoxically, the law gave universities a *too vast* and *undefined* autonomy: though now free to set up a whole array of new procedures and regulations, university agents felt that they needed more certainties over the State's role within a new framework of autonomy. In particular, they requested more specific guarantees on the full extent of their autonomy from the State's unregular and chaotic policymaking (Rebora & Turri, 2008). Further, whilst the 1989 law re-articulated universities' internal dynamics, it was not a

sufficient policy for remedying the sector's structural issues. Proposals did not appear comprehensive enough to address the sector's decline (Vaira,2011) and to remedy the specificities of low graduation rates. Seen retrospectively, those laws indeed appeared un-useful, given that as of 2003, for example, less than 13% of Italians aged 25-34 obtained a university degree (Vaira,2011). These aspects together created a bubbling context of debate over the future development of Italian higher education.

As a result of these perplexities on the law, a mixed group made up of academics and VCs gathered in Pontignano (in Siena) to discuss potential remedies. As Vaira (2011) notes, it is here that, in 1989, a reformist "advocacy coalition" formed. This "active minority" gathered to discuss the radical changes occurring, and eventually what had started as a debate on the reform's gaps developed into a *de facto* reformulation of its objectives and strategies altogether. Vaira's account of this specific group (2011) is a detailed insight into how policy alternatives were formulated outside government walls. These proposals are relevant also because they later entered ministerial corridors (Vaira,2011). Vaira identifies Luigi Berlinguer as the coalition's leader, the advocacy coalition's 'institutional entrepreneur' (2011) who mobilised agents around a reformist project that turned Ruberti's failures into an opportunity for constructing new ideas. During his active engagement at the Pontignano seminars, Berlinguer was VC at the university of Siena and also secretary of the CRUI, second only to its then president, Tommaso Scarascia Mugnozza. Berlinguer's leadership in Pontignano is key for capturing the CRUI's political engagement in the sector( Corradi, 1998), filling a gap in understandings of university governance ideas have also crucially developed from the bottom-up in the context of a transition phase. Two points emerge in this respect.

The first concerns how the advocacy coalition's composition shaped the development of a new and shared *"sectorial réferentiel,"* namely '[…] the perspective of an influential part of the academic world on the situation of the university sector and the proposals being made to face it' (Vaira, 2011, p.74). Berlinguer's leadership in fact gathered members which later covered political roles. This is the case for Romano Prodi, professor at the university of Bologna during the Pontignano seminars and then prime minister in 1996-1998 and 2006-2008. Similarly, Guido Martinotti, a sociologist, would become the guide of the "Martinotti group" wanted by Berlinguer for drafting educational policy when he was elected HE minister in Prodi's government in 1996 (Vaira, 2011; Pinto, 2009). The advocacy coalition thus indeed remained united following Pontignano, so much that its entrance in politics allowed it to build on ideas and projects developed during the seminars.

The second point concerns the coalition's work, which developed around two main goals relevant in the broader emergence of a managerial approach to university governance. Bom the

delusions of Ruberti's reforms, the group developed principles and practices for shaping university autonomy and evaluation. For the advocacy coalition, the new governance powers now being granted were so loosely defined that they opened too many possible interpretations of it (Vaira, 2011). Universities were not sure how this new autonomy could or couldn't be used. The 'oxymoron of the law on evaluation' (Vaira, 2011) effectively played out: it was felt that new freedoms introduced in the autonomy and evaluation law had to be better defined from the centre. This was paradoxical given the fact that the centre was meant to retreat and leave wider margins for space. Yet, it was an important reason for the very limited implementation of Ruberti's reforms (Vaira, 2011). Indeed, there were concerns that implementing new powers could backlash if the State didn't first do more to  concretely strengthen the long-awaited university autonomy. Pontignano's advocacy coalition therefore gave a go at defining the new autonomies in new and more specific terms in line with the university agents' perspectives.

Firstly, participants agreed that they could not foresee nor practically establish a more decentralized State-universities relationship based on the latter's autonomy if universities did not define what autonomy meant for them. Borrowing from European models (such as the UK), members felt that they wanted a relationship whereby the State would have a financial steering role and universities kept the rest of the responsibilities, especially those of internal management. Autonomy was the core prerequisite for all the new activities that universities were being asked to carry out in the future. This is particularly relevant in a political context like Italy's where incomplete laws are often repaired with successive laws and decrees, meaning that the implementation of new regulations depends on a complex deciphering exercise. Indeed, the confusion created by the ambiguity of the first autonomy law was  that it rested on old structures whilst introducing radically new principles, *de facto* diminishing its perceived transformative power.

Berlinguer's coalition thus elaborated shared definitions on the sector's objectives, amongst which also performance and quality. This was tied to the attempt of defining State and institutions' responsibilities (Vaira, 2011). The seminars defined external evaluation as a State responsibility and internal quality assurance as an autonomous institutional exercise. As early as 1989, academics had elaborated evaluation as a  management criterion for remedying lagging graduate rates, and, by extension, catch up to other European universities' productivity. The possibility that increasing student numbers would exert more pressure on teaching and research was addressed through discussions on efficiency procedures. If universities didn't' engage in this they risked aggravating their productivity (Vaira, 2011). Thus, the broader frame of defining autonomy allowed to develop shared ideas of types of evaluation channels to be established to monitor changes. Shared quality

standards were a much safer path than taking different development paths and possibly widening the social and economic inequalities in the sector.

The members and work gathered during the Pontignano seminars deserves much more scholarly attention, as this context says a lot about how a part of the university sector's academic community saw and confronted Italy's higher education policies in the 1990s, an aspect which crucially gives insights into the policies themselves, including on limits to their implementation. As has been shown in detail (Vaira, 2011), Pontignano's non-State, advocacy coalition drafted proposals on many aspects, including for example also Italy's integration in the European education context. It was in Pontignano that the first seeds of an academic-led articulation of the future of HE were fertilised, and were partially implemented politically in later years during Berglinuer's political career. Crucially, Berlinguer's leadership in these networks and his simultaneous affiliation to the CRUI helps to unpack and analyse how the CRUI's construction of evaluation fed from the reflections and the porposals discussed at Pontigano. In the hope of accelerating the conclusions of the Pontignano discussions, Berlinguer pushed the CRUI to act on them in its future work. A united academic front led by the institutional representatives could, it was believed, widen the scope for introducing and diffusing and evaluation culture. Simultaneously, it was the CRUI's chance to finally be more active and acknowledged as an actor in university policy following years of attempts to do so (see section 3.3.3).

### 1991 – 1993: Spreading an evaluation culture

Tommaso Scarascia Mugnozza's mandate (between 1987 and 1994) as the CRUI's president was characterized by a focus on evaluation, which turned into the committee's priority for years to come and had particular success during Luciano Modica's vice-presidency (between 1994 and 1998). The first steps were decided in a general assembly held in 1991 (see appendix K for a list of members), where the CRUI stated:

> '[…] the importance and the necessity that the entire system (public and private) be put in the condition    to be able to carry out evaluation of its activities as well as of the efficiency of the single universities […] to this end the Italian universities agree on the necessity of preparing an organizational model and methodologies as models that may allow to set off, within universities, internal evaluation units able to provide a support and assistance to both the Rector and the elected bodies that are concerned with the management of the university.' (Allulli, 1995, p.2)

At this stage, the CRUI was close to extraneous to issues around institutional management and evaluation largely due to the fact that , as discussed earlier, university governance was quite an informal process highly limited by the State's role in resource decisions (Rebora & Turri,2008).

As notions of autonomy, evaluation and self-governance began populating the academic and political debates, it became clear that these were key governance concerns, leading the CRUI to take them into its agenda. Because of its inexperience with the technicalities and implications of such notions, however, the committee hired an external consultant, Giorgio Allulli, who presented preliminary methods of educational evaluation. As discussed in chapter three Allulli was an active researcher at CENSIS and had importantly influenced the centre's methodological approach to the proposed school evaluation framework in 1991 (CENSIS,1991). At the CRUI, Allulli followed CENSIS' suggestion that the proposals for the school sector could also be integrated to HE. Along with his experience, he brought the CIPP model to the CRUI. In practical terms, this means tailoring CIPP to the university sector to start discussing possible self-evaluation and national evaluation methods and principles. This ambition first necessitated a better organization of the CRUI itself.

Before setting off research on evaluation, the CRUI set up an ad-hoc committee upon Modica's proposal. The committee was composed of VC delegated directly elected by VCs. The central idea was that members would try to mobilize university colleagues on evaluation issues and practices. The ad-hoc committee's role developed in two key ways. At the CRUI meetings, delegates discussed their university context. In universities, they worked to familiarise academic and administrative staff with the CRUI's evaluation projects. In particular, they engaged and assisted university agents in carrying out new tasks, such as filling out surveys and gathering statistical data on staff, students, and expenditure. This made VC delegates concrete conduits in the diffusion of emerging concepts around university evaluation. When the CRUI began to advocate for data-gathering exercises, some universities readily responded by establishing statistics offices working to improve data collection efforts (Rebora & Turri, 2008). This showed that this re-organization was succeeding in turning proposals into practice.

Having set up a new organizational frame, the CRUI began to set-off its evaluation projects. These embryonal experiments traced a path to build on in later years, with each successive report presenting more complex and far-reaching evaluation indicators. At a closer look, the very first reports show a very detailed articulation of how to construct internal evaluation in particular. Two key documents reflect this turn and the committee's ambition to actively spread an evaluation culture in the sector. The first document (CRUI, 1996) presented the first ever national set of university statistics, gathered by the Committee between 1992 and 1993 and published in 1996. It was in 1993 that the CRUI launched its annual survey of university data and indicators, "Documenti CRUI" (CRUI, 1996). The first survey was defined as a new instrument at universities' disposal: '[…] this document can be used by universities to comprehend their general position with respect to that of other universities' (CRUI,1996). The second document is a report

written between 1993 and 1994 and published in 1995. Here, the committee elaborated its specific understanding of evaluation indicators. It was the university sector's first and bottom-up articulation of a university-tailored path to internal evaluation which also discussed implications for external evaluation. The report, '*Organization and methods of the internal evaluation nucleuses in Italian universities: the proposals of the conference of rectors'*, (Allulli, 1995) was elaborated through Allulli's key inputs. It set out a strategy for initiating university self-evaluation and made proposals for IEU's Together, the reports convey the CRUI's political stance vis-à-vis the changing political-regulatory context. Examining the two reports in detail unveils the preamble of a new chapter of university governance which depended on the success of evaluation.

### 5.2.2. The first set of indicators

In a university sector which was rather inward-looking, statistics were usually extrapolated by Italy's national statistics institute (ISTAT) and mostly geared at analysing the impacts of public policies on universities. As such ISTAT's data mostly included levels of expenditure, employment, enrolment, and graduation rates. More sectorial and university-level specific information was absent, with calculations such as average costs across disciplines being carried out by third parties for specific studies by third parties. When the CRUI decided to undertake a nation-wide data gathering exercise, it therefore filled a substantial information gap. In its first experiment at devising indicators, the committee made use of the CIPP model, which guided the data survey on the 1992/1993 academic year (see figure 5). The data survey was wide-ranging, and gathered information from different organizational units: faculties, departments, degree courses and the university in its entirety. This unprecedented thrust had an overall response rate of 61.5% of the sector's public universities of the time (Allulli, 1995) and was key for the construction of the sector's first performance indicators.

The CRUI constructed performance indicators following Allulli's proposal to implement Stufflebeam's CIPP model (Stufflebeam, 1966,2000). The diagrams above show the use of CIPP categories for the survey sheets (figure 5) and for constructing indicators (see figure 6). Context data sought identified the 'target population' and the institution's key features (size, age etc.). Resource data asked universities to quantify incomes and expenditure (such as average expenditure per staff and student) as well as the quantities and types of students and staff. Research and teaching were categorised as both processes and products. Students' average score for each course year, for example, was considered a process type of data, whereas students' degree scores reflected product data. These categorisations drove the construction of indicators: indeed, the former set of data expressed how the teaching process developed throughout the course, whilst the latter showed the final outcome of that process by conveying the 'output'.
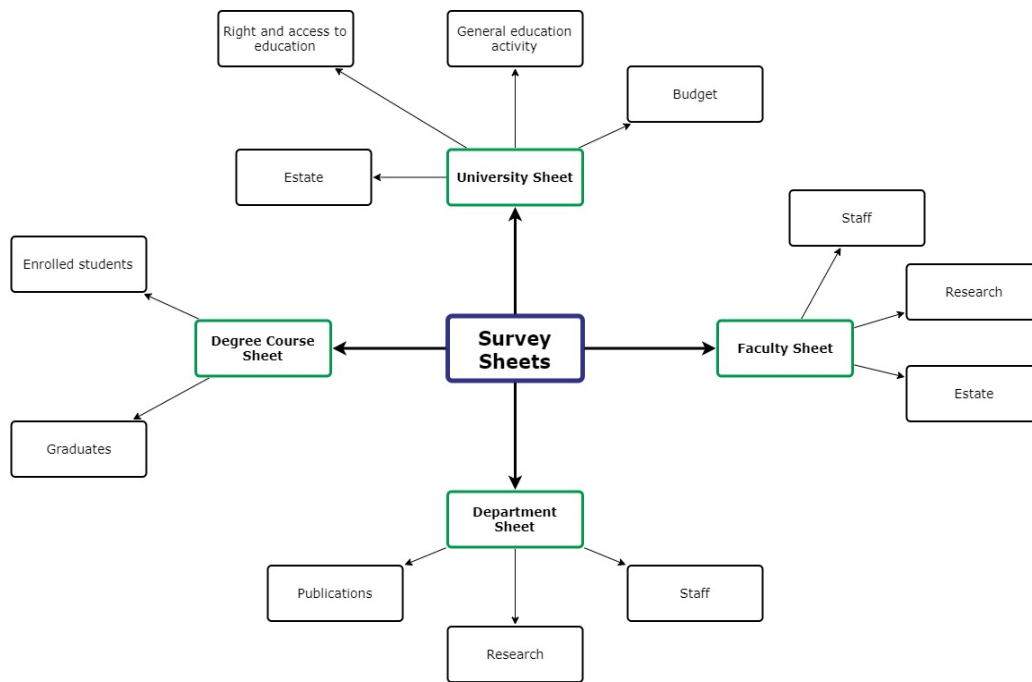
*Figure 5: The CRUI's survey sheets designed for the CRUI's 1992/1993 data survey (author's diagram from CRUI, 1996)*
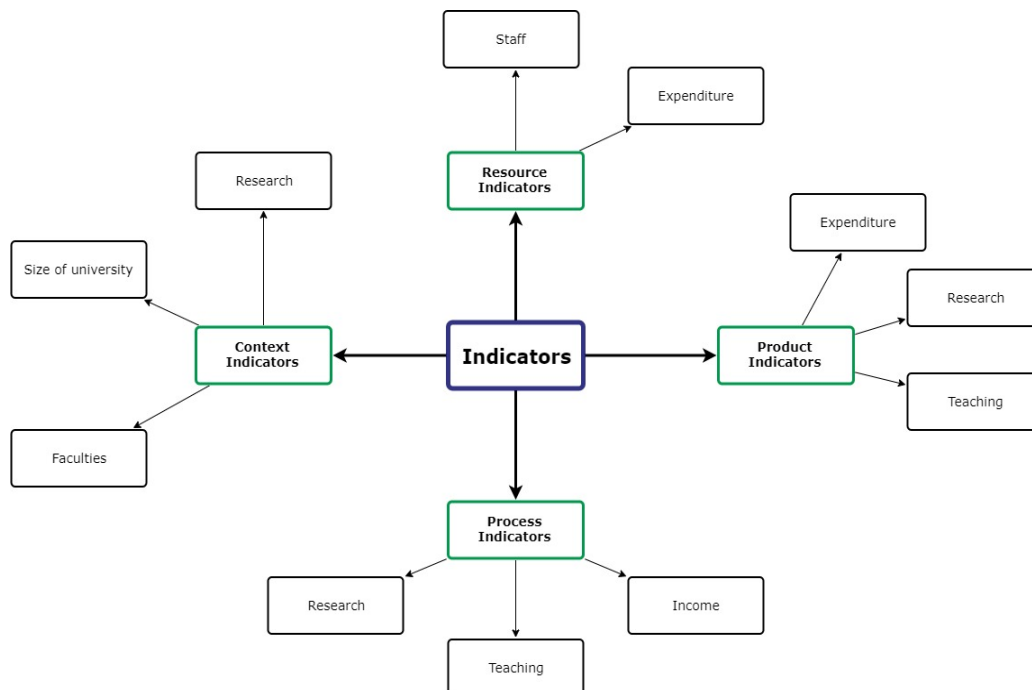


*Figure 6: Indicators devised by the CRUI between 1992 and 1996 (author's diagram from CRUI, 1996)*

200

This quantification of activities was a first go at quantifying staff performance in teaching and research. Examples are indicators such as 'average number of exams held by each staff' and 'average number of graduates per staff' (CRUI 1996). With regards to research, a process indicators devised was 'average number of staff for each department', and a product indicator was 'average number of articles published by each staff member' (CRUI, 1996). This more generic application of CIPP was further elaborated to produce more specific performance indicators (CRUI, 1996, pp.31-34). These were also expressed as averages and were thus very similar to the 'general' indicators. The key difference between 'general' and 'performance' indicators was the unit of reference: the former conveyed information on single universities and the latter expressed national-level information. The CRUI's first set of indicators was thus informative in scope rather than being strictly results-oriented. There was indeed no complex measurement of outputs in relation to inputs. Compared to conventional PIs geared at conveying the relationship between inputs and outputs often expressed in cost-benefit terms, the committee took a different approach and focused on sectorial analysis at this stage. This is reflected in its research PIs. To construct them, the committee summed the research products of each departmental faculties (excluding researchers) and divided it by the total number of universities participating in the survey (forty). The result was a national average for multiple types of research outputs, listed below:

– Average number of articles published by each staff member (average = 1.71)
– Average number of volumes published by each staff member (average = 0.25)
– Average number of conferences held by each department (average = 1.85)
– Average number of seminars per department (average = 9.25)

(CRUI, 1996)

Moving on from the 1992/93 averages published in 1996, the CRUI parallelly tailored this averages approach, proposing new research indicators which were added to the "resource" and "product" categories. Within "resource" indicators, a key element added was the source of income that each department/institute received for research purposes, and, in particular:

– Total income received
– 40% and 60% funding (the State's two channels for funding university research)
– Other public funding received
– International funding received
– Private funding received

Within the "product" indicators, the CRUI included results-oriented considerations in order to quantify *the type and number of outputs* in different departments and university research institutes. These new indicators were:

- Number of citations that appears on scientific journals per staff
- Number of contracts/conventions secured with public entities per department
- Number of contracts/conventions secured with private entities per department

(Allulli, 1995)

Through this first set of indicators, the CRUI used a *qualitative* approach by relating inputs and outputs in a way which escaped a strictly cost-oriented dimension which as we have seen characterized the CVCP's indicators. Surely, the CRUI *quantified* university inputs (staff, expenditure, students) and outputs (graduates, quantity of research output according to type) through the CIPP model to explore their relationship. However, this quantification exercise – as the definition of performance which the CRUI presented – did not yet serve a cost-efficiency purpose aimed at analysing resource use specifically. The committee's definition of performance at this stage served more of an informative scope, using averages as a means to identify similarities and differences between universities. This was in line with concerns over the sector's high student mortality rates, and indeed the CRUI's reports paid extensive attention to that data expressing numbers of students and their stay at university. The "context" indicators were thus considered central for understanding these context-specific factors shaping educational outcomes, in line with CENSIS' approach to school evaluation (CENSIS,1991). The major simplicity of school learning and curricula compared to university processes in fact makes the identification of the final 'product' (educational outcomes) easier. In higher education, university courses usually bring together multiple types of material and immaterial aspects which are not easily conducible to degree course content. Educational outcomes in higher education are thus *acquired skills*, that can hardly be exhaustively measured at a first go. The CIPP model was a first step to get as close as possible to an exhaustive assessment of HE's final 'product' (graduates) before during and after their university experience, with "context" and "process" information geared at identifying factors which could explain dropout rates. By the end of its data survey, the CRUI constructed a set of over seventy indicators which compared universities' activities at almost 360 degrees for it had gathered a very detailed set of wide-ranging institutional data (Rizzi & Silvestri, 2002).The CRUI effectively provided the first ever national snapshot of the university sector in statistical terms. Additionally, its indicators, were the very first systematic turn to a quantification of university activities, a step which opened the sector's "black box". It was through the committee that this databank now existed and was made accessible, opening up a new array of future and then still undefined uses.

The information and its elaboration into indicators came at a crucial moment in the broader context of the sector's transformations. The material gathered between 1991 and 1993 in fact

shaped the CRUI's political stance on the autonomy law 537 of 1993[101].. As I argue, the committee reacted innovatively to it because it conceptualised the use of indicators in ways different to those proposed by the legislators. In its view of the roles and structure of IEUs introduce by the law, the committee did not merely integrate an 'ex post' role for the nucleuses geared at verifying the optimal use of resources. Instead, the committee articulated evaluation in terms of its use for multiple dimensions of university governance, going beyond what universities were being asked to do. Its proposals (Allulli, 1995) indeed turned evaluation into a device for decision-making by emphasizing that IEUs would serve to make also ex-ante judgments and not simply ensure financial accountability and transparency. In doing so, it positioned itself at the foreground of future university governance as that body which could implement new decision-making procedures in the sector.

### Internal evaluation: a new practice

The CRUI's second report used the law's vague prescriptions on evaluation as a possible opportunity to present its own governance approach to evaluation. This emerges in both how evaluation was defined and how the committee later practically constructed it. The committee turned its criticisms of political demands into a premise for advancing its previous work on indicators and specifically on internal evaluation. It argued that policymakers were unable to provide a well-rounded notion of quality and efficiency, suggesting that this void left universities unequipped with practical strategies for setting up IEUs. To this criticism, the CRUI aimed to act as a 'gap-filler':

'[…] in the legislative text *the word quality is not cited*, and the tone is definitely oriented towards and economic type of evaluation […] Despite the lacking legislation and the slowness of application mentioned above, there is, in the Italian university system - and especially at those levels responsible for strategy and management - a large consensus on the necessity to rapidly reach a mechanism of evaluation of university activities, institutional, of research and teaching, and administrative.' (Allulli,1995, emphasis added).

It presented a wider range of indicators, going beyond an "economic type" of evaluation to provide a sectorial and common definition of quality. Indeed, it emphasized that:

'Although the formulation of the task [correct management of State resources] is shared not just on the basis of common sense but also for its political meaning, it appears much more

---

[101] https://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario

controversial when translated on an operational level. What is meant by "correct" management of public resources? What is the parameter of evaluation for "correctness"? […] Therefore, what appears as an objective to be verified (correct management) in reality is not such, because the enunciation lends itself to extremely diversified interpretations.' (Allulli,1995, p.13)

The problematization of weak operational guidelines was very legitimate indeed, it signalled that though universities were being asked to change, the State was failing to express what exactly they had to change *into.* The CRUI did not however treat this as an impediment to its role. It was seen as further evidence of the concerns discussed in Pontignano, and used criticism as a means to intervene directly in defining quality and evaluation. Picking up earlier discussions on autonomy, the committee framed quality definition as a corollary of universities' new autonomies, especially financial autonomy. Autonomy was a new and major responsibility, for '[…] the growing weight of autonomy […] will have to be sustained and qualified through a process of permanent evaluation and self-evaluation, also to avoid future excessive inequalities on the national territory' (Allulli, 1995 p. 3). For this reason, autonomy and evaluation were seen as two sides of the same coin. Responding to the 'growing weight' of financial autonomy implied identifying shared parameters of 'correctness'. It is precisely the fact that legislators were so ambiguous in defining evaluation that helps to see the CRUI's innovative action. All but adaptive, the committee responded to the law by asserting its definition. Its later approach was thus by no means an inevitable articulation shaped by political guidelines. All the contrary, it was an active reformulation of political demands made possible through a criticism of politics' own inexperience in evaluation. In effect, the law did not at all use the word quality, leaving indeed wide space for interpretation on its methodological articulation for the internal units being proposed in universities. Besides defining its necessity of seeing accountable behaviours, the law left wide margins for self-articulation of evaluation practices.

As a result of this the CRUI eventually presented a university governance model more complex than that set out by the legislator, proposing that IEUs make use of its previously devised indicators. This context could in fact become an opportunity to concretely diffuse the committee's stance on evaluation. In a managerial vein, the CRUI framed indicators as crucial for internal governance. Again, this was in contrast to the law which had only referred to the use of indicators for governance as far as financial accountability was concerned. For policymakers, the priority was that governing bodies provide resource information so that the Osservatorio could then formulate funding formulas. Further, the legislator had to safeguard organizational and financial autonomy being introduced by the law. The CRUI, instead, presented a circular evaluation model which included all of the sector's agents and bodies. In the model, universities' governing bodies would

work with IEUs, and the two would in turn collaborate with external bodies – the Osservatorio and the Ministry – to provide information on their performance. The scope for indicators in the circular evaluation model was dual: to aid internal governance decisions and to allow external users to assess the sector for allocation purposes. Having set its view of university governance as reliant on the use of indicators (Allulli, 1995), the CRUI then proposed the ones which it saw most fit for implementation: namely the 1992/1993 survey indicators which it had not yet published.

For IEUs, it suggested that '[…]indicators on *resources and processes* will allow to verify "the correct management of public resources", and "the appropriate functioning of administration" whilst the *product indicators* naturally refer to the "productivity of research and teaching"' (Allulli, 1995, p.14). Further, 'context' indicators would capture specific differences and similarities between university departments, '[…] provid[ing] a more comprehensive and balanced analysis of the system' (Allulli,1995, p.14). The CRUI argued that the IEUs' functions could be readily initiated by adopting its indicators (Allulli, 1995), for their horizontal use across the sector's IEUs would make them comparable and also allow to fulfil the use of evaluation for internal and external governance. IEUs were defined as both evaluating and evaluated units which for this reason necessitated shared meanings and possibly homogenous indicators. To encourage IEUs' establishment and work, the CRUI lay out those *operational* functions missing in the law. In so doing, the committee drew out its managerial approach, for IEUs were advised to devise benchmarks that would assist governing bodies in assessing universities from within on a variety of aspects:

'[…] it is the governing bodies of the university which have the duty of defining he objectives which they seek to pursue, whilst the IEUs ought to identify the indicators or enquiries for verifying their achievement.' (Allulli, 1995, p.13).

The neutral and objective skills of the newly emerging units allowed to justify that internal evaluation was an instrument of decision-making to be used by governing bodies. Simultaneously, this elevated the role of governing bodies, reserving to them the power of objective-setting. This subjective form of power could not be, arguably, delegated to IEUs, charged instead with a technical-objective type of power. However, the fact that IEUs were advised to use the CRUI's indicators crucially blurred the key distinction between academic objectives-setting and technical objectives-assessment. For whilst the CRUI stressed the distinction between the governing bodies' subjective powers and IEUs' technical-objective role of analysing goal achievement, proposing that IEUs make technical evaluations using criteria proposed by governing bodies could effectively extend subjective power over to the 'objective' sphere. Vice-chancellors' turn to evaluation can

indeed be conceptualised as pragmatic use of the context with an underlying intent to acquire new forms of leverage in university governance —arguably both within individual universities and at a sector level. Within universities, the new evaluation framework being proposed by the CRUI allowed vice chancellors to reaffirm their crucial role within governing bodies as objective-setters as well as 'evaluators' of their universities' development. More interestingly, internal evaluation was treated as a means for increasing the CRUI's role within the sector. This could occur through the institutionalisation of its own approach and indicators. The CRUI insisted on the importance of its role in evaluation. A further reason for what I here define as the committee's search of legitimacy in evaluation is the fact that there was an approaching limit to its influence in the sector. The law had in fact stated the establishment of the Osservatorio. Although this body only started functioning from 1996 onwards, as of 1993 this was not yet known. As a result, there was the concrete possibility that the Osservatorio's entrance in the sector could undermine the CRUI's previous work, for example by introducing different indicators. The Osservatorio would have benefited from greater legitimacy than the CRUI in its demands, for it was a state-devised body which was intended to follow ministerial guidelines.

The Osservatorio was meant to assess universities' adherence to national development plans set by the Ministry  with the aim of assessing universities' adherence to the national development plans set by the university Ministry (see Art.5, comma 23, of law 537/1993). It also had to define selectivity criteria for allocating the re-equilibrium part of the FFO. This represented a potential conflict between the CRUI's and the Osservatorio's indicators. The CRUI therefore tried to ensure that the Osservatorio's methods would turn out favourable to its own. As I argue, the CRUI considered this when insisting on its role and advocating for the implementation of its proposals. Legitimacy could importantly be built by acting as a 'first-stepper' in defining evaluation indicators and dynamics, and if universities practically acted on specific guidelines, it would also have made it more difficult for later bodies to modify procedures being set off or to introduce new methods and perspectives on evaluation. In its report, the CRUI emphasized examples of international experiences demonstrating that rectors were the legitimate and typical leaders in university evaluation. The VNSU in Holland, as the AAU in England[102] '[…] coordinate national evaluation and […] provide the methods and instruments of analysis […] they have a significative presence at a strategic and methodological level' (Allulli, 1995, p.4). These efforts succeeded, as the Osservatorio struggled to be set up and as the IEUs were instituted and began working through the committee's proposals. Indeed,  87.5% of universities which implemented internal evaluation following the 1993 law used CRUI's 1995 indicators (Osservatorio 1997, doc 5). The CRUI

---

[102] I was keen to analyse these aspects further but have been unable to do so partially due to the lack of a proper CRUI archive.

encouraged universities to continue on this path, asking IEUs to equip themselves with methodologies and follow its guidelines if they wished to keep up in a changing context (Allulli,1995). Meanwhile, the Osservatorio began criticising IEUs' work, creating new tensions. Its monitoring attitude signalled that it sought to scrutinize universities' decisions, as occurred in one of its first reports:

'[…] far from verifying efficiency and efficacy, it appears that even the respect of the legislative requirements, such as the establishment of the evaluation nucleus and the presentation of an annual report, have not been seriously considered, maybe due to the lack of explicit sanctions'(Osservatorio, 1997a, p.29).

Given that the work of the nucleuses depended on the CRUI's 1995 guidelines, these criticisms were also implicitly directed at the CRUI. The Osservatorio's pressures are partially explained by the fact that its members were keen to accelerate the work on selective fund distribution (Osservatorio,1997b) and implement the funding formula (Osservatorio, 1998) . Yet this could not be done without feedback from universities. The Osservatorio had to strengthen its coordinative ability given that the sector's three levels pursued different objectives. The ministry at the top had expenditure and development targets, the universities had attainment targets, and the IEUs within them had to work on an internal balance between resource-use and educational targets. For the Osservatorio, IEUs needed to do more to define institutional goals, otherwise its own function of assessing their achievement from an external position could not effectively play out. The CRUI reacted to these concerns by continuing its autonomous work on evaluation and strengthen its indicators by updating its set. In the process, it interacted with new agents and explored new university evaluation research. Between 1995 and 1999 this exploration path allowed it to acquire new skills, ramify its network of collaborations with industry, and carry out new projects. One such project was CAMPUS.

**Project CAMPUS**

Shortly after Romano Prodi's election as prime minister in 1996, the CRUI – now led by Paolo Blasi, VC of the University of Florence – interacted with this government which it saw favourable to its standpoint on HE development. It is worth recalling that Prodi was part of that group mobilized by Luigi Berlinguer at Pontignano (Vaira, 2011). When Luciano Modica presented the committee's work at Palazzo Chigi (the government headquarters in Rome), Prodi agreed that Italy's position was still dramatically lagging compared to other European universities, so much that

if there had been a Maastricht treaty for universities, Italy would have been excluded from it. As the Pontignano *advocacy coalition* (Vaira, 2011) had already highlighted, universities' productivity levels since 1989 were still extremely low. The sector necessitated urgent and major restructuring, as European universities were well into their evaluation experiences. As a result, and now with Prodi's support, Modica sought opportunities. Him and Rodolfo Zich –VC of Turin polytechnic between 1987 and 2001– came up with an innovative way to secure funding for evaluation projects from the European Union. Zich and Modica thus used Ruberti's recent introduction of the three-year degree to suggest that Europe's Social funds could be used to provide to improve the setup of these degrees and the professional training provided as part of these courses. Modica and Zich mobilized in the EU corridors to argue that their planned use of the funds coincided with the social fund's objectives and, following tedious negotiations, funds were eventually assigned to the CRUI, who invested them in the 'CAMPUS' project which stands for "Corsi Avanzati Mirati alla Preparazione Universitaria orientata agli Sbocchi professionali", advanced courses geared at the professional outcomes of university preparation). The project was officially set off in the academic year 1995/96 and prospected to continue also in academic years 1996/97, 1997/98 and 1998/9. CAMPUS was also financially supported by other parties besides the EU,: Confindustria (the Italian industrial confederation), Unioncamere (the trade chambers' representative association), ENEA (a public research body for environment, sustainability and competitiveness technologies) and different regional funders (Modica & Stefani, 1997).

In many ways, CAMPUS was an Italian articulation of a study of the European Commission drafted in 1991 and launched in 1994 (CRUI,1995). For this reason, when CRUI launched CAMPUS in 1995 it managed to catch up with the Commission's work by following its project. The two studies had important similarities. The Commission's project was geared at searching a shared method for the external evaluation of two key disciplines, engineering, and the arts, in Europe's more advanced countries: France, Holland, Denmark and the UK. Each country would draft proposals to be collectively discussed, shared, and implemented (CRUI,1995). To ensure the success of CAMPUS, the CRUI collaborated with the HE Ministry and piloted external evaluation at Bari and Turin polytechnics for engineering and at the University of Udine and Tuscia for the arts( the latter specifically for courses on the preservation of Italy's cultural heritage). The CRUI addressed the productivity issue by externally assessing employment outcomes of the two extremely different disciplines. Confindustria's inputs were key in this, for it advanced business and industrial quality management practices to be used in the CAMPUS framework. It suggested the use of the Total Quality Management (TQM) model, used by Italian industrialists to ensure the achievement

of the ISO 9000 standard[103]. In general terms, the TQM model is often summarised by 'turtle diagrams' reflecting how multiple stages of the production process (production, marketing, product design etc.) are evaluated through quality-oriented questions. In this way *quality is managed*, and controlled through a step-by-step verification for the accomplishment of a final, satisfactory quality product without defects. From a quality control perspective, the core of the Japanese management approach in industrial production is the continuous improvement of production – in Japanese 'Kaizen' – through the feedback system which the model itself provides. Besides many other grounding differences which will not be discussed here, a key one between the cost-attentive management frameworks of PPBS and TQM's is the latter's use of a *qualitative* feedback system. When the CRUI turned to Confindustria for advice, the two created a peculiar innovation for university evaluation (and specifically for courses' evaluation), namely the construction of a grid of quality-oriented questions inspired by the TQM approach. The rather general yes/no questions concerned the content and structure of courses, to be filled out by students participating in the courses being evaluated by the CAMPUS project (Modica & Stefani, 1997)[104]. The innovative element brought about by this grid was not just the appropriation of TQM approach for university evaluation. The partnership with Confindustria also led to the construction of a 1-4 ranking scale for external evaluators keen to hire graduates, such as industrial groups themselves. In this way, CAMPUS provided new internal and external evaluation criteria. For the CRUI this was a radically new path, '[…] the uniqueness of the grid […] has permitted to *quantify qualitative evaluation*, just as occurs in the business sector.' (Modica & Stefani, 1997, p.10) Importantly, it was a possible solution to the committee's earlier attempts at standardizing internal and external evaluation in a sector which still had no concrete external evaluation framework. The committee thus experimented this in CAMPUS by providing internal and external evaluators with the same evaluation checklists '[…]external evaluation has been conducted using a list identical to that used by self-evaluators' (Modica & Stefani, 1997, p.10). The steps taken through the project show how the committee was keen to use concepts different to those of CIPP. The far-reaching experimentation of course evaluation was a way to once more step into a gap of the sector, which had few instruments for this type of assessment, and the CRUI was acting once more at the forefront of new experiments on evaluation. Its success was evident when CAMPUS continued for longer than initially planned. CampusONE, its successor, was carried out between 2001 and 2005 and now involved over 500

---

[103] The ISO 9000 standard owes its name to the international organization for standardization, founded in London in 1947 following a seminar at the Institute of Civil Engineers in 1946. If businesses certify their production with ISO 9000, they effectively communicate adherence to a set of management principles dealing with quality assurance and a meticulous inspection of the production process more generally. Indeed, the ISO 9000 standard is very general and can thus vary enormously in different production sectors. See https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100080.pdf for the ISO 9000 management principles.

[104] My analysis of the CAMPUS project is thus exclusively based on the explanation provided in Modica & Stefani, 1997.

degree courses and 50000 students enrolled in a much wider variety of disciplines. These figures testify the project's reach and drastic expansion, the continued collaboration with industry and the success compared to the 70 engineering and applied sciences courses and 5000 students covered in the CAMPUS project between 1995 and 1999 (Carpita, 2016; CRUI, 2002; Tosi, 2003). As buzzwords of innovation and entrepreneurialism proliferated in the sector, so did the CRUI's managerial approach. In CampusONE, the committee further developed the TQM approach by advocating the use of William Deming's approach to the "Shewart Cycle" (Deming, 1967, 1982), characterized by the division of quality management processes into the Plan-Do-Check-Act (PDCA) categories (Mirandola et al.,2003).

This managerial experimentation was different to the CRUI's initial managerial approach to university governance. In the early 1990s, it manifested its managerial character by proposing a top-down decision-making process oriented around VCs' use of evaluation indicators and their strategic role in objective-setting. Now, the managerial approach was being characterised by the integration of industrial management evaluation templates and the broader adherence to the notion of the entrepreneurial university development model where university-industry ties are crucial. These new skills allowed the committee to return to research on indicators and seek to apply some of the new insights into its definition of university quality, shaping it through evaluation indicators. Indeed, the CRUI soon piloted the development of managerial skills in the academic body, hoping to reach a point where a large part of evaluation would be carries out internally at the root of the university production process: teaching. The guiding principle was that teaching activities needed more articulated feedback devices if outcomes were to improve. The CRUI provided practical steps by setting up training courses for future "teaching managers" ("manager didattici"). Overall, it recruited around 1,000 members mostly from a pool of technical-administrative staff in universities as well as through a public competition ("concorso publico")[105]. Teaching manager trainees would act as 'zippers' between the student body and teaching staff by assigning one teaching manager for each course of study. Teaching managers interacted with both the staff designing the courses and the student body receiving them, gathering opinions on perceived successes and failures of courses. This first feedback-gathering activity supported teaching managers' second role, that proposing changes to course content on the basis of feedback. In this way, the teaching managers would act as impartial auditors for single courses. Though they had no or little teaching experience because of their often very young age and/or administrative background, the CRUI considered teaching

---

[105] From conversation with CRUI members

managers as expert figures suitable for assessing the quality of teaching processes because of the fact that they used a mix of objective criteria and subjective feedback[106].

The CRUI communicated the developments of its pilot project to the European level, signalling that Italy was playing its part in evaluation and slowly catching up to other countries. Effectively, the CRUI's leadership was applauded as a significant effort in evaluation, such that the committee was quickly cited in European reports as Italy's key reference organization ( Frazer, 1997). Luciano Modica's active mobilization was bearing it fruits, equipping the committee with new inputs, agents, and practices around both internal auditing and external evaluation which were often readily experimented in universities. The CAMPUS project was a way for "exporting" the Italian experience into the European debate, in the hope that this would also allow the CRUI to lead future evaluation projects. (Modica & Stefani, 1997). The committee was indeed very proactive when it escaped the macro-analytical dimension of universities' productivity issues. Rather than merely re-expressing its preoccupation for Italy's backward situation and referring to the structural constraints shaping it, the committee made practical steps to understand if and how evaluation could be an analytical framework for deconstructing the complex intersection of structural and university-specific factors. As emerges from the CAMPUS project, the committee's evaluation framework was heavily influenced by the inflow of a managerial declination to evaluation, such as that of the European Commission's project and the interaction with Confindustria. The Committee was eager to continue along this path and build its expertise on evaluation, aware of the unprecedented detail and reach of its experiments. This required that its members also acquire new skills and knowledge on evaluation practices, a step taken by Modica, for example, who attended a training on university evaluation in Glasgow in 1998 and learnt about the functioning of the HEQC and the QAA.

The committee's use of different evaluation experiences is reflected in the indicators that it published during CAMPUS. Whilst the committee's first indicators were an important yet general snapshot of university performance, by the end of the 1990s the CRUI itself considered them close to obsolete for a constantly evolving context (CRUI, 1997). Its intention to lead university evaluation was now explicit, it aimed '[…] to acquire comprehensive indicators on the entire university system as well as efficiency and efficacy indicators for evaluating governance actions at the single university level', it applauded the fact that '[…] internal evaluation units have already widely used CRUI data, and now also the national Osservatorio proposes that units use the indicators elaborated by the CRUI' (CRUI,1997, pp.14-15). Under this enthusiastic thrust, the fifth volume of the "Documenti CRUI" series (CRUI, 1998; Modica & Stefani, 1997) reflects what I

---

[106] I have not been able to carry out more detailed research on the development of the teaching managers initiative, such as the courses audited and the universities in which it was carried out.  This is because Modica himself admitted that the

here define as the committee's "renewal" at the end of the 1990s, comprised of two key developments.

Firstly, the new indicators were more tailored to universities' objectives. Given the persisting urgency of remedying low graduation rates, the CRUI constructed the "ideal student" indicator, a measure of '[…] the 'intermediate' productive capacity of the university compared to the finished product [the graduate]' (Modica & Stefani, 1997, p.23). The idea was to calculate the number of students placed below, within or over the standard time of graduation – i.e., the "ideal student" – and more specifically convey the efficiency university teaching and staff performance. This indicator reflected the committee's appreciation of quantitative analysis, now considered more important than in the early 1990s. Indeed, the ideal student indicator could allow to calculate both student and staff performance in different universities. Additionally, ideal cost quantities could be devised for the 'ideal student', in this way allowing university administrators to quantify resource 'wastes', i.e., those resources being repeatedly invested on students that did not graduate in time or graduate at all. This echoed the Osservatorio's attempts at devising a 'standard cost' for full-time students when developing a new funding formula. The CRUI's advancing skills emerged also from the new proposal of an "intermediate productivity indicator", which, similarly to the "ideal student" indicator, would trace students' achievements at stages 25, 50, 75 and 100% of their course duration, calculated on the basis of the exams passed by students (and not on the expected duration of the course).

Secondly, this later report paid much more attention to resource use. It introduced new university questionnaires with more specific questions on resource use in relation to staff and student quantities. Differently to the previous lists of indicators, the committee now began to develop cost information and gather micro-level resource information. The 'fundamental data' of this kind could answer impingent questions such as 'how much is being annually spent, and for which objectives? How does this reflect in universities' national and international quality rankings?'. This would allow Italian universities to face that broader quality competition on the same axis of more developed European university sectors (Modica & Stefani, 1997). Finally, the CRUI also proposed that IEUs devise an indicator for conveying universities' capacity to attract non-public funds, in line with its intention to push the sector to open up to the entrepreneurial model of university-industry collaboration being set off through CAMPUS.

Finally, and in stark similarity with CVCP's National data study proposals, the CRUI suggested that Italian universities make use of management information systems because new organizational models were necessary. Performance information on costs and activities was defined as 'a new political instrument of the Vice-Chancellor' for governing universities (Modica & Stefani, 1997). By extension, it was up to the CRUI to transmit this information to external bodies

(MURST, ISTAT, industrial stakeholders, future employers) and ensure that these be kept up to date in making their judgments on resources and collaborations with the university sector (such as trainings within businesses in the case of engineering courses).The CRUI described its new wider network which ranged from the European Commission, to the CRE, to the national industrial representative associations and political groups (Modica & Stefani, 1997), signalling that it was determined to develop universities' governance and that it was methodologically skilled to do so. This exercise of consensus-building coincided with the CRUI's attempts to digitalise data collection and to spread university information more quickly and extensively outside the sector. The CRUI had a newly born collaboration with ISTAT which, through an informal joint working group, was developing the existing national database to '[…] predispose an information system for external evaluation of the university system' (Modica & Stefani, 1997, p.15).

All in all, by the late 1990s the CRUI increasingly tried to use external groups and their advice for developing new evaluation indicators, and especially external evaluation ones. At the same time, however, this thrust kept research evaluation indicators out of the broader developments. In part, this was because funding received by the committee was destined to course evaluation specifically. Also, the CRUI was hesitant to engage in research quality indicators because it shared British critics' objections to the methods of the first RAEs.

### 5.2.3. A disciplinary divide recast through evaluation

This hesitancy soon changed, and turned into an active search for an alternative, Italian turn to research evaluation in 1998. This turn was influenced by two key factors: a new law which re-structured disciplinary divisions and the Osservatorio's new funding formula. Between 1997 and 2000, Italian universities lived the decisive shift in research quality evaluation. On one side, the Osservatorio was developing its funding formula (see section 5.1) and insisted that selectivity was behind the corner. On the other, politics was stepping in after a few years of silence. When Ministerial Decree (D.M.) number 152 (passed in June 1997) intervened on research[107], both the CRUI and the Osservatorio intensified their work on research quality indicators. The law merged those scientific-disciplinary research fields ("settori scientifico-disciplinari") which had been introduced just a few years earlier in 1994[108]. From a situation where university research was categorised into 441 scientific-disciplinary fields, Minister Berlinguer's law (with approval of the national university council) introduced 17 broad macro-sectors which aimed to categorise research fields into broader disciplinary headings (see appendix L for the list). This reconfiguration was part

---

[107] Law 127/1997 (specifically article 17, comma 99) https://www.gazzettaufficiale.it/eli/id/1997/05/17/097G0161/sg was active from 18.5.1997. The Ministerial Decree (D.M.) number 152, passed shortly later on 23.6.1997, set out a new list of scientific-disciplinary sectors. https://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario

[108] The presidential decree at its time intervened on a law which established scientific-disciplinary fields used to identify university courses. Specifically, the decree modified article 14 of law n.341 of 19.11.1990). The decree can be found at https://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario

of an administrative law (law n. 127/1997)[109], which sought to 'streamline administrative activity' as well as 'the processes of decision and control' and had significant qualitative effects on how research evaluation was later formulated. Further, the 1998 legislative decree (n. 204)[110] introduced the country's first national research evaluation body, the CIVR ("Comitato per l'indirizzo della Valutazione della Ricerca"), charged with evaluating both university and non-university research. Through the CIVR the government emphasized the importance of a national research planning effort, making the first concrete steps at steering it 'from above' (Reale, 2008; (Rebora & Turri, 2008). As such, research quality evaluation now appeared as a central exercise which would be guided by the Ministry itself. Not only was CIVR located within the Ministry, but the later establishment in October 1999 of yet another body for university evaluation, the CNVSU ("Comitato Nazionale per la Valutazione del Sistema Universitario"), made this perception a concrete possibility[111]. For attentive audiences such as the CRUI and the Osservatorio, these laws opened up a new possibility to articulate the governance and evaluation of university research. This possibility meant intervening from below to make their views and evaluation methods visible and institutionalise them in the newly emerging bodies. However, in a context filled with notions of rationalization, simplification and efficiency, the emergence of these research evaluation bodies also constrained such actions. The CRUI needed to act quickly if it sought to pre-empt centralized evaluation. **Both the methods proposed by the CRUI and the Osservatorio between 1998 and 1999 reflect each organization's struggle at institutionalising its methods in light of the Ministry's centralizing force.** This section explores what emerged from this tension, arguing that the CRUI made a very radical proposal on research quality evaluation criteria.

The Osservatorio first reacted by recognising the incompleteness of its previous evaluation proposals, defined as '[…] a series of coarse simplifications which will be overcome when the [evaluation] experience will be repeated' (Osservatorio, 1999, p. 2). It defined research quality as importantly based on research outreach and diffusion. The methods which it proposed were aimed at '[…] evaluating the *means of diffusion* of scientific publications produced, without evaluating their intrinsic quality'(Osservatorio, 1999, emphasis added). In fact, evaluation was voluntarily

---

[110]The decree establishing the CIVR dates 5 June 1998 and was published on 1 July 1998. However, this decree only formally established it, delineating as follows: '[…] the committee is composed of no more than 7 members [and] operates for the support of quality and improvement in national scientific research […]. It has the role of indicating general criteria for evaluating research results and of promoting experimentation, application and diffusion of methods, techniques and practices of evaluation' (law decree n.204/1998, art.5, comma 1). Indeed, the members of the CIVR were officially nominated in November 1999 through ministerial decree n.643. Hence, the CIVR was officially active one year and five months following its establishment.
[111] The CNVSU was established by the law n.379 of 19 October 1999, but its functioning was only officially formulated by the legislator in April 2000, through ministerial decree n.178, and, more importantly, the members of the committee were nominated in May 2004, through ministerial decree n.101. This means that in practice, the CNVSU was not operative between 1999 and 2004, although it published some reports.

detached from a quantitative analysis of research output. To evaluate research diffusion, the Osservatorio advanced a 1 to 5 rating scale, where values were assigned according to types of publications and their national or international reach. To accurately convey research quality along this scale, research outputs were categorised specifically in terms of publications in a foreign language and in foreign journals/editions. Book publications in foreign languages were therefore the highest-ranking, whilst Italian conference publications or working papers were the lowest-ranking (Osservatorio, 1999). For each point of the 1 to 5 scale, the Osservatorio also proposed two weightings reflecting *prestige* of research based on the publishing house or journal in which research output was published: "A" for high prestige and "B" for modest prestige. The method was called 'VPS' (evaluation of scientific production, in Italian "Valutazione della Produzione Scientifica") and was inspired by the UK's 1-5 scale used in the 1992 RAE, which also judged publications' quality according to their national or international degree of excellence. Yet there were differences. Firstly, whereas the RAE rankings evaluated upwards from 1 to 5, the Osservatorio's scale moved in the opposite direction, with grade 5 representing lowest quality. A second qualitative difference was the role of expert groups of evaluators, the panels. Differently to the RAE, the Osservatorio did not place panels' judgments at the foreground of evaluation. The role of external evaluators was defined as *verificatory* rather than *evaluative*. Instead, it was universities who were asked to evaluate their own research outputs along the scale and its respective weightings. Expert groups would come into the picture only at a second stage, to exert a '[…] purely formal control over the awarding made' (Osservatorio, 1999, p.5). The scale itself was *the* quality evaluation such that additional methods  conventionally used by expert panels were not considered[112]. Ranking research according to its diffusion and weighing its prestige were thus the Osservatorio's key methods for a first go at evaluating research quality. With regards to confronting the issue of disciplinary differences, the Osservatorio followed the HEFCE's decision not to assign different weights to the same research output of different disciplines (i.e., weighing a book publication in history more than its equivalent in engineering). This was arguably a wise choice given that there was still a lot of work to be done in aligning the 441 scientific-disciplinary fields find their respective parent equivalent amongst the newly established 17 macro areas. The Osservatorio asked universities to evaluate their research output by using the 17 macro-areas as the disciplinary references, justifying this decision on the grounds that the 441 scientific-disciplinary fields were more reflective of teaching activities than research. This also followed the UK's RAE, where units of assessment incorporated sub-disciplines for the scope of assessment. However, the

---

[112] This is in part explainable by the fact that the Osservatorio (operative officially only from two years) wanted speed up research quality evaluation rather than slow it down through procedures such as the recruitment of a wide array of panels and the setting up of a solid and credible structure for external evaluation – for the latter did not yet exist.

UK had 69 units of assessment for disciplines and sub-disciplines, whereas Italy had 17 macro areas –the equivalent of UoAs – for 441 disciplines.

**The CRUI's research quality indicators**

Parallelly to this proposal, in 1998 the CRUI was drafting its own and radically distinct set of research quality evaluation indicators. Importantly, the CRUI's proposal was published in January 1999 (CRUI, 1999), only 7 months following the establishment of CIVR in June 1998 (see appendix M for the members list of the CRUI's research quality indictors committee). This explains why, upon the Osservatorio's publication in February 1999, the CRUI was able to elaborately attack the core of the VPS framework, namely the decision to keep disciplinary weightings out of research quality evaluation:

> '[…] the [VPS] model proposed [by the Osservatorio] has not found an application in research evaluation […] for it does not offer different scores one the basis of disciplinary fields […] and contains a difficult, if not impossible, assumption, namely that it ranks the type of publication according to a non-objective criteria which is neither detectable nor applicable, that of the degree of prestige' (CRUI, 1999, p. 40).

From the CRUI's perspective, the VPS program presented an objectivity gap because it did not consider how different disciplines tend to have different publishing output types. **This critique was also strategic because the CRUI's quality indicators were based on conveying disciplinary differences.** In the committee's view – similar in this to the Osservatorio's – disciplines' research quality could be judged through the new macro-sector research categories. This meant that smaller or interdisciplinary research would be evaluated through criterion devised at a more general-ised level, similarly to the British units of assessment. However, for the CRUI this did not translate into an equal method for all disciplines. Whilst on one hand it proposed the same research-quality formula across disciplines, it pointed out that it presented different weightings which depended on the dominant research output types in different disciplines. In the Italian university context of the late 1990s, this was a radical and unprecedented suggestion. Further novelty can be attributed to the breadth of the committee's framework, which addressed research production, diffusion, costs and profitability and management, considering the internal and external evaluation dimensions of each aspect. Each of these aspects, the indicators and the broader framework, will be briefly explained in turn to show how the committee approached the question of accounting for disciplinary differences when conveying quality. In the case of research criteria specifically, the committee's weightings were communicated in an obscure manner which did not help to justify assumptions on arbitrary

decisions taken for conveying different disciplines' research quality. Indeed, there were crucial differences in weightings which required more thorough elaboration and explanations to the academic community in case of eventual implementation. Eventually, these weights were not considered in the first national research quality exercise, probably because indeed they were not seen as transparent and suitable enough. Further research, for example through interviews, could further investigate this aspect.

Applying different weights for disciplines was the committee's key innovation. In its calculations, the CRUI tried to statistically convey the relationships between the key factors ("inputs") influencing research output quantity and quality. These inputs were: types of publications, disciplinary contexts, and types & quantities of research staff. Some applied equally across disciplines. Like the Osservatorio, the CRUI assigned major value to publications diffused internationally, assigning major weights to publications in foreign journals and languages. Additionally, full-time staff was weighed (with a weight of 1) more than part-time staff, (with a weight of 0.5) (CRUI, 1999) and the same applied to full-time and part-time researchers (see appendix N for the staff weightings proposed). Types of research output were categorised in a more elaborated way than the Osservatorio's, seeking to fill the "objectivity gap" and overcome perceived distortions caused by the prestige weightings. This was done by looking at publication types *between* disciplines. In addition to those publication types listed by the Osservatorio – books (both authored and edited), articles in academic journals, working papers and conference proceedings – the CRUI included: book chapters, articles specifically published on journals of the Institute for Scientific Information (ISI), multimedia products, patents, geological maps, and critical reviews, amongst others (for the full list see CRUI, 1999). This degree of detail served to apply very specific weightings, transmitting a fairer quality evaluation which acknowledged how different disciplines tend to have different dominant outputs. However, it is important to problematize whether the CRUI's method was really more objective through its weightings. On one hand considering the weightings for different publication outputs was more balanced because it kept into account the key difference –as the SCRI highlighted also in the UK – that social sciences tend to publish more in books and articles as opposed to the hard sciences. On the other, as I argue, the weights eventually chosen conveyed a stark bias in favour of the hard sciences.

Some simple examples explain this argument, such as a comparison of the CRUI's weightings for economics (macro-sector P) and industrial engineering (macro-sector I), respectively (see appendix O for the different macro-sector's weightings). The weightings were in all respects a statistical exercise. Firstly, each publication type in each macro-sector was assigned a weighting (p1,p2,p3…). Secondly, to convey research quality these weightings were multiplied by the quantity of publication types within a given macro-sector. The final set of weightings for each discipline in a

macro-sector was therefore represented by Px (e.g. p1 x 2). Each macro-sector would then have an aggregate P value, the sum of the weighted calculations (P1+P2+P3…), such that the P value expressed the total weighted and normalized number of publications for each macro-sector. This was a crucial value which fed into the CRUI's equation for measuring research quality, expressed as the relationship between the total weighted number of research outputs (P) and the total weighted number of research staff (N[113]), to give the value of the research quality indicator R1, where R1 = P /N. The P value was weighted and then normalized to avoid distortions created by the different range of weights in different macro-sectors. Following this rationale, the weightings would account for different distributions of publication types across disciplines and thus, arguably, convey quality by keeping account of this. Some examples of weights were the following: an Italian article published by a researcher in economics was assigned a weighting of 1 (1.5 if published in an international journal), whereas an article published in industrial engineering was divided into multiple categories: an article published in a high-ranking journal was assigned a weight of 10, on an average journal a weight of 6. Furthermore, for industrial engineering weights were not divided according to national/international categories).With regards to a second type of publication, 'scientifically divulging books', contributions from economics were weighed 2 for Italian publications and 3 for foreign ones, whereas in engineering 'scientifically divulging books' was assigned a weighting of 10 –and a weight of 20 for 'scientific books' (CRUI, 1999). This difference could, to an extent, be supported by the argument that economists are more likely to publish books than engineers whose scientific books must undergo very thorough checking from the academic community before being published. This, however, should have been explicated in the report rather than left unmentioned and therefore leaving academics to deduce differences in weightings for the same type of publication output. This is also because the rationale of output types is less applicable and straight-forward for a case like conference contributions, where the differences in weights were telling. In economics, conference contributions weighed 0.5 for Italian contributions and 1 for international ones. For industrial engineering, conference contributions were weighed a maximum of 5 and a minimum of 1. *The highest possible weight for an economics conference contribution in a foreign language* (with a weight of 1) *was thus equivalent to the lowest-ranked contribution for an Italian conference contribution for industrial engineers.* Clearly, the committee pointed that these weights would be normalized in the final calculations of R1, so to avoid distortions. To return to the example of book publications weightings, whilst differences in weights may account for the fact that economists are broadly more likely to publish scientific books than engineers, at the same time the stark gap between the weightings (10 for a book in industrial engineering and 3 for

---

[113] Logically, as mentioned above, the same weightings applied to staff across the disciplinary macro-sectors in order to reach the final N value. Examples include n1=1 for full time professors (I and II fascia), n3= 1 for full time researchers and n4=0.5 for fixed term researchers (amongst other N weightings)

economics). Furthermore, the weightings still did not clarify why an economics conference deserved less weighting than an industrial engineering one.

I argue that these differences indicate more than the mere acknowledgment of different disciplines' dominant research outputs. On one hand it may in fact provide useful and valid to distinguish, for example, between publications of the earth sciences (such as 'geological maps') through weights, as these types of contributions would not be dominant ones in other disciplines. Within the intention to establish evaluation according to weights of research outputs, assigning weights to geological maps would therefore allow to avoid discriminating the earth sciences in relation to other disciplines. On the other, however, the *specific weights* assigned by the CRUI in 1999 significantly discriminated the social sciences against the hard sciences. The example between industrial engineering's much higher weights than economics, combined with the scarce explanation of such differences for contributions such as conferences which are, assumingly, quite probable in both disciplines, shows how new discriminations were made between disciplines in the criteria's construction phase and precisely through the weightings which were intended to convey a more objective dimensions of quality. These differences are made more problematic by the fact that weightings decided in the report were scarcely explained and accounted for. Indeed, no part of the report explained the weightings decisions, leaving open space for making assumptions or deductions on weightings decisions rather than transparently explicating why weights varied so much in more problematic aspects where output types tend to vary less, as in the example of conferences made above. Elaborating on such weightings would have made the weighting decisions more explicit and crucially would have also exposed the committee to justifying its decisions and itself be accountable for them. The lack of such elaborations points to the validity of critical perspectives on evaluators' powers, 'the dogmatists of transparency' who 'most need shadow' when constructing a new normativity through evaluation (Zarka, 2020, p.42). The report discussed that it mostly used the existing data-bases available for its weightings, such as the SolWebSIR[114]. This database had been constructed by Pavia University's IEU to gather information on university research and was later extended to other universities (CRUI, 1999). The committee added that'[…] the indicators proposed are not exhaustive and can result redundant. Their re-view and simplification will be carried out in light of the data which the CRUI will obtain' (CRUI, 1999, p.27).This suggested that the weights and the respective indicators were only a first round of experimentation with evaluation through disciplinary weights. It is likely, therefore, that the CRUI tried to devise new scales of value through the data already available in order to rapidly set-off the quality evaluation within internal evaluation units as a response to developments occurring at the Osservatorio. What still remains obscure in light of these possibilities is the committee's

---

[114] See https://www2.crui.it/CRUI/repertorio/VI.%20AnagrafeRicercaAteneo.pdf

introduction of such a complex articulation of research quality accompanied by a very limited explanation of the principles and logics behind the evident distortions created by the weightings. In part, the scarce and non-mathematical explanation of the weightings can be attributed to the strictly technical background (mathematical, statistical, engineering) background of the research indicators working group (see appendix M). Still, it does not suffice to answer the question as to why further or later efforts to make the disciplinary weightings method more transparent were not made, and, more importantly, why they favoured the hard sciences compared to social ones in a hypothetical evaluation scenario where quality would have used these criteria.

As seen, the committee's indicators used weightings to signal research quality. If indicator R1 applied weightings by relating staff and numbers of research outputs, indicator R2, related P to Q, a descriptor devised to capture the degree of external research collaborations. Q was a number (below or equal to P) calculated by subtracting a department's research affiliates from co-authored contributions. This descriptor was meant to indicate the external reach of disciplinary departments' work, signalling quality in terms of non-internal research collaborations. Potentially, it offered a more exhaustive quantification of a university's quality declined in terms of diffusion than that provided by the Osservatorio's prestige indicator. The rest of the CRUI's indicators transmitted quantifications of research funding distribution (R3,R4,R5,R6,R7,R8,R9,R10), exploring for example, the relationship between income and high-ranking professors and researchers (R3). An interesting indicator was R11, a quantification of high-level research calculated by dividing high-ranking research staff over the total number of staff . The committee's use of these descriptors shows an attempt to define research quality in a variety of different dimensions that could appeal to all sorts of stakeholders and/or interested parties who could use different indicators when deciding to invest or collaborate. In fact, this set of indicators provided a highly codified research quality assessment *system,* where given that quality could be conveyed in a multiplicity of ways, there was the possibility of relating different indicators to different contexts open (CRUI, 1999).


### Descriptors and Normativity

As discussed, the CRUI's calculations largely re-elaborated existing research data. The relationship between the use of weights and descriptors to convey quality through indicators is worth elaborating in some detail. The weightings proposed by the CRUI (CRUI, 1999) brought to light, once more, the divisions between the social and the hard sciences, and were presented as overcoming the Osservatorio's objectivity gap. Indeed, evaluation of research in different disciplines should keep account of the research outputs differences across disciplines. However, the further gap created between disciplines through weights effectively obscured that very objectivity which the CRUI sought to introduce in research quality evaluation. Applying Zarka's critical insights (Zarka, 2020)

on the dynamics of the process of constructing evaluation, one can unveil that this process also entails the construction of a new normativity. The committee's subjective judgments –obscured behind the CRUI's weightings – reflect validations of a given "truth" at the cost of "other truths" (Zarka,2019). If we agree that objectivity is a specific normativity or normative understanding which evaluation seeks to assert, then a critical analysis of the CRUI's criteria shows how the crystallization of normativity occurs not just in the 'judgment phase' when indicators are *used* by panels during the research evaluation stage, but also, and crucially, at the earlier construction stage. Thus, it is not merely the indicators that convey powers of articulation of value vested as objectivity (Zarka, 2019), but crucially also its contents. As we see, contents were devised **subjectively** by a specific group, the CRUI's working group, and they conveyed a specific understanding of research quality. The committee tried to use weights to convey diversity, yet it favoured the engineering discipline at the cost of others. This use of the weightings exerted the committee's power of articulation and definition.

A second important insight into how power was conveyed in the construction of these indicators concerns the surrounding structure which was proposed. This structure referred to the organization of evaluation responsibilities supporting the evaluation exercise. The structure depended on the CRUI at both the initial and final stages of quality evaluation, thus in the internal and external evaluation stages. In other words, the committee complemented its indicators with its own coordination of them at different levels, placing itself as universities' coordinating body –or conduit– between their internal research objectives and the still undefined external evaluation framework. The CRUI was stepping far beyond what it was officially requested to do, designing the criteria's functioning and diffusion of and thus not 'just' their content. As the report stated,

'The […] indicators defined here by the CRUI aim to stimulate a quantitative definition of parameters which can help to implement departments' self-evaluation procedures […] a process which, *starting from the focalization on the choice of the objectives* […] verifies the research process (which may lead to the re-definition of objectives due to unexpected factors or to major resource needs than those that are allocated). Self-evaluation concludes with a final balance of the objectives that have been reached, with reference to: the means of diffusion of research amongst the scientific community, the effective costs, the appropriateness of resource allocation, the transfer of results in teaching activities as well as towards the external socio-economic environment, the setting up of programs which review human, financial and instrumental resource in relation to the lacunas which emerge during the conduct of research and, finally, *the design of new research objectives in light of the results obtained*. It should be remembered that the objective is to achieve evaluation

exclusively in the context of the specific macro-sector that is of major interest for the facility in question [department]. Once the data on the first round of implementation of the method proposed here will be available, there will be a revision and eventual simplification of such method, as well as the definition of a method capable of comparing between macro-sectors.' (CRUI, 1999, p.4, emphasis added)

This encapsulating force of evaluation was turning evaluation into a universal steering device, that could crucially re-organize research work in departments. The indicators proposed were intended to play a role in re-formulating internal research priorities depending on results obtained, and in this way act to take decisions on research activities. This stage, then, would use the CRUI's indicators for internal research assessment to facilitate a second stage where 'outside' groups – which remained unspecified in the report (CRUI,1999) – would use indicators to make their judgments, deciding to focus either on the cost dimension or the diffusion dimension, for example, depending on their interests in university research. In this way, the CRUI's indicators would be used also externally:

'[…] the method […] entails a first phase of self-evaluation followed by a departmental report, according to the scheme prepared by the CRUI. A second phase entails the visit of a peer committee to the structures, which examines the congruence between the contents of the self-assessment report and the local reality to then express a critical judgment on the report and suggest eventual measures for improving the department's research activity.' (CRUI, 1999, p.4).

This "*organizationalization*" of evaluation (Dahler-Larsen, 2020) facilitated a blurring between the managerial and the evaluative functions of the indicators proposed, importantly problematizing their objectivity in a second way. Surely, it is part of external evaluation processes that research results be confronted to their context. However, external evaluation need not necessarily require institutional information so specific and varied, in other words it needn't become a generic assessment of as many institutional features as possible to understand research quality, for this contextual information can eventually backlash if it is used to suggest for example that resources be used more efficiently rather than being used to understand what research emerges from specific resource endowments. The CRUI's claim that indicators had to be both internally organizing devices for research and used to rank research between institutions when used by external groups was, I argue, a way to facilitate their diffusion as instruments that would rest on university decision-makers' responsibilities, such as, importantly, that of defining objectives. The CRUI's leading role

in the proposed processes of research evaluation as well as in the evaluation of evaluation (i.e. monitoring how universities used evaluation indicators) highlights the committee's attempt to exert significant power over the process of evaluation, which was no longer just a quality evaluation exercise, but also a managerial one, devised *by and for* the committee. The committee did not assert itself as the direct evaluator of research products, but rather indirectly through diverse procedures.

These insights elucidate on the modalities through which *measurability develops into normativity*, of how the practical and discursive construction of evaluation strictly concerns an exercise of power. This discussion (Pinto, 2012; Leydesdorff, 2005, 2011; Zarka, 2020; Dahler-Larsen,2019; 2020) still finds no single theory capable of deciphering the complex interaction between quantification, the cognitive sciences, and the new epistemologies can be mobilized to turn power into an institutionalised technology waiving objectivity yet exerting subjective power(Pinto, 2012,2019). The quantification of quality in the form of indicators surely obscures their process of construction and restricting the ability of non-experts to debunk evaluators' hidden interests. Along these technical features run the organizational implications for the evaluated, asked to assist by providing information on research and lose the direct control over it, for in this case such control is to be ideally handed to the CRUI deciding the functioning of evaluation at a sectorial level and to single VCs for reorganization of activities at the institutional level. The suggestion that quality be analysed also as an organizational phenomenon (Dahler-Larsen, 2019, 2020) adds to the complexity of approaching the issue in merely theoretical terms, and points to the importance of conceptualising managerial governance through evaluation as a constantly changing process. This process is dependent on contestations over the definition of indicators and control over their use. The CRUI's 1999 report, in fact, was an attempt amongst others to make the research quality indicators proposed fit into a future research evaluation framework which was still being developed. As such, though this proposal was later left unattended, it equally left an important bulk of work and principles *around research quality evaluation as well as around research evaluation as a managerial practice*. As such, contextual analyses of attempts to institutionalise indicators can help to understand how groups may propose to organize and coordinate knowledge in order to portray their indicators as more objective than others. Further, these attempts leave lineages which must be studied in order to understand how these contestations play out in a given context. The relationship between the micro-data necessary to construct evaluation indicators, on one hand, and the macro-categories used to compare research produced in extremely diverse contexts on the other may otherwise risk remaining unresolved if evaluation indicators remain treated in the same generic terms advocated by their proponents. The result of a generic analysis is a process of de-politicisation whereby the specific features are not studied nor understood in relation to their proponents, leaving the construction of power through evaluation criteria in the shadows. As has

been suggested from different standpoints (Pinto,2012; Dahler-Larsen 2019,2020), analytical frames would begin from the acknowledgment that the construction of evaluation criteria involves specific decisions that involve power. The power of defining quality surely lies in the ability to create sets of indicators in the first place, but also in the corollary ability to create '[…] logical systems of relationships between them' (Pinto, 2009 p. 145). To put in in other terms,

> '[…] contemporary practices of evaluation are strictly interrelated with management and organizational structures of modern societies […] With organizationalization and institutionalization of evaluation, now evaluative information comes in streams. The ambition is to weave these streams together with streams of funding, streams of legitimacy, and streams of decision-making at many levels' (Kristiansen et al,.2019 as cited in Dahler-Larsen, 2019, p. 51).

### 5.2.3.1. The turn to Citation indexes

The CRUI's work on research quality continued when it decided to use citation indexes in 2002 and devise a list of indicators for research publications. A working group headed by Elena Breno first presented the state of the art of universities' research quality using the ISI's data, obtained thanks to negotiations with the MIUR which had bought the institute's database (Breno et al., 2002). The CRUI's intent through this work was to '[…] measure oneself through indexes, not measure with indexes' (CRUI, 2002, p. 9). In its first research on citations, the committee was once more innovative. Instead of publishing ISI data, it largely re-elaborated it and constructed a specifically Italian use of the citation index that could fit the Italian university research context. The committee therefore mapped Italian researchers' institutional affiliations for over 100.000 generalised bibliographic references on Italian research (CRUI, 2002 p.46), given that the ISI database connected authors only to the city in which they published. From the perspective of the CRUI, devising a citation index *for each university* in the sector could allow to compare research performance nationally and internationally in a new and more exhaustive way (CRUI, 2002). Well aware of the differences between the impact factor (IF) and the Citation Impact indicators; the committee used the latter to analyse universities' research impact. As it noted,

> '[…] the IF refers to a journal, and an average between journals has no sense, especially if in different disciplinary fields […] the IF is calculated annually on the basis of the two previous years' data, whilst the usefulness of a publication in many scientific fields extends beyond the two-year frame' (CRUI, 2002, p.48).

The CRUI's approach altered the traditional structure of the citation index (for reconstructions of the development of this indicator, see De Bellis, 2005;Leydesdorff, 2008), thus it detached citations from the journal in which they appeared. Following its affiliations mapping it took specific methodological decisions. For example, it only analysed publications and citations between 1995 and 1999 in order to have the most updated state of the art and also allow for future longitudinal comparisons. Additionally, it distinguished between universities with medicine faculties and those without, for '[…] biomedical researchers produce a much higher number of publications and thus has a higher average number of citations' (CRUI, 2002, p.30). Finally, it distinguished between polytechnics and "schools of excellence", aware that their structural differences could create analytical distortions when compared to research from traditional public universities. The information eventually gathered is listed below:

– total number of publications (articles, notes, proceedings, reviews) in the technical-scientific disciplines
– total number of citations received by the university (taken from the publications above) in the technical-scientific disciplines
– total number of cited publications

and the following data on staff (taken from the Ministry):

– quantities of research staff (I and II ranks) from the technical-scientific disciplines

This data allowed to devise impact, productivity, and presence indicators for university-level information on research performance(see table 2 below).

*Table 2. The CRUI's university-level citation impact indexes(author's table)*

| Index | Description | Calculation |
|-------|-------------|-------------|
| **Impact Index** | Impact of Italian publications | Citations / Publications |
| **Productivity Index** | Average Publication per research staff | Total number of Publications / Total number of research staff |
| **Presence Index** | Average Citations per research staff | Total number of citations / total number of research staff |

**Citation impact as a measure of quality**

The Committee's use of citations for the purpose of conveying impact was further elaborated at a disciplinary level. This means that the committee calculated the same three indexes (impact, product, and presence) for a range of disciplines. Specifically, it chose to apply the indexes to 9

disciplinary areas If the ISI database was useful for the citation information which it provided, its journal-oriented character meant that the institute's organization of research categories took the journals as the starting point. Using citations to measure research impact and Italian research performance specifically was hazardous, for the ISI's 106 categories (grouped in 24 'fields') did not match different university sectors' disciplinary divisions. As far as Italy was concerned, in 2002 a newly passed ministerial decree (D.M.4/10/2000) had further re-articulated the 1999 17 macro-sectors into 14 disciplinary areas (see appendix P for the list). Therefore, the committee built an equivalence table in which the ISI's 106 journal-based disciplinary divisions were matched to Italy's 14 disciplinary areas. The result was that each of the Italian university sector's 14 disciplinary areas, 'contained', in the CRUI report, multiple ISI categories. This could allow for a comparative quality evaluation of Italian university research based on citation analysis. The CRUI's first go at this was presented in the report, where, however, indexes were compared only for the hard sciences (9 out of 14 disciplinary areas) as these were the ones for which the ISI's database provided more accurate bibliographic information (CRUI, 2002). After this step, the committee perfectioned the three university-level indexes into disciplinary-indexes. Now universities could compare their citation impact, productivity, and presence for the scientific areas 1 to 9. The CRUI advocated that these considerations become part of universities' own understanding of their research and find space in the work of the internal nucleuses so that they could have a better understanding of how to organise research resources (CRUI, 2002). The new indicators devised are summarised below. This work continued in 2005 (Breno, et al., 2005) when citation impact data was updated. The committee hoped to see this used in national evaluation to '[…] propose a *standard methodology* for the evaluation of universities' scientific production, a methodology based on the analysis of citations.' (Breno et al., 2005, p.5, emphasis added). However, differently to the 2002 report the CRUI now conflated university and non-university research under the broader heading of 'scientific and technological research'. Whilst research publications were quantitatively increasing - placing Italy as the 4th in Europe - at a qualitative level the proportion between citations and quantities of publications placed it 10th within the European Union (then composed of 25 member States). Hence the suggestion that, from a citation analysis perspective, '[…] Italy produces a lot but is scarcely cited, that is, its contribution to scientific and technological research is less significant than that of northern European countries' (Breno et al., 2005, p.11). The CRUI was very attentive to northern-European trends which historically disadvantaged Italy's position, with the UK leading the way in both qualitative and quantitative terms. The CRUI saw citation indicators useful for understanding how to correct Italy's research competitiveness. The committee suggested that low competitiveness in citation impact was a direct result of low national R&D investments, particularly when compared to investments in Denmark, the UK, Germany, and Sweden. Whilst

Italian research had undoubtedly drastically internationalized since the late 1990s (CRUI, 2002) because an increasing proportion of it was now part of ISI journals, the situation was still lagging. The committee demonstrated a significant degree of familiarity with citation analysis techniques as well as knowledge on factors influencing research diffusion on ISI journals, yet it made little consideration of how factors other than investments shape citation impact, such as for example language (Dahler-Larsen, 2018), intra-State and inter-State differences between university and non-university research funding channels. In overlooking these factors, it generalised its approach to research quality. The report made unproblematized considerations on ideal quantities and types of research investments. The key proposal was to emulate the USA, to '[…] invest more in the more internationally competitive areas, in order to support them, push them towards excellence and accelerate Italian international competitiveness' (CRUI, 2005, p. 32). For the committee, citation analysis was sufficient for reflecting on conditional research funding. This echoes what much later became the 'differentiation' argument in Italian higher education debates (Capano, Regini, & Turri, 2017), whereby it is desirable to discriminate non-excellent research by: '[…] adopting strictly meritocratic criteria – based on the evaluation of scientific products – for the allocation of available resources, in order to prize the most significant research institutions' (Breno et al, 2005, p.32). This would encourage Italy's more internationalized research to further produce and catch up to virtuous models such as the USA, where investments '[…] monetize every increase in research funding into a productive investment, generating a strong entrepreneurship based on innovation.' (Breno et al., 2005, p.32). These conclusions appear problematic for two reasons.

Firstly, the committee's argument that citation impact become a future criterion for selective funding was exclusively supported by comparative, national-level evidence on R&D investments as a percentage of GDP, strengthening the assumption that high citation impacts in universities depend on R&D funds. Based on the conflation between university and non-university research, this assumption ignored the fact that R&D investments are often directed - particularly those for the 'technical-scientific' fields - at non-university research.

Secondly, the CRUI's proposals easily adapted Italy's university sector to other countries' contexts. To give an example, UK universities' research funding importantly came from research councils as well as from the State, which funded research from its education funds and not from R&D expenditure (Shattock, 2012). In Italy, R&D could reach universities through existing yet still weak alliances between universities and the industrial sector. The weight of funding sources in universities' different disciplinary areas is hard to generalise because it varies significantly. With regards to the scientific- macro areas, usually considered that area which can more easily attract external funding; 'other funders' accounted for 27% of research funding in 2008 (Reale, 2008, p. 69). Exceptions confirming the rule existed, such as the engineering macro-area which received

about 48% of its research funding from non-State funders. The committee's underlying and under-problematized assumptions were thus readily conveyed in terms of a direct causality between R&D investments and research citation impact, seeking to bring attention to its continuing work on research quality rather than elaborating the issue in its complexity.

All in all, this report arguably shows how the committee was losing the rigor with which it previously treated the specificity of university research (CRUI, 2002) and the issues which emerge when trying to convey impact through citation analysis in universities. It began to marginalise a due articulation of the stark differences between countries' higher education sectors. Furthermore, the CRUI had by now side-lined its previous view that evaluating the same research output generically across different disciplines caused risky distortions of research quality. Considered within the changing national evaluation context, this report reflects how the committee endeavoured to keep up through a bottom-up articulation of the CIVR's intentions to integrate bibliometric evaluation by using the ISI database in the 2004-08 VQR (initially proposed for future use by the CIVR, see CIVR, 2006). This mixed with further strains posed on the CRUI's influence coming from the CNSVU, which began to question whether nucleuses double role as internal evaluators and key bodies interfacing with external evaluation was sustainable in the long-term (CNVSU, 2005). As Italy moved to what was then considered the most radical attempt in Europe to assess all of the country's research in one single framework (Potì & Reale, 2005), universities' room for self-regulation came under strain.

## 5.3. Research as a strategic Asset

Whereas throughout the duration of the Prodi government the CRUI find relative space for its proposals within political circles thanks to Berlinguer's role as education minister, with the advent of D'Alema's government at the close of 1998 this certainty risked crumbling as when the new education minister, Zecchino, was elected. In higher education, the new government initiated the principle of steering at a distance, beginning with the announcement of a first, selectivity-funding research evaluation exercise (VTR). The CIVR was the body in charge, and  by the mid-2000s it was considered an '[…]intermediary to put policies into action' (Reale & Potì, 2009, p.90). When Berlusconi's government later replaced D'Alema's in 2001, public sector governance evolved into a policy of 'command and control'. This last section explores how research quality evaluation changed as the State gradually tried to shape it. Two key points emerge. Firstly, that the CIVR ignored most of the CRUI's work on research quality. Secondly, that it prioritised disciplinary factors to institutional ones for research quality assessment.

### 5.3.1. The first national quality evaluation exercise

In 2004, the CIVR published a circular introducing the first and upcoming national research evaluation exercise, initially planned on a triennial frequency. In the VTR, the CIVR eventually opted for its own methods with the aim of acquiring a concrete, and documented understanding of how to improve Italy's research capacity. As has been noted (Reale, 2008), the CIVR benefited from a large pool of academic consensus, both upon its establishment and following its comparison to later national evaluation exercises. This is largely because the CIVR focused on evaluating research outputs rather than the management of endowments and of the research environment more broadly. The types of research output evaluated were: books and book chapters, academic articles on scientific journals, patents, projects, compositions, drawings and design, exhibitions and performances, and artistic pieces (CIVR, 2006). In the circular, the CIVR followed the structure of the British RAE's 0-5 scale (CIVR, 2004), with some modifications. Differently to it, the CIVR's scale indicated specific percentage benchmarks for panel members to work with, and, further, it decided to indicate specific weightings (see appendix Q)

As table 3 below shows, the approach was rather straight-forward. The priority of the VTR was to carry out a disciplinary-level evaluation through the evaluation scale. Attention to qualitative evaluation was supported by the fact that the only method of evaluation used was panel members' recognized expertise in their field. To convey fairness for the fact that evaluation included both university researchers' and non-university researchers' work, the VTR applied final weights of 0.5 to the former and 1 to the latter(CIVR, 2006). As has been noted (Reale, 2008), the guidelines ensured that comparisons would not be made between disciplinary areas avoiding possible competitive and therefore distortionary behaviours among panellists. This guiding principle on the disciplinary focus of evaluation was crucial, it guaranteed also in broader terms that the evaluation framework would not extend beyond the analysis of research and research-influencing factors, thus avoiding the slippery path of conflating of research quality with resource management capacities. In fact, the evaluation of research outputs weighed 6/9 overall with the assessment of other factors around research quality weighing a cumulative 30%. Indeed, the other factors weighed importantly, yet the research context's qualitative features rather than strictly cost elements were used. This gave the CIVR information on research capacity. This was starkly different to what occurred in the UK's first RAE, where it was mostly universities' ability to efficiently use resource income that was treated as an indicator of high research quality. Indeed, research context indicators of the VTR looked at research diffusion and departmental networks broadly. The mobility factor, for example, conveyed research internationalisation with less reductionism than in other cases both in later Italian national assessments and in the UK's first one. Further, the CIVR made no use of the

CRUI's proposals on disciplinary weightings (CRUI, 1999) and citation impact (CRUI, 2002). It limited itself to pointing out that IEUs assign to the submitted publications a disciplinary area by using the CRUI's equivalence table (CIVR, 2006).Besides this, it was evident to the CRUI that the CIVR would not institutionalise its research quality indicators.

*Table 3. Evaluation Criteria for research institutions in the 2001-03 VTR*

| Criteria | Weight |
|---|---|
| A – Quality (of weighted research products) | 4 |
| B – Property of excellent products[115] | 2 |
| C – Researchers with international mobility | 1 |
| D – Capacity to attract external financial income for research | 1 |
| E – Researchers in training | 0.5 |
| F – Capacity to invest own financial income for research | 0.5 |
| **Total** | 9 |

If research output evaluation had an important weight in the final research quality evaluation, it was not the only indicator. Indeed, the CIVR's criteria also included an assessment of the research context. In a transparent way, the CIVR focused on human and financial resources, the former analysed through mobility and training and the latter through attractiveness to external funding and ability to re-invest public funds for research purposes. The evaluation of these factors was carried out for each disciplinary area, and the institutional figures were displayed as their sum, without recurring to additional formulas or weightings besides those regarding different types of researchers (full-time/part-time; university/non-university). These elements were judged through the IEUs reports (CIVR, 2007). The CIVR found a way of integrating internal evaluation without overlapping internal and external indicators. Indeed, IEUs were, along with the CIVR and the panels, key units in the evaluation process, their reports formed the basis of the context evaluations. IEU members working for the VTR were chosen by universities and authored a three-year report on the institution's research aspects. This information made up the 4/9 of the overall evaluation, and through it the CIVR pushed universities to articulate how their efforts and endowments translated into research capacity. Information requested was on research plans, proof of training of

---

[115] This criterion was an index: a calculation of the percentage of excellent judgments by the total products judged. In the 2001-03 VTR this calculation was done according to disciplinary areas of each university and research structure.

researchers, capacity to attract new resources, and, importantly, the way that the university used nucleuses' internal evaluation for taking decisions (CIVR, 2006). These are examples of how the CIVR tried to understand the functioning of the management of research resources without imprinting a managerial approach to research quality evaluation.

### 5.3.2. A new evaluation agency

In 2010 this radically changed (Rebora et al.,2015; Turri, 2014; Pinto,2009, 2012, Banfi&Viesti, 2015) when law 240/2010 (known as the "Gelmini law") reformed both internal university governance and the national evaluation structure (Rebora, 2012; Turri, 2014). This law almost entirely re-shaped the university sector's structure from scratch with institutional management and research evaluation as its primary organizing principle. A new agency, ANVUR, was established to lead the scrutinization of universities' internal management procedures (including recruitment) *as well as* research quality, these powers were gradually built on in following years and added significant administrative and financial powers to the ANVUR's monitoring role (Paolini & Soverchia, 2014; 2015, 2017). Additionally, a new structure of skills tests known as the National Scientific Habilitation (ASN, in italian "Abilitazione Scientifica Nazionale") was introduced, becoming compulsory for individuals aiming to enter the university as professors of highest academic teaching ranks (I and II ranks). This evaluative imprint on the academic profession became the new and essential scientific certification for employment advancement, without which also staff employed for years within the sector can find itself officially unqualified to scale up[116]. As Rebora has put it, '[…] by 2012 all of the sector's teaching staff and aspiring candidates were mobilized for evaluation' (2012, p.9) for habilitation and for research purposes. The arrival to this point is filled with a series of interlocking and complex dynamics, which have been exhaustively debated in their multiple dimensions (Barone et al., 2010; Capano, 2014; Minelli, Rebora, & Turri, 2015; Turri, 2014; Viesti,2016, 2018;Pinto,2012;). What is of most interest here are the political interests of State and university groups behind the emergence of ANVUR are the political interests behind it. From a situation where the CIVR was unfavourable to institutional-level evaluation and prioritised disciplinary-area-level evaluation (Reale, 2008) in fact, ANVUR has gone in the opposite direction, making conditional funding intrinsically tied to an analysis of the institution's processes (Banfi & Viesti, 2016) where '[…] each university is favoured by the scarce performance of others' (p. 345). Specifically, evaluation of research quality was, as represented in the ANVUR's

---

[116] Interestingly, this has not facilitated Italy's notorious issue of an ageing academic staff population, already evident in the mid-2000s (Reale, 2008). Quite the opposite, it has facilitated the permanence of older staff at university, with over half of the proportion of full professors aged above 60, and 47% of associate professors aged between 47 and 57 years old. In terms of access to the academic profession, also fixed-term researchers are ageing, showing the difficulty of advancement within the university: as of 2017, only 0,6% were aged below 30 compared to an almost 2% in 2013. (ANVUR,2018).

first report, defined also as the ability of universities to prove their management and accountancy strengths through reporting, which gradually turned into an essential element of research quality, even though researchers have no direct influence over resources as in other European university sectors such as Germany's, for example. The bureaucratization process set off by ANVUR's evaluation framework (Rebora, 2012) has been defined highly problematic, creating a much more complex evaluation process which lacks in transparency. It is difficult to define ANVUR's framework in generic terms, for it is still debated as either managerial-technocratic or as a State bureaucracy (Turri, 2014). The agency's establishment is thus often seen as a de-regulation of the sector along the NPM model primarily because it occurred in a context of major university autonomies to be monitored (but arguably not controlled) through performance criteria. However, ANVUR's role in the first VQR showed that in practice the rhetoric of efficiency importantly functioned to legitimize the scrutinizing functions of a body concerned with steering universities rather than protecting their autonomy. As I argue, the debate over how to best define ANVUR (as either managerial-technocratic body or a State-bureaucracy) has valid arguments on both sides and gets to the core of the dilemma of higher education governance through evaluation/ gets to the core of evaluation as a governmental instrument in higher education. However, it is extremely difficult to disentangle this debate in generic terms.

On one hand, ANVUR appears a State-arm which evaluates to centralize the governance of the sector into its powers and steer it according to political priorities (Pinto,2012). This role appears to have been the politically-desired one if one looks at the law establishing the agency (D.P.R. N.76 of 2010). On the other hand, the agency's managerial and technical criteria are complex and articulated to the point that they have been used to construct a logic which is very distinct from the political one. The agency uses such a high degree of complexity and breadth in the collection, elaboration and manipulation of data that it is increasingly difficult to fit its powers within the traditional political 'skills'. One can try to untie this knot by taking on a historicist approach. Specifically, it can be useful to break down ANVUR's actions and evaluation assessments since 2010 into multiple phases and study each VQR in comparative terms to understand changes to the criteria as well as changing narrations and principles mobilized each time that evaluation was used for new purposes (Viesti, 2016). Here, I take on a first attempt in doing this by looking at the emergence of the new agency. Coherently with the theoretical agency-historicist framework which I have so far used to conceptualise the rise of a managerial class in HE, I propose that there be more scholarly attempts to analyse ANVUR in contextually-specific terms rather than generic ones. This is even more compelling given the stark differences of each successive VQR, on one hand, and the political governments steering the agency, on the other. As VQRs often changed, so did the criteria being used, and, more often than not, this created new demands on universities to integrate the new

criteria into their organizational structures and their internal hierarchies. The changing political priorities, secondly, have often translated into an erratic funding of the sector when seen in longitudinal terms, with the FFO and its selectivity (now called "premial") part creating uncertainties at each round of the VQR. In other words, if literatures agree that both the criteria and the political interests of the agency have continuously changed since its establishment, there is little reason to continue analysing the agency's role and character in generic terms, for this means distorting and further obscuring its development by taking on the assumption that it is progressively developing (just as the debate on evaluation instruments broadly claims) either towards neoliberalisation or technocratization-managerialisation of the sector. I propose instead to give to contextually define its initial features and roles, those that the agency exerted upon its establishment and in its first VQR. ANVUR's establishment was defined by a combined articulation of the agency: a law proposal of 2006 (Act n.3773) first articulated it, and a definitive law (D.P.R. n.76) 2010 eventually established it. Looking comparatively at these two documents, each with their group of proponents, allows to grasp the set of interests behind the rise of the agency and make a first set of initial comments on how the contestations over quality evaluation played out when defining the agency's future powers.

### A political void filled with political debate

Between the years 2006 and 2011 (after the first VTR was carried out and until the first VQR by ANVUR began in 2012) the Italian university sector had no national evaluation agency and did not witness any national evaluation exercise. Indeed, the second VTR for the years 2004-2008 which the CIVR intended to carry out – and for which guidelines were indeed made – did not take place[117]. However, this does not imply that this period was politically void, for there was significant 'fermentation' of proposals for the body that would eventually replace the CIVR, and discussions over this agency were more heated than is often assumed when narrating ANVUR's rise as a natural succession of CIVR's. Two key and contrasting proposals lie behind the future agency: one from the Ministry and one from a group of CRUI members which had by 2006 moved from HE walls to political corridors. Prodi's return as prime minister in 2006 was in fact a new opportunity for the CRUI, which entered the political arena. This is the case of Luciano Modica, who substituted Luigi Berlinguer as senator of the "Ulivo" party (left democrats) on 31 October 2002 following Berlinguer's resignation. Shortly later, from 7 November 2002, Modica participated in the Senate's

---

[117] This can be ascribed to a set of reasons. In 2006 a law (n.286) actually abolished the CIVR and the CNVSU and provided that they be substituted by a new evaluation agency (ANVUR). However, this agency did not become politically defined at all until 2010 (D.P.R. n. 76 of 2010), and this fact could explain why the ministry clumsily continued to send over guidelines for the VQR 2004-2008 as of 19 March 2010 (prot. n.8/2010).

political enquiry survey ("indagine conoscitiva"), launched by the 7th permanent Commission[118] ("7a commissione permanente") on November 6th, 2002. This inquiry investigated the degree of practical implementation of a law which had been passed in 1998 (number 204) and which sought to regulate national planning of national research along European guidelines and trends. Having participated in the entirety of the inquiry (November 2002 to February 2006), Modica's presence ended with stark disapproval over the Commission's final document (Modica, 2006). He expressed delusion for the fact that the Ministry Letizia Moratti was wrongly associating the British RAE with the CIVR's VTR, arguing that her conclusions on the enquiry were based on a privatised view of research funding and development. He added that she had '[…]provided false evidence' (Modica, 2006) which portrayed a much more positive picture of the national research sector than was the case in order to maintain political support. Further, he vocally expressed that the commission had lost an important occasion for '[…] provid[ing] the country with instruments for understanding and confronting the malaises of the research sector' (Modica, 2006[119]). This sentiment was in accordance with that of three other critical members, Albertina Soliani, Alberto Monticone and Fulvio Tessitore[120], who also voted against the final publication. Given their minority within the commission, however, the document was still passed, yet the group which was formed was bound to last. In February 2006, Modica led the minority[121] – now enriched by the presence of Giampaolo D'Andrea and the Commission's ex-president and vice-president, Acciarini and Mauro Betta – in Senate, where they proposed a law for the establishment of a new national evaluation agency, what would become ANVUR. This law proposal is worth considering in some detail for this and two other reasons. Firstly, because it was drafted by the inputs of two ex-members of the CRUI, Modica and Tessitore, who had an array of information and previous research on university evaluation at their disposal. Secondly, because even though the proposal was not readily approved, the abolishment of both CIVR and CNVSU in November 2006 (law n.286) following this proposal can be seen as reflecting an immediate political step towards the direction then being proposed by Modica's group. Indeed, the abolishment of these bodies left the grounds free for the establishment of a new agency, the sector could not be left without one and both the academic and the political world were aware of this. These aspects allow to consider how the CRUI committee's proposals could hope to find future political legitimation. The eventual law (D.P.R. n. 76 of 2010) establishing ANVUR has been interpreted as the successful culmination of Modica's proposal because it articulated the organisation and functioning of ANVUR along similar lines. Whilst

---

[118] https://www.openpolis.it/parole/cosa-le-commissioni-parlamentari-perche-importanti/
[119]See http://www.senato.it/japp/bgt/showdoc/frame.jsp?tipodoc=SommComm&leg=14&id=176410&part=doc_dc-sedetit_pi for the minutes of the meeting
[120] 'shocked that the minister finds only advantages and no critical aspects of our national research sector' (Tessitore, minutes of meeting of the commission held on 08/02/2006)
[121] Modica was in fact 'first signatory' of the proposal.

agreeing on the fact that the proposal's instituionalisation into a law could have allowed the CRUI to importantly see its definition of evaluation and university governance legitimated, I suggest that considering the law as an alignment to Modica's proposal is misleading. There are significant differences in how the proposal and the law defined the powers of the future evaluation agency. Indeed, the law proposal proposed more power for universities in evaluation than was eventually granted to them. At a closer look which accounts for the broader historical trajectory of the CRUI's interests and work in university evaluation, *the D.P.R. appears was actually a partial failure in the legitimation of the 2006 proposal* (Act n.3773[122]). This emerges evidently through a detailed comparison between the proposal and the law. Whilst many differences emerge, the central one concerns the two documents' prescriptions for the agency's role in internal university governance and evaluation. *The law proposal aimed to safeguard university autonomy in internal evaluation much more than what the law eventually guaranteed.*

From the prescriptions of the D.P.R. n.76, it emerges how the legislator aimed to centralize the IEUs' activities through the new agency rather than leave them to autonomously function within universities. For the legislator, ANVUR was entitled to carry out an external evaluation of internal processes(D.P.R. n.76). Not just research therefore, but also teaching and management, were processes to be evaluated externally. Specifically, the agency could evaluate each activity's process, results, and products through international standards to evaluate teaching, peer review methods and income from external funders to evaluate research –quality and the universities' competitiveness (D.P.R. n.76). The agency was eventually charged to scrutinize internal decision-making processes and their outcomes and base efficiency judgments by evaluating the work of the internal evaluation units, a step which starkly opposed CIVR's non-interference with IEUs' work. After the CRUI's stimulus (increasingly strong since 11995) to develop as universities' decision-makers' key information providers, by 2010 IEUs were well-established in the majority of universities. D.P.R. n. 76 now aimed to obtain close to full control over their operation by suggesting that ANVUR could re-define IEUs internal evaluation criteria (such as those for teaching, student services, and courses' financial sustainability). Opposed to this, the law proposal reserved an article (Act n.3773, art.13) to internal evaluation precisely as a means to exclude this activity from the agency's evaluative functions. The only aspect of internal evaluation which the CRUI members' proposal explicitly assigned to the future agency was research quality evaluation. The agency would thus ideally carry out evaluation of research products through the national evaluation exercise, and would limit itself to propose, and not assess, internal research quality assurance criteria. For Modica's group, research would ideally be both internally and externally evaluated. The rest of universities' activities, however, would be developed in congruence with universities' own assurance of management and

---

[122] http://www.senato.it/leg/14/BGT/Schede/Ddliter/testi/23790_testi.htm

quality processes on teaching, administrative management, and services (Act n.3773). Internal evaluation units were that key structure guaranteeing universities' autonomy from external interference. The proposal did not merely suggest that IEUs be responsible for evaluating the abovementioned activities, it also sought to safeguard them by explicitly defining the terms of their functioning and organization in relation to the agency's new power. As I argue, this demonstrates the CRUI's legacy and its intent to maintain university governance in the strict control of governing bodies and VCs. An example is the suggestion that '[…] units formulate reports and proposals to the governing bodies of the university' (Act n.3773, art. 13, comma 8), echoing the CRUI's very first proposals on units as central to the well-functioning of university governance (Allulli,1995). This perspective on future university governance was based on the VC's leadership, which would ideally be supported by a stronger administrative council. Considering that Modica was the key advocate of the law, this makes much more sense. As we have seen, the CRUI was aimed to keep roles of IEUs as strong as possible, because these were not only closely tied to a managerial governance role that VCs could freely act in, but also, because these were the key bodies which allowed universities to defend themselves from central interference on aspects of resource use and management. Modica's group's proposal sought to guarantee this autonomous function precisely to avoid that governing bodies' actions be scrutinized, as their resource use and decision-making powers. Instead, legislators' D.P.R. N.76 eventually gave ANVUR the right to control internal procedures through the assessment of virtually the majority of internal governance powers, namely over '[…] procedures, results and products'. The proposal had strongly limited the agency's functions in this respect to those of '[…] evaluating that they achieve the intended aims set out by the governing bodies' (Act n.3773).

Contrarily to an easy success of its proposals, therefore, the outcome of the law proposal as embodied later in ANVUR's establishing law marked a new obstacle for Modica's group, evidently attacking the CRUI's broader intentions to solidify internal governance through the roles of the VC, the IEUs and the board of directors –all working in tight collaboration in governing committees to manage resources through internal performance indicators. If in part the group's proposals (Act n.3773) were later integrated in the Gelmini law 2010 because it established an independent external evaluation agency assessing universities as the group had proposed, it is de-politicising act to suggest that the emergence of ANVUR *can be directly traced to the group's ideal view of a national evaluation agency*. This has crucial implications for how we can fully comprehend the development of managerial governance through evaluation in universities and try to understand its internal and external dimensions in detail. This is even more important if seeking to better comprehend VCs' relative successes and failures vis-à-vis political powers and their changing surrounding context more broadly. The way that ANVUR was eventually articulated by the

legislator (D.P.R. n.76) actually dismantled the core of the group's proposal (Act n.3773) by granting the State-devised agency the ability to interfere directly in evaluation criteria which in the group's proposal, were powers reserved to IEUs and at the discretion of internal university politics in the name of its autonomies. It is s thus more accurate to suggest that the group's proposal was to an extent willing to bargain external research evaluation academics' work in exchange for maintain control over those internal governance powers and processes which the CRUI had been constructing through internal performance indicators and evaluation. Similarly, the legislator's articulation of ANVUR in 2010 (D.P.R. n.76) reflects, by large, *an unintended effect* of Modica's 2006 proposal, with rather adverse effects when analysed from VCs' perspective.

It was rather the Gelmini law (n.240 of 2010) which eventually comforted VCs' governance aspirations for it lay the grounds for the verticalization of internal university governance, making an '[…] explicit attempt in changing the structures and the distribution of powers […] prescribing a greater role to central governance organs – the rector and the board of directors – at the cost of the representative organ of the academic community, the Senate.' (UNIRES, 2014, p. 8). This new opportunity was quickly grasped by many VCs, who enlarged their teams by hiring more collaborators. Only 8% of universities decided, for example, *not* to create pro-VC figures in their internal governance bodies (UNIRES,2014 p.65), and instead over 30% of members in universities' board of directors were hired from private enterprises. (UNIRES, 2014, p.67). The CRUI had in fact articulated its view of the role of the board of administrative directors as central in governance, suggesting that

'[…] the rector should therefore have the possibility to enhance the governance of the university through the collaboration of an organism, the board of directors, with whom he can express the real capacity of university governance. In the context of autonomy, statutes define the competences and the composition of the board of directors, which should, however, be a slim and operative body which goes beyond the mere representative logics of the university' (CRUI,2004, p.5).

To assert the managerial features of governance, the CRUI emphasized that it would function according to check and balances criteria whereby '[…] the administrative director should operate closely with the rector in a relation based on trust'. The rector thus was seen as the figure for nominating the administrative director, finding a figure who would be '[…] actively involved in a constant organizational innovation of the administrative structure of the University, taking care of putting into practice through efficiency and efficacy the objectives set by the governing bodies of the University' (CRUI, 2004, p.6).

## 5.4. Conclusion

The CRUI's first turn towards evaluation was the result of Berlinguer's concerns on Italy's lagging position and missed chances to engage in European-wide projects on higher education. In 1991, the CRUI devised its own means for diffusing an evaluation culture in the sector and worked intensely to structure its functioning and methods, beginning with internal evaluation which stressed VCs' roles in objective-setting and in designing evaluation criteria. From the perspective of the State, IEUs were instead meant to become that pivotal body through which politics could have a window of control over universities' resource use without intervening on quality criteria. For the CRUI, the broader system of evaluation was not merely an exercise of accountability. The CIPP model gave the committee a way to frame evaluation as a new organizing principle which would orient internal resource distribution and empower decision-making, with the CRUI's "circular evaluation model" allowing VCs to reap the fruits of the newly conceded autonomies. This 'first-stepper' behaviour can be in part explained by the CRUI's determination to become increasingly influential in shaping university policy, as the advocacy coalition had tried to do at Pontignano (Vaira,2011).Throughout the 1990s, the CRUI was well ahead of the State in evaluation criteria and principles, as we have seen anticipating the work of the Osservatorio which struggled to be set up. The committee was active in trying to institutionalise its indicators, with relative success in the early 1990s on internal evaluation and a relative unsuccess on research quality indicators in the late 1990s. As has been shown, if in a first moment in 1999 the committee paid unprecedented attention to weighting the disciplines' different output types, in 2002 and 2005 this concern was suddenly swept away by the work on citations and the construction of productivity indexes based only on research published in journal articles. These moments of sudden reversal indicate that institutionalisation of the committee's work was often prioritised to the concrete content of evaluation. As has been argued throughout, this reflects the deep contestation dynamics which make evaluation an instrument central to university development because of the power which its control can bring about rather than because of the concrete benefits which the groups involved see in evaluation itself as a means to develop the sector's quality. Quality has often been distorted by power struggles over the control and the definition of evaluation, a dynamic which cannot be marginalised to a case of a top-down progression of the power of a managerial State. By showing these numerous power struggles I have rebalanced the perspective that the Italian State was a fully-fledged evaluative state when it first set up university autonomy and evaluation procedures, an argument that many recent studies continue to make (Pinto, 2012, 2019;Turri, 2014;) . It is instead more appropriate to acknowledge that in their repeated attempts VCs have played a political role of seeking to define and control evaluation and thus the relationship between this part of the university sector and political demands ought to be considered in a constant negotiating dynamic whereby the bottom-up and inside-out developments

reveal much more than a top-down linear understanding of university governance (Moscati, 2010 in Barone et al., 2010).

# Chapter VI

## Comparing the two trajectories

The chapters so far have looked at the formation of managerial groups in different contexts. They have traced the bottom-up development of a tight nexus between educational evaluation and managerial governance. As highlighted throughout, the key feature sparking it historically was the innovative approach of systems analysis, which framed performance assessment as a practice carried out by top-level decision-makers. In this respect, the CIPP and PPBS models shared the approach, portraying evaluation as an instrument at the disposal of decision-makers, with advocators asserting that information gathering and elaboration be used for re-organizing an institution's activities.

However, tracing the weaving of this relationship does not directly imply that managerial classes form each time that systems analysis-inspired models are used. This perspective distorts the complex process of building a new managerial expertise by alluding that any group using managerial practices of this kind will inevitably succeed in governing through evaluation. This further devoids managerialism of the meaning espoused here, namely governing an organization by using performance assessment as an instrument for acquiring the powers to determine the space of action of an organization's lower levels. By exploring how different groups at different levels and in different contexts have attempted to do so, I have undertaken the complex task of seeking to untangle how educational evaluation specifically may have developed into an instrument of managerial governance, and, secondly, which group of agents succeeded in creating a framework in which this could play out. I have thus aimed to unveil how state agents made use of performance assessment practices to govern the PA and assessed their relative successes and failures. I have then highlighted the failures of political groups in governing universities through performance assessment by showing that VC committees were relatively more experienced in this practice than political groups when broader NPM concepts first emerged in each political-economic context. In focusing on these moments of emergence, the claims made so far refer strictly to the *rise* of managerialism in HE and do not extend to what is a complex open-ended development. The emergence of managerial governance in universities was marked by the VC committees' success at articulating evaluation methods and principles in more specific and tailored ways than political groups. The study has conceptualised this as a social innovation. VC committees actively moulded their advantage over policymakers in educational evaluation precisely by emphasizing policymakers' gaps in defining quality and quality evaluation. It was by stressing these gaps and filling them with (often more complex) alternative definitions that the two VC committees managed

to exert significant leverage over university governance during the initial development of the NPM paradigm.

The argument that VC committees constitute a managerial class thanks to their active engagement in constructing evaluation does not stem merely from evidence that they used practices of performance assessment, nor because it is assumed retrospectively that this form of evaluation has become dominant in the two sectors. As I have highlighted throughout, the definition and use of educational evaluation practices was frequently and intensely contested. Rather, VCs can be conceptualised as an emerging managerial class because they consistently attempted to create very elaborated evaluation practices, roles and processes which could support committees' control over universities. This often involved defining quality and evaluation in context-specific ways that could support top-down governance powers for VCs. As shown, VCs often acted as first-steppers in the attempt to quickly diffuse and institutionalise their proposals. These attempts often succeeded as universities followed committees' guidelines on evaluation, particularly those on internal evaluation.

I here summarise this process by comparing the two countries' trajectories of managerial university governance and drawing similarities and differences. The implications of the differences in the two trajectories are both theoretical-methodological and historical. On a theoretical level, the key implication is that context-specific historical analysis on different groups' use of evaluation practices is crucial for conceptualising the formation of managerial governance in university sectors and how evaluation may be instrumentalised to develop into a managerial technology empowering decision-makers. On a historical level, the central lesson so far is that the historical international lineage of educational evaluation practices played an important role, so it should be historicized also for other university sectors to explore how it diffused and how it was used. In this study, the fact that this international lineage developed previously to the advent of NPM reforms helps to account for VCs' vanguard proposals on evaluation when efficiency policies were introduced. Overall, these aspects point to the importance of framing VC committees as political groups rather than as mere subjects of State governance.

## 6.1. Methods

The CRUI and the CVCP had two different approaches to educational evaluation shaping their set of indicators. The former built on the principles of the CIPP model to draw out performance indicators focused primarily on productivity intended as institutions' ability to achieve final outputs (i.e., graduates for analysing teaching and research articles for analysing research). The latter built on the principles of the PPBS model to devise cost-focused indicators assessing institutions' ability to efficiently use resources. Although both committees turned to quantifying universities' outputs at a micro-level by calculating indicators 'per staff' or 'per student', they did so differently. This

explains why the CRUI's first performance indicators had a more qualitative dimension than the CVCP's. The latter first devised indicators such as 'cost per FTE student' or 'equipment costs per FTE staff', whereas the CRUI did not carry out such costing analysis, quantifying instead the activities of each staff, such as 'number of articles per staff member'.

In terms of research quality criteria first devised, the key difference in different quantification efforts concerns the unit of analysis that VCs gathered information from. The CVCP constructed indicators by gathering information at a cost-centre level, whereas the CRUI gathered it mostly at a university level. This has qualitative implications because the CVCP's cost-centres redefined departments as cost-incurring units, whereas the CRUI maintained a more traditional view of departments –even though it devised indicators about 8 years after the CVCP and so could have borrowed from existing cost-frameworks had it wanted to. The CVCP's key indicator showing this difference is 'research income per FTE academic staff' (CVCP/UGC, 1986), with each cost-centre reporting this information for each staff member. The CRUI, instead, constructed this indicator at a university level, and did not search for staff-specific quantification. The CRUI's micro-level indicators (i.e., indicators calculated per staff member) defined outputs as "products" rather than as wealth, with indicators such as 'number of articles per staff member'. The CVCP had no indicators on publication outputs per staff member due precisely because of its cost-focused approach geared at assessing resource management. To be fair, the CVCP did include this indicator amongst its set of 'analysis of publications' indicators (CVCP/UGC, 1986), yet did not include it in its first and second PIs list, suggesting that it was considered secondary to other PIs. Even in the first RAE, in fact, publications were not asked per staff member, but rather on a cost-center basis. The RAE eventually used a criterion whereby a cost-center had to decide the five publications to be assessed, rather than using an indicator such as publication per staff member which could also have conveyed research production in different cost-centres. For the CRUI there was an evident focus on attainments in relation to faculties' overall staff endowments and, in the case of teaching especially, student numbers. It is the differences between the CIPP and the cost-benefit model which account for the committees' different articulations of what performance ought to convey in the performance indicators devised. These differences played out also in the indicators proposed for universities' institutional management.

Taking the CRUI's resource and product indicators to understand its attention to resource use, it emerges that the cost-benefit structure of PIs was not as elaborated on costs as the CVCP's. The CRUI's resource indicators were calculated on a university level (CRUI, 1996), CIPP did not extend to quantifying departments' efficient resource use. Its departmental-level indicators were aimed at qualitative assessments such as average scores and average number of exams, using mostly staff/student ratios to convey resource distribution. Following the construction of university-level

indicators, the committee then produced national averages to enhance comparisons between universities. The CVCP's detailed attention to cost-efficiency is to an extent explained by the fact that universities already had an array of qualitative indicators such as completion rates entry scores, etc. The two committee's degree of attention to costs is thus partially explained by Italian universities' backwardness in the attainment of university data albeit the significant time which passed between the construction of CRUI's indicators and the CVCP's. After having devised more 'basic' indicators, the CRUI did in fact turn to a cost-focus, as the 'ideal student' indicator shows. The CVCP's managerial imprint, however, is more marked that the CRUI's also when keeping account of these differences. This is because the CVCP carried out a meticulous scrutinization of costs. It emphasized the reduction of administrative costs as well as non-administrative ones, such as library costs ('book expenditure per FTE student'), 'telephone expenditure as a % of general expenditure', and 'computer service expenditure per FTE student' etc. (CVCP/UGC, 1987a). In line with the attempt to raise the scope for managerial scrutiny, these indicators would be used by university planners to monitor resource use and thus the CVCP's definition of university performance. Further, the committee fuelled the view that educational evaluation can and should serve as an instrument of managerial governance also outside internal university affairs. Indeed, it attempted to extend cost-focused indicators to research quality. assessment when, in 1991, it advanced indicators such as 'publication per £1000 of external income' (CVCP, 1991). This step affirmed the primacy of cost analysis to convey research quality –and simultaneously re-affirm the primacy of internal university planners who would ensure the maximisation of publication outputs – ,for it was *also through publications* that evaluation could become an instrument for empowering managerial roles.

If on one hand the indicators show that the CVCP's evaluation was more managerially-oriented than the CRUI's, on the other it would be very hazardous to limit the comparison of the two committees' managerial character strictly to indicators. Doing so would imply that managerial governance can be said to exist whenever groups using managerial practices are identified. This would obscure the indirect ways in which managerial groups often expand (Nicoli, 2010, 2014), particularly strong when evaluation is used (Power, 1997; Borrelli & Giannone, 2020). These indirect ways include discursive devices which can support specific indicators and their functioning and build on the nexus between evaluation and managerial steering. Arguably, this is an intrinsic role of evaluators in the first place: by evaluating they inevitably assert judgment upon others or things, a judgment oriented at changing existing arrangements for improvement. However, the point of this research is precisely to problematize why and how VCs took on this evaluator function and the ways in which this fuelled a managerial definition of evaluation in university governance. Otherwise, the issue would be to return to square one: conceptualising evaluation and

managerialism as inextricably tied and thus managerial governance through evaluation as an inevitable progression of a broader paradigmatic NPM turn. As shown, the committees employed innovative articulations, constructing – and later building on – them to sustain their proposals. There are visible similarities in how the two committees narrated the instrumental use of evaluation for managerial purposes. Both committees used evaluation to suggest that VCs become and act as the primary formulators of goals. The CRUI articulated such goal setting in terms of attainment targets whilst the CVCP in terms of expenditure control, yet these differences do not underplay an instrumental framing of evaluation supportive of new managerial roles. By re-defining goal setting through evaluation, both committees undermined academics' role in key academic functions such as deciding on educational requirements (the need of extra resources, extra staff or educational contents/attainment levels). As has been discussed, the construction of this nexus played out especially in internal evaluation, where VCs could *de facto* imagine to act with less constraints compared to the sector-level influence that they could try to exert through evaluation. The CVCP, as discussed, made an outright attack at academic senates to raise the importance of managerial and executive powers. In Italy, the CRUI importantly stressed that IEUs be used for governance and follow governance goals even though policymakers intended IEUs mostly as neutral-technical bodies assessing resource use. In its proposal of the 1999 research quality indicators, further, the CRUI proposed to act at both the beginning and the end of universities' evaluation procedures, ensuring in this way that its role be both that assign research weights for university research and of verifying how well universities self-assessed their research by assigning weights. This marks how the evaluation-managerial nexus was shaped not merely by indicators but importantly also by the narrative and organizational devices mobilized to support their functioning (Pinto, 2012; Dahler-Larsen, 2020). The fact that VC committees acted to construct such devices is a significant lesson learnt from this research, which tells that much more attention must be paid to analysing the surrounding context within which VCs insert their articulations of evaluation to understand how it is used to managerially govern universities. This last point is strictly relevant also for a second way in which VCs' construction of a managerial role through evaluation can be analysed and compared, networks.

## 6.2. Networks

Networks have played a key role in shaping committees' path to evaluation and managerial governance. They have allowed to develop methods and political influence.

VC committees' methods largely borrowed from the work of key conduit agents, Geoffrey Lockwood and Giorgio Allulli, who actively participated in international research on educational evaluation between the 1970s and the 1980s. Interestingly, the fact that at the time these agents were not part of the VC committees suggests that their later membership was importantly shaped by

the experience and notoriousness acquired through the international projects. So far, this aspect points to the importance of historicising the international lineage of educational evaluation research and practices in order to understand how its diffusion may account for national-level developments in evaluation. At the same time, it is thanks to an agency-centred historicism that the committees' use of two different streams of international research has been traced. As discussed extensively in previous chapters, the committees' conduit agents carried with them not just their research experience, but also the deeper historical roots which made such research possible in the first place. When Giorgio Allulli presented the CIPP model to the CRUI, he carried CENSIS' research legacy on school evaluation, and, tracing further backwards, the centre's own inheritance of post-war OECD projects on manpower planning in Italy. Similarly, the Jarratt committee availed itself of key propositions on managerial governance made by Geoffrey Lockwood (Lockwood & Fielden, 1973; Lockwood,1972b,1981), who transmitted the insights of IIEP's and CERI's research on institutional management in universities. The latter was in turn a PPBS-inspired rearticulation of human capital theory, made possible by the mobilization of RAND networks within the OECD and UNESCO (Hirsch, 1965;1968; Coombs, 1968). If on one hand the two committees borrowed from different research streams, on the other these streams both conduce back to the common historical lineage of the systems analysis approach in education which was first pioneered by US military intellectuals in the late 1950s (NDEA, 1958; Kershaw & McKean, 1959). Comparing the two committees' indirect international networks thus so far supports the argument that the rise and initial development of output-governance in education is strictly connected to those alliances between military intellectuals and researchers which emerged at the OECD during the Cold War context (Granata, 2022; Alachevic & Granata, 2019). The social-demand political approach to education began to transform as it was increasingly challenged by new methods seeking to shape that expansion for goals that were not strictly educational, such as manpower building for the development of military technologies and the institutionalisation of a class of  the early post-war expansion reforms. This crucially enriches our understanding of why education has become instrumentalised independently of "neoliberalisation" (Giannone, 2019) processes of the mid-1980s. Historizing the legacy of this historical development on the work of VC committees permits to acknowledge how their partial embeddedness with the Cold War context importantly accounts for their vanguard position vis-à-vis States' NPM reforms. This study has embarked the scholarly suggestion of historically investigating the linkages between international-level developments in output-governance and national lineages (Bürgi, 2016b) by focusing on university networks, and the conclusions presented here draw out preliminary insights which suggest that the early stages of NPM in higher education were influenced by such international-national connections. Without the international legacy brought

about by their respective conduit agents, the CVCP and the CRUI would not have been able to readily and elaborately develop a significant part of their evaluation principles and methods.

National networks were also important in shaping a growing managerial leadership of the committees. VCs actively constructed political ties to advance their research and build consensus over it. Interestingly, there are specific networks which enhanced the diffusion and use of VCs' work more than others. Indeed, the study so far reveals that, broadly, key moments of VCs' success coincide with moments of strong ties with these networks, bringing to light the political nature of the highly contested process of evaluation-building and diffusion. Secondly, this process hints to how bottom-up innovations institutionalised in the sector, revealing that VCs actively confronted top-down constraints rather than adapting to them. All but hegemonic during the early stages of NPM reform, top-down discourses and practices on quality were pragmatically exploited by VCs and their networks to step into a changing policy context and modify it. This crucially helps to explain not just VCs' actions, but the fact that evaluation criteria frequently changed (as they still do) precisely because part of a power struggle between different groups.

The CVCP benefited from the political trust of Keith Joseph, who conceded that universities carry out their own efficiency study, and, further, agreed on the famous "concordat" which allowed the committee to benefit from a degree of freedom in elaborating and then proposing methods of efficiency for universities. In the Jarratt report, the CVCP was innovative in rearticulating the changing funding situation as an emergency context to be remedied *from and for* universities in response to political demands. Furthermore, the freedom to carry out the efficiency study can be said to have made VCs feel comfortable in explicitly attacking academic roles as the Steering committee did in the Jarratt report. For the CVCP, embracing cost-efficiency in its own articulation of evaluation was therefore facilitated by the trust of Keith Joseph and by Jarratt's own ties with Ibbs, who was also in the committee. Further, the CVCP used this trust to try and side-line the DES's influence. This indicates that the committee was making an opportunistic use of political support to actually propose a distancing from politics through the collaboration with the UGC. As shown, part of the work with the UGC institutionalised in the cost-biased PIs which were used in the 1986 RAE. Differently to this initial success, the CVCP's external support declined in the 1990s with the end of the binary line, an aspect which is reflected in the HEFCE's refusal of the CVCP's 1991 indicators. As the group of agents participating in evaluation widened to include also polytechnic representatives and members of the HEFCE, the CVCP lost a degree of its previous influence. Connections between VCs were therefore important, but so were those with other groups of agents, for such networks could facilitate the diffusion of the committee's indicators and approach more broadly.

For the CRUI, this research has found that important ties between VCs developed at the turn of the 1990s thanks to Berlinguer's active role at Pontignano and in shaping an advocacy coalition. It was here that networks were built also with future politicians such as Prodi. The relationship ties with Prodi remained solid when the CRUI's renewal phase (through CAMPUS) occurred. Here, Prodi's political help allowed the CRUI to enter into the European Projects and to have support during the CAMPUS project. With political consensus on the project, the CRUI's ties with industrial networks was facilitated, possibly explaining its continuation into CampusONE. Uncoincidentally, Modica replaced Berlinguer in Senate when Prodi was once again prime minister in 2006. During this function, Modica was able to use his experience and knowledge on university evaluation to advance the law proposal for an independent evaluation agency. Importantly, however, Modica did not make the proposal of the agency on his own and neither on behalf of the CRUI. At the same time, there are important continuities with the CRUI's broader vision of international evaluation dynamics and VCs' role in them. This is clearly different from suggesting that the proposal explicitly acted in the CRUI's interests, but thanks to a historical perspective on the proposal this argument can be considered seriously as a step to try and defend the committee's long work on internal governance through evaluation.

These insights show that the ways that Clark's categories (1983) are used (Capano, 2008; Amaral et al., 2002) to explain universities' responsiveness to the NPM paradigm is unfitting, for they do not explore institutional agency and account for the difference that it makes in shaping universities' behaviours. To get a concrete insight into university actions, it is not sufficient to look at the different autonomy margins introduced by States in general terms (Capano,2008) because it does not explain how autonomy was used by universities regardless of how it was envisioned by policymakers. Furthermore, when Clark revisited his model, he suggested that academic and managerial roles are increasingly intersecting, forming a 'strengthened steering core' (Clark, 1998). The suggestion made is thus that managerial functions gradually incorporated academic ones, and this applies indeed at a general level if looking at how academics have been increasingly pushed into carrying out bureaucratic or monitoring tasks. However, this is quite different from suggesting that administrative functions and academic ones have become a united front in universities' internal management. Indeed, this study reveals that VC committees, and particularly the British, were very keen to maintain administrative roles *separate* from academic ones precisely to reinforce their managerial powers over the latter. As I have shown, evaluation helped this process. Whilst asking academics to participate in information gathering and reporting exercises, this was done to maintain the power to construct indicators that academics had to work with, as well as the institution's overall planning decision-making powers, as a restricted exercise for administrative-managerial groups only. This research has so far looked only at the initial stages of how managerial powers

developed in universities, so further research would need to continue this investigation for later years. At the same time, the insights of this study are in line with observations made on the recent developments occurring US sector where there has been 'the fall of the faculty': an unprecedented and continuing rise and detachment of administrative groups from academic powers, most exemplified in the significant income gaps between the two (Ginsberg, 2011). As the study shows, therefore, Clark's categories and particularly their use to explain developments following the 1980s are too deterministic about how given types of internal organization are the result of external types of authority, and this is evident in how VCs used autonomy differently to what was expected given their surrounding external authority. VC committees made specific innovations on their prescribed autonomous powers. The CRUI worked to modify legislative prescriptions, using autonomy as a premise for drawing out complex evaluation dynamics which went beyond the responsibilities requested politically (Allulli, 1995). Indeed, evaluation empowered the autonomous powers of internal evaluators and VCs by giving a new pretext for action in universities. Similarly, the CVCP's ties with the UGC up until its demise widened the margins for the committee's work on PIs and also brought to concrete attempts to side-line the political interference of the DES by bringing the UGC to its side. As a result, Clark's categories can serve to generally contextualise state-HE relationships rather than to explain how they play out and develop. Furthermore, in Clark's model British universities tend towards marketization, yet university VCs first used evaluation to develop planning and control devices rather than to expand universities' entrepreneurial character. As seen, those most marketization-inclined proposals in the UK (Lockwood & Davies, 1985) discussed market dynamics of universities' external context to advocate stronger planning responsibilities and improved ability to plan image building and through internal executive leadership in universities. The potentially emerging university market was therefore a reason for pushing universities into using evaluation as an organizational device. In the Jarratt report, the possibility of an increasingly uncertain funding context directed towards increasing competition in the sector was treated as the premise for using evaluation to protect universities from the dangers of being exposed to such competitive dynamics. This emerged also later on when the CVCP emphasized that universities' prestige vis-à-vis that of polytechnics lie in the fact that universities' quality audit was advanced enough to protect them from being subjected to different auditing and assessment procedures. Therefore, the structural assumptions made on the British model do not account for the way that university agents acted to keep the market outside of the university sector and instead use evaluation to maintain a grasp over their functioning. This makes scholars' emphasis on the attention to the specificity of changing State-HE relationships (Neave, 1982,1988,1998; Cave et al., 1997) and university networks (Paradeise et al., 2009; Kogan & Hanney, 2000) central to how studies on higher education sectors ought to be framed. In future

research, for example it would be interesting to analyse how the CVCP's abovementioned behaviour persisted or changed within the CVCP as state demands and funding changed following 1992. The establishment of the Russell group, for example, could be conceptualised as an attempt to create an oligarchic-like internal market for research between the most prestigious universities in the sector, creating new competitive logics which would side-line other universities' competition axis.

# Conclusion: Managerial governance through evaluation: an inevitable development?

This research has examined the rise of managerial governance through evaluation in English and Italian universities by engaging with the *problematique* concerning the rise of a managerial State in the two countries. From existing studies, it seems that there are contradicting uses of evaluation for States' governance of the public university sector. If some scholars emphasize evaluation's centralizing powers (Pinto,2012), others delineate that it has instead facilitated a process of marketization (Tapper & Palfreyman, 2014;McGettigan,2013). This ongoing debate depends importantly on the assumption that evaluation personifies how neoliberalism's logics have changed since the advent of the new public management paradigm. Paradoxically, neoliberal states seem to have turned to planning and evaluation instruments in order to construct market dynamics, going against neoliberalism's founding principles. As I have argued, this debate needs more concrete historical work around the question of how managerial governance of the public sector has emerged, for the answers given so far for higher education remain focused on a contradictory reading of neoliberal States' intentions: use managerial practices to centralize power or to retreat? As discussed in this study, this perspective creates new shadows which scholars struggle to shed light on. By focusing extensively on untangling State intentions, the result is that answers have become generic: neoliberalism is seen as a broader phenomenon which can account for managerialism in higher education. Analyses often turn to defining new shades of neoliberalism to account for managerial governance in the sector rather than studying the latter in its own historical and context-specific terms. Although managerial practices and groups have been diffusing within a neoliberal context as authors suggest, this does not imply that the emergence of managerial classes in diverse contexts be treated as functional to neoliberal State governance. Scholarly studies making this argument (McGettigan, 2013; Tapper and Palfreyman, 2014; Deem et al., 2007, Pinto, 2012; Slaughter & Rhoades, 1997,2004) have a structuralist approach which ascribes key transformations to state power, taking the focus away from the politics of new public management. As I suggest, they do not allow to break down how different groups have tried to gain leverage and control over universities through evaluation. Though problematizing the several types of neoliberal governance goals driving NPM reforms, authors share the assumption that evaluation was managerial from the onset, and that managerial change occurred as states diffused evaluation practices. This hinders, as I have argued, critical and context-specific research on the role played by institutional groups such as vice-chancellor committees in shaping a specifically managerial connotation of educational evaluation. For this reason, this study has historically investigated the complex interaction between State and VC

powers in a specific historical moment, that of the emergence of evaluation in the two university sectors.

## Where does university evaluation come from?

As emphasized throughout, university evaluation is a broad category which consists of performance assessment and institutional management. NPM policies emphasize to maximise universities' efficiency. To historically understand how and why performance assessment instruments emerged in the two contexts, it has been crucial to first explore their origins and initial developments in the United States (Amadae, 2003). As chapter three has shown, an array of educational evaluation criteria developed at the OECD after having been imported there from key American defense intellectuals, who re-articulated war strategy methods into a framework for budget planning. This innovation, in turn, sparked an array of research into ways for expanding educational sectors and their ability to produce new scientists and technologies for competing with the USSR during the Cold War period. The OECD became, in this way, a key locus for the ramification of the principle of educational productivity (Elfert, 2019; Bürgi, 2016a, 2016b, 2017b). However, this was articulated in a multiplicity of diverse – and at times clashing – ways, which depended importantly on the interests of research networks. Whilst the broader construction of a Western bloc significantly sparked the search for ways to plan and evaluate scientific productivity – material and immaterial – it did not determine the specific methods devised by different research groups, particularly when methodological models were piloted within universities. American influence in the post-war international space of the 1950s and 1960s thus initially shaped the framework within which methods were articulated yet could not extend into the processes of re-articulation of the very framework.

As this research discusses, re-articulations owed much to a combination of international and context-specific factors. English and Italian researchers working in international projects based their understandings of the scope and methods of educational evaluation also through the lens of national-level developments. Their views on university governance and educational evaluation were thus influenced by both international and national debates. A key peculiarity of the PPBS model developing in the OECD's corridors between the 1960s and the 1980s is the predominantly non-political character of its researchers, which made research on evaluation a field which was importantly constructed by diverse groups who from different backgrounds experimenting with innovative methodologies and principles such as program budgeting. The CIPP model similarly developed through an emerging international group of evaluation scholars who were importantly concerned with secondary education. As this study shows, acknowledging how these international developments intersected with national research and experiments on educational development allows

to better analyse the role of State vis-à-vis that of university agents during the construction of university evaluation frameworks. Though caught up in developing NPM practices, the two States analysed in this study were more concerned with using evaluation and managerial practices in order to cut the costs of running central government through key reforms.

### What did universities adapt to?

States indeed attempted to use performance assessment practices to formulate expenditure saving policies. This was no easy task given that the functioning of the PPBS framework requires a significant restructuring of the State's internal organization, namely the relations of power between departments/ministries (Fry, 1988; Rebora et al., 2016). As political resistances shaped failures at using performance assessment for running PAs, managerial transformations remained quite limited within the State machinery. This in turn hindered States' ability to formulate tailored performance-oriented evaluation frameworks or evaluation-based governance practices in public sector institutions. The two countries' first NPM policies instead articulated performance evaluation in relation to decentralization, with the result that new autonomies being introduced more often than not left States free to retreat from the burden of dealing with the details of performance assessment and concern itself more with aspects of sectors' financial sustainability. Indeed, States' first political demands of efficiency in higher education were foremost concerned with universities' resource-management accountability and much less with implementing specific processes of teaching and research quality. This argument has been supported by an in-depth reading of the two States' first quality-oriented university reforms, which conveniently avoided an elaborated articulation of quality evaluation, leaving this task to autonomous universities. The emergence of a managerial class within the two States was therefore a much more complex and contradictory process than NPM literatures suggest (Deem et al., 2007; Deem & Brehony, 2005; Tapper and Palfreyman, 2014; Ferlie et al., 2003).

Both in England in the mid-1960s-1970s and Italy from the mid-1990s, policymakers tried to gear central government and the Public Administration towards performance assessment by using different PPBS-inspired frameworks. These were Program and Analysis Review in British departments, and –though limitedly so- the internal evaluation nucleuses (SECINs) in Italian ministries. As this research has investigated, these two frameworks were either only partially implemented or abandoned altogether after being formally introduced. In Italy's case, though evaluation was introduced to monitor civil service staff's performance, ministries were left to elaborate specific internal assessment methods. This amplified legislators' own difficulties in implementing performance assessment and the broader organizational transformations required to introduce program-budget centrally (Rebora et al., 2016). The British PAR framework was more

elaborated than Italy's SECINs, yet faced other types of challenges. These can be summarised as the inherent political dimension of program budgeting reform, first pointed out by Wildavsky (1964), to explain the endurance of more traditional State budgeting structures despite the developments of PPBS in central government. Although PPBS's technicality has been treated as an a-political exercise of organizing people and resources towards new quantitative targets, the cases analysed have shown that at a practical level it was a highly political process. Struggles formed as given State groups such as Rothschild's CPRS tried to use PPBS to redefine decision-making processes around the distribution of State resources between departments (Gray & Jenkins, 1982; Lewis, 2011). The two cases analysed have shown how the articulation of PPBS (and PPBS-inspired management practices) for reshaping the relationship between politicians and civil service around notions of performance was a highly open-ended process when it was first approached by States. If asking civil service staff to work with templates and targets were indeed effective steps towards achieving compliance to political agendas – as PAR tried to do – the gap between the formal use of performance templates and the practical outcomes of their use remained wide (Lewis, 2011). This, in turn, allowed resistances to "rational" policymaking strategies to build up and further impede the rise of managerial governance through performance assessment. As has been discussed, the result was that "performance policies" in the public sector did not really place quality considerations at the forefront. States' first introduction of managerial practices in government achieved the reduction of public expenditure rather than a more centralized governance of the university sector, on one hand, or its marketization, on the other.

Universities were in fact defunded through the mantras of cost-efficiency and expenditure accountability. In England, the mismanaged 1981 cuts were made to 'starve the beast' and pressure the sector into performing with less resources. In Italy, the autonomy and evaluation law of 1993 was driven by concerns on universities' ability to productively use State funds in light of their newly acquired self-governance autonomies. The dominant view of efficiency and quality in these NPM-oriented policies associated these concepts to public institutions' ability to communicate expenditures transparently and achieve public savings. At the same time, however, this is as far as the initial stages of the NPM reforms can be considered as a determinant force shaping universities' turn to educational evaluation. All but adaptive, university *VCs performed evaluation rather than being evaluated by performance* guidelines introduced by policymakers. This is most evident in the evaluation methods and processes which VC committees constructed, which carry starkly different lineages to those of State reforms. Further, this difference is strong even when keeping account of the fact that VC committees interacted with the changing policy environment and thus were aware of political developments. As I have argued, the historical lineages of VC committees' first educational

evaluation methods and principles owed much to developments occurring in international research arenas, especially at the OECD.

## Evaluated through performance of performing evaluation?

VC committees' articulation of the international CIPP and PPBS models emphasized the instrumental use of evaluation devices. The committees often defined and constructed evaluation as that tool through which top university levels could induce new behaviours which would be defined, first, and assessed, secondly, by managerial roles. By emphasizing resource planning in the UK and the safeguarding of autonomy in Italy, VCs innovatively tailored international evaluation models to their specific university context. In this way, during the early phases of the NPM reforms, VCs responded to policymakers' demands of change and efficiency in a somewhat opportunistic way. As discussed, they re-defined political demands of financial efficiency on new axis which were different to the political ones. In so doing, vice-chancellors innovated the models to their disposal. Both committees constructed evaluation responsibilities as intrinsically tied to those of executive leaders. By defining evaluation practitioners as planners and governance advisors, British vice-chancellors tried to build consensus around the notion of the vice-chancellor as an executive manager who requires evaluation criteria to carry out functions. The CRUI, too, used the CIPP evaluation model and strengthened its managerial purpose when proposing that internal evaluation place governing boards at the centre of defining evaluation indicators and institutional objectives.

Though influenced by international models, VC committees did not merely import or replicate these in their respective national contexts. They actively re-defined power hierarchies so that performance assessment could function as a top-down exercise where VCs would have new powers of defining the spaces of action of academics thanks to evaluation. Internal evaluation was recast from an accountability exercise substantiating resource use – as States had requested – to a managerial one, where performance criteria played a part in reshaping the internal organization of university activities. What began as a limit, that of demonstrating the appropriate use of funds, developed into an opportunity for exerting more and new kinds of power within universities. In this way, committees frequently took it upon themselves to expand their role in evaluation by defining not just processes of internal evaluation, but also the boundaries of satisfactory and unsatisfactory academic work, namely of teaching and research. Indeed, they engaged importantly in the construction of research quality indicators and actively attempted to institutionalize them – successfully so in the case of the CVCP's role in the 1986 RAE. Building this nexus between evaluation and managerial governance, however, was not a simple nor inevitable process. As shown, VCs were often constrained by challenges coming from within the sector and from politics. These, for example, took the form of alternative evaluation criteria being proposed by other bodies (such as the Osservatorio's "VPS" research evaluation model), academic criticisms to their content (such as

British academics' critique of the 1986 RAE's cost-bias) and, finally, pressures coming from State-devised evaluation and funding bodies (CIVR and HEFCE). These challenges often impeded the legitimation of VC committees' criteria, highlighting the complexity of what is often considered a progressive intensification of managerialism in higher education. As the study has discussed, these difficulties were even more marked in the case of VCs' proposed research quality evaluation criteria, with respective States often trying to alter each round of national evaluation exercises by establishing new evaluation powers in the sector or by emphasizing political control over research funding. The development of research evaluation deserves thus further agency-centred historical research on the specific developments of State-HE relationships. It would be ideal in future research to explore the suggestion made in this study that UK policymakers were, to an extent, latecomers to university evaluation and began to step into existing evaluation arrangements as they began to realise that this practice could in some ways become a university sector-led exercise. In other moments, abovementioned challenges pushed VC committees to intensify their research on evaluation and find means for better institutionalising a university-tailored definition of evaluation.

Scholars have scarcely studied the multiple paths taken by vice-chancellor committees on management and evaluation models, convinced that the rise of evaluation in the university sector is exhaustively explained by State policy and ideology. When embarked on through an agency-centered approach, however, a historical investigation of the vice-chancellor committees reveals the difference that their constant and proactive engagement made to the broader development of managerial governance in universities. The research has argued that, in their respective national contexts, vice-chancellor committees were more methodologically informed than political agents when the first efficiency and evaluation policies were advanced. This has also allowed to gain perspective on the ways in which *managerial practices and evaluation ones developed together*. The different evaluation methods proposed by VCs borrowed from existing international evaluation and management models that provided cues on decision-making. Yet, VC committees' further elaboration and research imbued these 'international' models with additional and specific interests which shaped them into concrete governance devices. These interests are hardly generalizable, and are more comprehensible when analysed in their context. When reified a-politically as a case of adaptation to strong external pressures, VCs' interests in evaluation simply don't emerge. At a general level, a common interest that can be identified between the two vice-chancellor committees was that of acquiring major governance powers over universities more generally. This emerges from the ways in which the committees constantly re-defined evaluation and management to ensure that their methods be taken into consideration both within the universities and within the education ministry/department.

**Neoliberalism and managerial governance: towards more historical research**

The insights of this study allow to build on existing problematizations on the relationship between neoliberalism and managerialism (Knafo et al., 2019, Eagleton-Pierce & Knafo, 2020; Giannone, 2019;Moini, 2017; Duménil & Levy, 2015). By tracing the distinct historical lineage of managerial and educational evaluation practices, the study has questioned whether State policies – and especially state retreat – conventionally associated to neoliberal governance can be seen as the driving force for the rise and proliferation of managerial governance in universities. In line with more recent scholarship (Paradeise et al.,2009; Bleiklie & Ferlie, 2009), the study shows how the rise of managerial university networks importantly influenced the initial stages of how NPM played out. This rebalances the widespread yet still very ambiguous view that NPM be conceptualised as a State-led appropriation of managerial devices for the purpose of either market-construction or centralization, or both (Palfreyman & Tapper, 2014, McGettigan, 2013;Pinto, 2012;Pollitt, 1990; March & Olsen, 1989). Rather than exploring the rise of managerial governance in university sectors by seeking to re-define neoliberal governance itself, this study has used an agency-centred approach to historicise the lineage of managerial practices. By showing that managerial models which used by university networks first emerged in the immediate Cold-War context for the purpose of planning and decision-making, I have therefore simultaneously informed the debate on States' relationship to this lineage, which was not as marked as that of universities and groups of researchers. Further research concerned with exploring States' managerial turn within NPM should thus attempt to understand if and how the systems analysis lineage was developed by State actors, an effort already undertaken by some (Amadae, 2003; Dutta et al., 2021). This study has given its contribution in this by exploring States' early managerial oriented reforms and capturing the successes and failures of NPM-oriented university policies specifically. By showing that the early stages of NPM in universities were marked by contrasting views on the use of managerial practices at the State and university levels, the study has shown how a managerial governance paradigm was already developing by the time neoliberals came to power in Italy and England. Performance assessment models within English universities were experimented and piloted at the turn of the 1970s, and by 1985 the Jarratt committee had elaborated extremely detailed notions for putting managerial governance into practice. What emerges in the Italian case is that the CRUI promptly re-articulated the CIPP model first used for Italian education in the 1970s and that, by the time the evaluation law emerged, it had already set-off data gathering exercises which led to the construction of the university sector's first ever performance indicators. These conclusions, in turn, recast the question of the relationship between neoliberalism and managerialism as a case of studying the historical lineages and developments of these two governance paradigms and relating them in order to understand their intersection. This historicist effort can be a crucial step in resolving the ongoing marketization vs. centralization debate which marks NPM scholarship, which as I have argued is

too engaged in forcing historical analysis into the categories of neoliberal goals. Although often conflated due to their assumed historical proximity and their evident discursive affinity, these two distinct intellectual terrains and practices of governance should continue to be studied as such. This is an exercise which will allow to better inform contradicting views on both managerialism and neoliberalism, and, simultaneously, to better conceptualise the rise and development of managerial practices in other public sectors. Further research could attempt to use agency-centred historicism to rebalance the notion that public sectors' governance turned managerial following evaluative States' pressures. Investigating the temporality and diffusion of managerial practices in relation to how different groups may have used it for governance can in this way help to debunk how agents use managerial practices to empower their role in given contexts. By instrumentalizing evaluation, VC committees importantly re-articulated existing managerial models to try and – not always in a successful way – exert leverage over internal university groups as well as sector groups and state groups. This exercise of appropriation and re-articulation of managerial models has been captured mostly thanks to this study's attention to how groups of agents interacted within a context of contestation over the definition and use of evaluation. More often than not, novelties lie in reformulations rather than in the construction of new paradigms.

# Bibliography

Abromeit, H.(1986). Privatisation in Great Britain. Annals of Public and Cooperative Economics, 57(2), pp. 153-179.

Adams, T. W., & Murphy, T. P. (1967). NASA' s University research Programs : Dilemmas & Problems on the Government- Academic Iterface. *Public Administration Review*, *27*(1), 10–17.

Alacevich, M., & Granata, M. (2019). Development and education. The OECD and southern Europe in the Mediterranean Regional Project. *Contemporanea*, *22*(2), pp. 225–256.

Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. The American economic review, 62(5), 777-795.

Allulli, G. (1995). *Organizzazione e metodi dei nuclei di valutazione nelle università italiane. Le proposte della Conferenza dei Rettori*.

Amadae, S. M. (2003). *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism*. The University of Chicago Press.

Amaral, A., Jones, G. A., & Karseth, B. (2002). Governing higher education: Comparing national perspectives. In *Governing higher education: National perspectives on institutional governance* (pp. 279-298). Springer, Dordrecht.

Amaral, A., Meek, V. L., Larsen, (Eds.).(2003). *The higher education managerial revolution?* (Vol. 3). Springer Science & Business Media.

Archibugi, F. (1974). A progress report: The "quality of life" in a method of integrated planning. Aspects of an Italian research project, "Progetto Quadro." *Socio-Economic Planning Sciences*, *8*(6), 339–345.

Archibugi, F. (1996). Towards a new discipline of planning. *Socio-Economic Planning Sciences*, *30*(2), 8–102.

Archibugi, F. (2006). Dal controllo di gestione alla programmazione strategica: passi necessari per rendere credibile l'attuazione dei programmi politici nella PA italiana Franco. In *I controlli per l'attuazione dei programmi governativi* (pp. 1–22). CNEL, Parlamentino, Roma.

Archibugi, F. (2008). *Da burocrate a manager: la programmazione strategica in Italia: passato, presente e futuro*. Rubbettino.

Aston, T. H., & Philpin, C. H. (Ed.).(1999). *The Brenner Debate*. Aakar Books.

Aucoin, P. (1990). Administrative reform in public management: paradigms, principles, paradoxes, and pendulums. Governance, 3(2), 115-137.

Bacevic, J. (2019). Knowing neoliberalism. Social Epistemology, 33(4), 380-392.

Balderston, F. E., & Weathersby, G. B. (1972). PPBS in Higher Education Planning and Management: Part I, an overview. *Higher Education*, *1*(2), pp. 191–206.

Balestracci, D. & Mirri, M. (Eds). (1991).*Università oggi: l'avvio di una fase costituente*. Franco Angeli.

Ball, S. (2009) Privatising education, privatising education policy, privatising educational research: network governance and the 'competition State'. *Journal of Education Policy*, 24:1, pp. 83-99

Banfi, A., & Viesti, G. (2015). *Antonio Banfi e Gianfranco Viesti "Merito" e "bisogni" nel finanziamento del sistema universitario italiano*. Palermo.

Barblan, A. (2002). Academic co-operation and mobility in Europe: How it was and how it will be. *Higher Education in Europe*, *27*(1-2), pp. 31-58.

Barone, C., Boffo, S., Pietro, F. Di, Moscati, R., (2010). Le trasformazioni dei modelli di governance nelle università europee, *Inchiesta*, 3

Bassanini, F. (2010). Vent'anni di riforme del sistema amministrativo italiano (1990-2010). *The Journal of European Economic History*, *1*.

Becher, T. et al (1977) "Maintaining academic quality in a period of retrenchment", paper prepared for a meeting of SRHE with CNAA and UGC, 5 November

Becher, T., & Kogan, M. (1980). *Process and Structure in Higher Education*. London: Heinemann.

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. New York: National Bureau of Economic Research

Bence, V., & Oppenheim, C. (2005). The Evolution of the UK's Research Assessment Exercise: Publications, Performance and Perceptions. *Journal of Educational Administration and History*, *37*(2), pp. 137–155.

Berdahl, R., & Shattock, M. (1984). The British University Grants Committee 1919-83: changing relationships with government and the universities. *Higher Education*, *13*(5), pp. 471–499.

Berdahl, R. (1990). Academic freedom, autonomy, and accountability in British universities. Studies in higher education, 15(2), 169-180.

Berle, A.A., & Means, G.C. (1932). *The modern corporation and private property*. Macmillan

Berrill, K. et al. (1983). *Excellence in Diversity. Towards a New Strategy for Higher Education*. SRHE.

Bleiklie, I. & Ferlie, E. (2009). *University Governance: Western European Comparative*

Bockman, J. (2011). *Markets in the Name of Socialism: The Left-Wing Origins of Neoliberalism*. Stanford University Press

Boffo, S. (1997). Evaluation and the Distribution of Power in Italian Universities. *European Journal of Education, 32 (2)*

Boffo, S., & Moscati, R. (1998). Evaluation in the Italian Higher Education System: many tribes, many territories... many godfathers. *European Journal of Education*, 33(3), 349-360.

Boltanski, L., & Chiapello, E. (2005). The new spirit of capitalism. *International journal of politics, culture, and society, 18(3)*

Borrelli, D. (2015). *Contro l'ideologia della valutazione. L'Anvur e l'arte della rottamazione dell'università*. Jouvence.

Borrelli, D., & Giannone, D. (2020). La neovalutazione al governo della società/Neoevaluation at the Government of Society. *Cartografie sociali. Rivista di sociologia e scienze umane*, *2*(8).

Bottani, N. (1990). The background of the CERI/OECD project on international educational indicators∗. International Journal of Educational Research, 14(4), 335-342.

Bottani, N. (2009). Il difficile rapporto tra politica e ricerca scientifica sui sistemi scolastici. *Fondazione Giovanni Agnelli Bottani Working Paper n.17*.

Brenner, R. (1976). Agrarian class structure and economic development in pre-industrial Europe. *Past & present*, 70(1), pp. 30-75.

Breno, E., Fava, G. A., Guardabasso, V., & Stefanelli, M. (2002). *La ricerca scientifica nelle università italiane: una prima analisi della banca dati ISI*. CRUI.

Breno, E., & Guardabasso, Vincenzo Fava, Giovanni A. Stefanelli, M. (2005). *Un aggiornamento sull' impatto della ricerca scientifica e tecnologica italiana in ambito internazionale (1981-2004) Analisi preliminare*. CRUI

Briggs, A. (1991). *The Story of the Leverhulme Trust: For Purposes of Research and Education*. The Leverhulme Trust.

Brosio, G. (1993). *Scenari di Spesa Pubblica al 2000*. Ministero del Tesoro, Commissione tecnica per la Spesa Pubblica, 3

Brown, R. (2004). *Quality assurance in higher education: The UK experience since 1992*. Routledge.

Brown, R. & Carasso, H., (2013). *Everything for sale? the marketisation of UK higher education*. Routledge.

Bruff, I. (2014). The rise of authoritarian neoliberalism. Rethinking marxism, 26(1), 113-129.

Brydan, D., & Reinisch, J. (2020). Organizing the 20th-century world: International organizations and the emergence of international public administration, 1920-1960s. Bloomsbury Academic.

Bucciarelli, C. (1982). Study and Work: The Pilot Projects of the European Community. *Western European Education*, *14*(3).

Bürgi, R. (2016a). Systemic management of schools: The OECD's professionalisation and dissemination of output governance in the 1960s. *Paedagogica Historica*, *52*(4), pp. 408–422.

Bürgi, R. (2016b). The Free World and the Cult of Expertise: The Rise of the OECD ' s Educationalizing Technocracy. *International Journal of Higher Education*, *6*, pp. 159–175.

Bürgi, R. (2017a). *Die OECD und die Bildungsplanung der freien Welt. Die OECD und die Bildungsplanung der freien Welt*. Verlag Barbara Budrich.

Bürgi, R. (2017b). Engineering the Free World: The Emergence of the OECD as an Actor in Education Policy, 1957–1972. In M. Schmelzer & M. Leimgruber. (Ed.).*The OECD and the International Political Economy Since 1948*. Springer International Publishing.

Bürgi, R. (2019). Learning Productivity: The European Productivity Agency—An Educational Enterprise. In C. Ydesen (Ed.).*The OECD's Historical Rise in Education*. Palgrave Macmillan, Cham.

Bürgi, R., & Tröhler, D. (2018). Producing the 'right kind of people': the OECD education indicators in the 1960s. In Lindblad, S., Pettersson,D., & Popkewitz, T. S. (Ed.). *Education by the numbers and the making of society: The expertise of international assessments*. Routledge.

Burnham, J.(1941).*The Managerial Revolution. What is happening in the world*. John Day Company, NY

Capano, G. (1991). Tendenze recenti della riforma amministrativa in Italia: alcune riflessioni in chiave comparata. *Rivista trimestrale di scienza dell'amministrazione*.

Capano, G. (2002). La riforma universitaria. L'anarchica attuazione di un disegno tecnocratico. *il Mulino*, *51*(6), pp. 1154-1163.

Capano, G. (2003). Administrative traditions and policy change: When policy paradigms matter. The case of Italian administrative reform during the 1990s. *Public Administration*, *81*(4), pp. 781–801.

Capano, G. (2008). *Il governo degli atenei.* In R. Moscati, M. Vaira (Ed.), *L'università di fronte al cambiamento*. Il Mulino, Bologna.

Capano, G. (2014). The re-regulation of the Italian university system through quality assurance. A mechanistic perspective. *Policy and Society*, *33*(3), pp. 199–213.

Capano, G., Regini, M., & Turri, M. (2017). *Salvare l'università italiana. Oltre i miti ei tabù*. Bologna: Il Mulino.

Carpita, M. (2016). Presente e futuro dei sistemi di valutazione dell'università. *Statistica & Società*, VI(1).

Cassese, S. (2013). L'ANVUR ha ucciso la valutazione, viva la valutazione!. *il Mulino*, 62(1), pp. 73-79.

Catalano, G., & Silvestri, P. (1992). *Il finanziamento e la distribuzione delle risorse nel sistema universitario Italiano. Ministero del Tesoro. Commissione tecnica per la spesa pubblica.* (Vol. 2).

Cave, M., Hanney, S., Henkel, M.(1995). Performance measurement in higher education revisited. *Public Money & Management*, 15(4), pp. 17–23.

Cave, M., Hanney, S., Henkel, M., & Kogan, M. (1997). *The Use of Performance Indicators in Higher Education* (Third Edit). London: Jessica Kingsley Publishers.

Cave, M., Kogan, M., & Hanney, S. (1989). Performance measurement in higher education. *Public Money and Management*, *9*(1), pp. 11–16.

CBO. (1993). *Using Performance Measures in the Federal Budget Process*.

CENSIS. (1991). *Per un servizio nazionale di valutazione. Esperienze straniere e proposte per l'Italia*. (G. Allulli, Ed.), *Collana Studi e documenti degli Annali della Pubblica Istruzione*. Le Monnier.

Centeno, V. G. (2018). *Elaboration of the OECD's Educational Agendas: Framed from Above, Fed from Below, Determined in Interaction. A study on the Recurrent Education Agenda*. Peter Lang.

Centeno, V. G. (2019). The Birth of the OECD's Education Policy Area. In C. Ydesen (Ed.), *The OECD's Historical Rise in Education*. Palgrave Macmillan.

Chamayou, G. (2018). La société ingouvernable: une généalogie du libéralisme autoritaire. La fabrique éditions.

Checkland, P. B. (1980). The origins and nature of hard systems thinking. In R. Spear & M. Lockett (Eds.).*Organizations as systems*. Taylor & Francis.

Checkland, P. (1983). OR and the systems movement: mappings and conflicts. Journal of the Operational Research Society, 34(8), 661-675.

CIVR. (2004). Bando VTR 2001-2003 del 16 marzo 2004.

CIVR. (2006). *Linee Guida per la Valutazione della Ricerca*.

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

Clark, B. R. (1995). *Places of inquiry: Research and advanced education in modern universities*. University of California Press.

Clark, B. R. (1998). *Creating entrepreneurial universities: organizational pathways of transformation. Issues in Higher Education*. Elsevier Science Regional Sales, New York

Clowse, B. B. (1981). *Brainpower for the Cold War: The Sputnik crisis and National Defense Education Act of 1958*. Greenwood.

CNVSU. (2005). *Nono Rapporto sullo Stato del Sistema Universitario*.

Coombs, P. H. (1968). *The World Educational Crisis*. Oxford University Press.

Coombs, P. H. (1970). *What is Educational Planning?* Paris. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000076671

Coombs, P. H., & Jacques, H. (1972). Managing Educational Costs. Retrieved from https://eric.ed.gov/?id=ED074585

Corradi, S. (1998). *Le Conferenze dei rettoi delle università, italiana (CRUI), comunitaria, paneuropea (CRE), mondiale (AIU-IAU), Storia e documenti*. ARACNE editrice.

Cozzolino, A. (2019). Reconfiguring the State: executive powers, emergency legislation, and neoliberalization in Italy. *Globalizations*, 16(3), pp. 336–352.

Cozzolino, A., & Giannone, D. (2019). State Transformations and Neoliberalization in Italy: A Critical Discourse Analysis of Governments' Political Economy, 1988-2009. *New Political Science*, 41(3)

CRE (29 May 1986a). List of participants of the Maastricht Seminar of 27 June - 2 July 1986. [List of participants] (CRE.2.1.1-159). Historical Archives of the European Union, European University Institute. Accessed at https://archives.eui.eu/en/fonds/473855?item=CRE-159 on 20 May 2019.

CRE (1986b). Checklist for preparatory Reflection of the Maastricht Seminar (CRE.2.1.1-159). Historical Archives of the European Union, European University Institute. Accessed at https://archives.eui.eu/en/fonds/473855?item=CRE-159 on 20 May 2019.

CRE (1986c). Notes on the Management Seminar (CRE.2.1.1-159). Historical Archives of the European Union, European University Institute. Accessed at https://archives.eui.eu/en/fonds/473855?item=CRE-159 on 20 May 2019.

CRE (1994). The CRE/IMHE Management training seminar for newly appointed executive heads of universities. Origins, evolution, perspectives. A report to UNESCO (contract no.105.012-4).Historical Archives of the European University Institute. Accessed in person on 1-3 December 2021.

CRUI (1995).*Progetto Pilota Europeo valutazione della qualità nel settore dell'istruzione superiore. Italia: rapporto Nazionale.* CRUI

CRUI. (1996). *Dati universitari 1992/3. Vol. 1*.

CRUI. (1998). *Dati universitari 1995/96. Vol.4*

CRUI. (1999). *Indicatori e schemi di valutazione della ricerca scientifica. Il metodo di valutazione per aree disciplinari della conferenza dei Rettori.*

CRUI. (2002). *Nuova università al lavoro*. Roma.

CRUI, (2004). Sulla Governance: Principi fondamentali e linee Guida. *Documento sul futuro assetto governativo degli Atenei italiani approvato all'unanimità dall'Assemblea Generale della CRUI, il, 17.*

Cuenin, S. (1987). The Use of Performance Indicators in Universities: An International Survey. *International Journal of Institutional Management in Higher Education*, 11(2), pp. 117-39.

CVCP. (1968). [Italian British Rectors meeting, documentation] (MSS.399/3/IBR/1). Modern Records Archive Centre, Warwick University, UK.

CVCP. (1985a). *Report of the Steering Committee for Efficiency Studies in Universities*. London.

CVCP .(1985b). *National Data Study*. London.

CVCP. (1986). *Reynolds Report. Academic Standards in the Universities*. London.

CVCP/UGC. (1986). *Performance Indicators in Universities: A first Statement by a joint CVCP/UGC working group*. London.

CVCP/UGC. (1987a). *Performance Indicators in Universities: A second Statement by a joint CVCP/UGC working group*. London.

CVCP/UGC. (1987b). *University Management Statistics and Performance Indicators*. London.

CVCP (1993). Research Performance Indicators: Annual Publications survey. First Annual publications survey: calendar year 1991. N/93/51. London: CVCP.

Dahler-Larsen, P. (2018). Making citations of publications in languages other than English visible: On the feasibility of a PLOTE-index. *Research Evaluation*, *27*(3), pp. 212–221.

Dahler-Larsen, P. (2019). *Quality: from plato to performance*. Springer.

Dahler-Larsen, P. (2020). Lost and found, and Lost Again? Tracing Moral-Political Values in Understandings of the Evaluation Wave. *Cartografie sociali. Rivista di sociologia e scienze umane*, 2(8).

Dainton, F. (1971). Two Views of Brnish Science. *Nature*, 234.

Davies, J. (1985). The Agendas for University Management in the Next Decade. In *Universities: The Management Challenge*. Society for Research into Higher Education.

Davies, J. L. (1980). The Role of the University Rector in Policy Formation in the Contemporary context of steady State and uncertainty. *CRE Information New Series*, 51(3rd Quarter).

Davies, J. L., & Melchiori, G. S. (1982). Developing the Image and Public Reputation of Universities: The Managerial Process. *International Journal of Institutional Management in Higher Education*, 6(2), pp. 87-108.

Davies, W. (2016). *The limits of neoliberalism: Authority, sovereigny and the logic of competition.* Sage.

De Bellis, N. (2005). *La citazione bibliografica nell ' epoca della sua riproducibilità tecnica:*

*Bibliometria e analisi delle citazioni dallo Science Citation Index alla Cybermetrica.*

Deem, R., & Brehony, K. J. (2005). Management as ideology: the case of 'new managerialism' in higher education. *Oxford Review of Education*, 31(2), 217-235

Deem, R., Hillyard, S., & Reed, M. (2007). *Knowledge, higher education, and the new managerialism: The changing management of UK universities*. Oxford University Press.

Deming, W. E. (1967). *What happened in Japan? Society of Quality Control Engineers*.

Deming, W. E. (1982). *Quality, productivity, and competitive position*. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

DES, 1987. Higher Education: Meeting the Challenge, Cm.114, London: HMSO.

DES, 1991. Higher Education: A new framework, Cm. 1541, London.HMSO.

DfE. (1992). Further and Higher Education Act. https://www.legislation.gov.uk/ukpga/1992/13/contents?view=plain

Di Mascio, F., & Natalini, A. (2018). *Oltre il New Public Management: Le riforme amministrative tra meccanismi e contesti*. Carocci. Retrieved from http://library1.nida.ac.th/termpaper6/sd/2554/19755.pdf

Dill, D. (1998). Evaluating the 'Evaluative State': Implications for Research in Higher Education. *European Journal of Education*, 33(3), pp. 361–377.

Directorate for Scientific Affairs. (1968). *Budgeting, programme analysis and cost-effectiveness in educational planning*.

Dorrell, S. (1993). Public sector change is a world-wide movement. *Speech by the Financial Secretary to the Treasury to the Chartered Institute of Public Finance and Accountancy, London*, *23*.

Duménil, G., & Lévy, D. (2015). Neoliberal managerial capitalism: Another reading of the Piketty, Saez, and Zucman data. International Journal of Political Economy, 44(2), 71-89.

Dutta, S. J., Knafo, S., & Lovering, I. A. (2021). Neoliberal failures and the managerial takeover of governance. *Review of International Studies*, 1-19.

Eagleton-Pierce, M., & Knafo, S. (2020). Introduction: The political economy of managerialism. *Review of International Political Economy*, 27(4), pp. 763–779.

Edwards, P. N. (1996). *The closed world: Computers and the politics of discourse in Cold War America*. MIT Press.

Elfert, M. (2019). The OECD, American Power, and the Rise of the "Economics of Education" in the 1960s. In *The OECD's Historical Rise in Education*. Palgrave Macmillan.

Elfert, M., & Ydesen, C. (2020). The Influence of the United States on the Rise of Global Governance in Education: The OEEC and UNESCO in the post–Second World War Period. In *Organizing the 20th-Century World: International Organizations and the Emergence of International Public Administration*, 1920-1960s.

European Education Ministers. (1976). Resolution of the Council and of the Ministers of Education concerning measures to be taken to improve the preparation of young people for work and to facilitate their transition from education to working life. *Official Journal of the European Communities*, *C,38*(13.12.76).

Eustace, R. (1982). British higher education and the State. *European Journal of Education*, 17(3),

pp.283–294.

Felice, E. (2015). *Ascesa e declino: storia economica d'Italia*. Bologna: Il Mulino.

Ferlie, E., Fitzgerald, L., & Pettigrew, A. (1996). *The new public management in action*. Oxford University Press, USA.

Ferlie, E., Hartley, J., & Martin, S. (2003). Changing Public Service Organizations: Current Perspectives and Future Prospects. *British Journal of Management*, 14(s1)

Ferraresi, F. (1982). Il rapporto FORMEZ sullo stato della pubblica amministrazione. *Rivista Trimestrale Di Scienza Dell'amministrazione*, (1), pp. 55–65.

Fielden, J. (1969). *Analytical planning and improved resource allocation in British universities: a critical evaluation of American PPBS and its potential in Great Britain; a research report*. London.

Fielden, J. (1973). Management Information in Higher Education: The Gulf Between Theory and Practice. *Educational Administration Bulletin*, 1(2), pp. 16–23.

Fielden, J. (1978). The Cost of Innovation and Change in Education. *Programmed Learning and Educational Technology*, *15*(1), pp. 16–25.

Fourcade, M. (2006). The construction of a global profession: The transnationalization of economics. *American journal of sociology*, 112(1), 145-194.

Frazer, M, & Craft, A. (1992). *Quality Assessment in Higher Education: Proceedings of an international conference in Hong Kong, 1991*.

Frazer, Malcolm. (1997). Report on the Modalities of External Evaluation of Higher Education in Europe: 1995-1997. *Higher Education in Europe*, *22*(3), pp. 349–401.

Fry, G. K. (1988). The Thatcher Government, the Financial Management Initiative, and the 'New Cml Service.' *Public Administration*, *66*(1), pp. 1–20.

Fulton, L. (1968). Fulton Committee Report-the Report of the Committee on the Civil Service (Cmnd. 3638).

Giannini, M. S. (1979). *Rapporto sui principali problemi dell'amministrazione dello Stato*. *Il Foro italiano*.

Giannone, D. (2019). *In perfetto Stato: Indicatori globali e politiche di valutazione dello Stato neoliberale*. Mimesis.

Giarda, P. (1966). Le dimensioni dell'unità amministrativa locale: Motivi tecnici economico e finanziari. *Rivista Internazionale Di Scienze Sociali*, *37*(Anno 74), pp. 43–67.

Giarda, P. (1993). L'unviersità italiana tra diversificazione, inefficienza e autonomia finanziaria. *Rivista Internazionale Di Scienze Sociali*, 101(1), pp. 69–90.

Gibbs, P. (2001) Higher Education as a Market: A problem or solution? *Studies in Higher Education*, 26(1), pp. 85-94

Ginsberg, B. (2011). The fall of the faculty: The Rise of the All-Administrative University and Why It Matters Oxford University Press.

Giovinazzi, L. (2019). The re-emerging debate over the rise of the evaluative State in the British university sector. *Cartografie Sociali. Rivista di sociologia e scienze umane*, 4(8).

Granata, M. (2022). Planning in cold-war Italy. *Journal of Modern Italian Studies*, 27(1), pp. 63-91.

Gray, A., & Jenkins, B. (1982). Policy Analysis in British Central Government: the experience of PAR. *Public Administration*, 60(4), pp. 429–450.

Gray, A., & Jenkins, B. (2002). Policy and program evaluation in the United Kingdom: A reflective State. *Interational Atlas of Evaluation*.

Greenwood, R., Hinings, C. R., & Ranson, S. (1977). The politics of the budgetary process in English local government. *Political Studies*, 25(1), pp. 25–47.

Guba, E. G., & Stufflebeam, D. L. (1970). Evaluation: The Process of Stimulating Aiding and Abetting Insightful Action. *Monograph Series in Reading Education*. Retrieved from https://files.eric.ed.gov/fulltext/ED055733.pdf

Hasley, H., & Trow, M.(1971). *The British Academics*. Faber and Faber

Henkel, M. (1991). The new "Evaluative State." *Public Administration*, 69(1), pp. 121–136.

Hirsch, W. Z. (1965). *Education in the program budget*. RAND Corporation.

Hirsch, W.Z. (1968). The Budget as an Instrument for Medium and Long-Range Planning and Programming of Education, *Directorate for Scientific Affairs, OECD*

Hirsch, W. Z. (1973). Program Budgeting in the United Kingdom. *Public Administration Review*, pp. 120–128.

Hitch, C.J & McKean, R.N. The Economics of Defense in the Nuclear Age. Santa Monica, CA: RAND Corporation, 1960. (Retrieved from https://www.rand.org/pubs/reports/R346.htm)

HMSO. (1971) A Framework for Government Research and Development. Cmnd. 4814, HMSO: London

HMSO. (1988). Education Reform Act.

Holmes, G., (1993). Quality Assurance in Further and Higher Education: A sacrificial lamb on the altar of managerialism. *Quality Assurance in Education*, 1(1), pp. 4-8.

Hood, C. (1991). A public management for all seasons?. Public administration, 69(1), 3-19.

Hood, C., & Jackson, M. W. (1991). *Administrative argument*. Dartmouth Publishing Company.

Hughes, O. (2003). *Public Management and Administration*. Palgrave Macmillan.

Human Synergistics (1985). *The project planning situation*. Plymouth, MI: Human Synergistics, Inc.

IMHE (2002). *Higher Education Management*. Journal of the Programme on Institutional Management in Higher Education, 13(2)

Jarratt, A. (1986). The management of universities. *Journal of the Royal Society of Arts*, 134(5363), pp. 742-753.

Jones, H. C., Lockwood, G., & Mackenzie, N. (1969). *The University of Sussex. IIEP seminar*

Jones, H.C., Lockwood, G., & MacKenzie, N.(1971). The University of Sussex. in Onushkin, 1971 *Planning The Development of Universities,Vol.I,* IIEP, UNESCO

JPIWG. (30 September 1993). Report on the conversion of the data submitted to the 1992 Research

Assessment Exercise into performance indicators. [Report of the Research sub-group of the Joint

Performance Indicators working group] (UWA/PUB/11/6/1-7) Modern Records Archive Centre, Warwick University, UK.

Judt, T. (2005). *Postwar. A history of Europe since 1945*. Penguin Books.

Katz, D., & Kahn, R. L. (1966). *The Social Psychology of Organizations. Wiley.* Wiley.

Kershaw, J. A., & McKean, R. N. (1959). *System analysis and Education*.

Kevles, D. J. (1995). *The physicists: The history of a scientific community in modern America*. Harvard University Press.

Khan, A. G. (1971). Programme on Institutional Management in Higher Education. CERI. Technical Report.

Khurana, R. (2010). *From higher aims to hired hands: The social transformation of American business schools and the unfulfilled promise of management as a profession*. Princeton University Press.

King, A. (2007). *Let the Cat Turn Round: One Man's Traverse of the Twentieth Century. CPTM.* CPTM.

Knafo, S. (2010). "Critical Approaches and the Legacy of the Agent/Structure Debate in International Relations." Cambridge Review of International Affairs 23(3):493– 516.

Knafo, S. (2020). Neoliberalism and the origins of public management. *Review of International Political Economy*, *27*(4).

Knafo, S., Dutta, S. J., Lane, R., & Wyn-Jones, S. (2019). The Managerial Lineages of Neoliberalism. *New Political Economy*, *24*(2), 235–251.

Knafo, S., & Teschke, B. (2017). *The rules of reproduction of capitalism: A historicist critique*.

Kogan, M. (2018). *Educational policymaking: A study of interest groups and parliament*. Routledge.

Kogan, M., & Kogan, D. (1983). *The Attack on Higher Education*. Kogan Page.

Kogan, M., & Hanney, S. (2000). *Reforming higher education* (Vol. 50). Jessica Kingsley Publishers.

Konings, M. (2010). The pragmatic sources of modern power. *Archives Européennes de Sociologie/European Journal of Sociology/Europäisches Archiv für Soziologie*, 55-91.

Kristiansen, M. B., Dahler-Larsen, P., & Ghin, E. M. (2019). On the Dynamic Nature of Performance Management Regimes. *Administration and Society*, *51*(6), 991–1013.

Labour Party (1963) *Report of the sixty-second Annual Conference of The Labour Party*. Transport House

Lafferty, J. C., & Pond, A. W. (1974). The desert survival situation: A group decision making experience for examining and increasing individual and team effectiveness. *Human Synergistics*.

Lanzalaco, L., (1995). *Istituzioni, organizzazioni, potere: introduzione all'analisi istituzionale della politica*. La Nuova Italia Scientifica.

Lewis, J. (2011). The Search for Coordination: The Case of the Central Policy Review Staff and

Social Policy Planning, 1971-77. *Social Policy and Administration*, *45*(7), 770–787.

Leydesdorff, L. (2005). Evaluation of research and evolution of science indicators. *Current science*, 89(9), 1510-1517.

Leydesdorff, L. (2008). Caveats for the use of citation indicators in research and journal evaluations. *Journal of the American Society for Information Science and Technology*, *59*(2), 278–287.

Linsi, L. (2020). The discourse of competitiveness and the dis-embedding of the national economy. *Review of International Political Economy*, *27*(4), 855-879

Locke, R. R., & Spender, J. C. (2011). *Confronting managerialism: How the business elite and their schools threw our lives out of balance*. Zed Books Ltd.

Lockwood, G. (1972a). Planning in a University. *Higher Education Quarterly*, *1*(4), 409–434.

Lockwood, G. (1972b). *University planning and management techniques*. Directorate for Scientific Affairs, OECD

Lockwood, G, & Fielden, J. (1973). *Planning and Management in Universities: a study of British universities*. Chatto & Windus.

Lockwood, G. (1979). The role of the registrar in Today's University. *Higher Education*, *8*(3), 299–320.

Lockwood, G. (1981). *An Analysis of the Planning Process in 1968-73 in the Contexts of the History of the University of Sussex and of the Management of Universities*.University of Sussex. PhD Dissertation.

Lockwood, G, & Davies, J. L. (1985). *Universities: The management challenge*. Guildford, Surrey: Society for Research into Higher Education.

Madaus, G. F., Stufflebeam, D., & Scriven, M. S. (1983). Program evaluation. A Historical Overview. In Madaus, G. F., Stufflebeam, D., & Scriven, M. S. (Eds.).Evaluation models (pp. 3-22). Springer, Dordrecht.

Madaus G.F., Kellaghan T. (2000) Models, Metaphors, and Definitions in Evaluation. In: Stufflebeam D.L., Madaus G.F., Kellaghan T. (Eds.).Evaluation Models. Evaluation in Education and Human Services, vol 49. Springer, Dordrecht.

Marris, R. (1964).*The economic theory of managerial capitalism* . Macmillan.

Marrow, A. (1975). Management by participation. In Cass, E. & Zimmer, F.(Eds.):*Man and work in society.* Van Nostrand Reinhold

Martinoli, G. (1962). *L'università nello sviluppo economico italiano*. Roma: Giuffrè.

Marx, K., Engels, F. (1956). The Holy Family, trans. R. Dixon.

Meyer, H. (2002). The new managerialism in education management: corporatization or organizational learning? J*ournal of Educational Administration*, 40(6), pp. 534 - 551

McGettigan, A. (2013). *The great university gamble: Money, markets, and the future of higher education.* London: Pluto Press.

Mcsweeney, B. (1988). Accounting for the audit commission. *The Political Quarterly*, *59*(1).

Miliband, R. (1970). The capitalist State: reply to Nicos Poulantzas. *New Left Review*, *59*(1), 53-60.

Miliband, R. (1973). Poulantzas and the capitalist State. *New Left Review*, *58*(December), 83–92.

Minelli, E., Rebora, G., & Turri, M. (2015). Quest for accountability: exploring the evaluation process of universities. *Quality in Higher Education*, *21*(2), 103–131.

Mirandola, R., Facoltà, P., Squarzoni, A., & Facoltà, G. (2003). *Guida alla valutazione dei corsi di studio*.

Modica, L., & Stefani, E. (1997). *Valutazione delle attività didattiche universitarie. Le esperienze condotte dalla CRUI.*

Modica, L. (2006) Minutes of the meeting of 7th permanent commission, consulted at http://www.senato.it/japp/bgt/showdoc/frame.jsp?tipodoc=SommComm&leg=14&id=176410 &part=doc_dc-sedetit_pi.

Moini, G. (2017). New Public Management e neoliberismo. Un intreccio storico. *Economia & lavoro*, 51(2), 71-80.

Moodie, G. C., & Eustace, R. (1974). *Power and Authority in British Universities*. Allen and Unwin

Moodie, G. (Ed.).(1986).Standards and Criteria in Higher Education Guildford: SRHE/NFER Nelson

Moscati, R., & Vaira, M. (Ed.). (2008). L'università di fronte al cambiamento: realizzazioni, problemi, prospettive. Il Mulino.

Murray, P. (2012). 'Can I write to you about Ireland?': John Vaizey, the Ford Foundation and Irish educational policy change, 1959–1962 [document study]. *Irish Educational Studies*, *31*(1), 67–75. https://doi.org/10.1080/03323315.2011.601889

National Research Council. (1995). Education Indicators Project (INES). In *International Comparative Studies in Education: Descriptions of Selected Large-Scale Assessments and Case Studies*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9174

Neave, G. (1982). The changing boundary between the State and higher education. *European Journal of Education*, 17(3), pp. 231-241.

Neave, G. R. (1985). Elite and mass higher education in Britain: a regressive model? *Comparative Education Review*, 29(3), pp. 347–361.

Neave, G. (1988). On the cultivation of quality, efficiency, and enterprise: An overview of recent trends in higher education in Western Europe, 1986-1988. *European Journal of Education*, 23(1-2), pp. 7–23

Neave, G. (1995). On living in interesting times: higher education in Western Europe 1985-1995. European Journal of Education, 30(4), pp. 377-393.

Neave, G. (1998). The evaluative State reconsidered. *European Journal of Education*, *33*(3), pp. 265-284

Nicoli, M. (2010). Regimi di verità nell'impresa postfordista. Esercizi filosofici, 5(1).

Nicoli, M. (2011). Il fascismo del manager. Aut Aut, (350), 61.

Nicoli, M., & Paltrinieri, L. (2014). Il management di sé e deglialtri.

Novick, D. (1954). Efficiency and economy in government through new budgeting and accounting procedures (Vol. 254). Rand Corporation.

Novick, D. (1973). *Current practice in program budgeting (PPBS): analysis and case studies covering government and business*. New York: Crane, Russak.

OECD, CSTP (1964). Educational Investment Planning: The Need for a Critical Review of Problems and Methods (STP (64)38, august 28, 1964), 3.

OECD. (1966). *The European reconstruction, 1948–1961: Bibliography on the Marshall plan and the organisation for European economic co-operation*. Paris.

OECD. (1967). *Methods and statistical needs for educational planning*. Paris

OECD. (1968). *Budgeting, programme analysis and cost-effectiveness in educational planning*.

OECD/CERI. (1971). *Centre for Educational Research and Innovation: Purpose, Programmes, Progress*.

OECD (1974). *Indicators of Performance of Educational Systems*. Paris

OECD. (1982). *OECD List of social indicators*. Paris

OECD. (1997). PUMA, Benchmarking, evaluation, and strategic management in the public sector, 1996 Meeting of the Performance Management Network of the OECD's Public Management service, OECD/GD (97)50. Paris.

Olssen, M. & Peters, M.A.(2005). Neoliberalism, higher education, and the knowledge economy: from the free market to knowledge capitalism, *Journal of Education Policy*, 20(3), 313-345.

Ongaro, E. (2006). The dynamics of devolution processes in legalistic countries: organizational change in the Italian public sector. *Public Administration*, *84*(3), 737-770.

Ongaro, E., & Valotti, G. (2008). Public management reform in Italy: Explaining the implementation gap. *International Journal of Public Sector Management*, *21*(2), 174–204. https://doi.org/10.1108/09513550810855654

Ongaro, E. (2009). *Public management reform and modernization: Trajectories of administrative change in Italy, France, Greece, Portugal, and Spain*. Edward Elgar Publishing.

Onushkin, V. (1969). Some Problems in the Planning of Higher Education in the USSR.

Onushkin, V. (1971). *Planning the development of Universities - Vol.I*. IIEP. Paris.

Onushkin, V. (1975). *Planning the development of Universities - Vol. IV*. IIEP. Paris.

Osborne, D., & Gaebler, T. (1992). *Reinventing government: How the entrepreneurial spirit is transforming the public sector from schoolhouse to Statehouse*. Reading, Mass., Addison-Wesley.

Osservatorio. (1997a). *Incontro Nazionale sulla Valutazione del Sistema Universitario. Ruolo, organizzazione e attività dei nuclei di valutazione interna delle università*.doc 5/97

Osservatorio. (1997b). *Incontro Nazionale sulla Valutazione del Sistema Universitario. La valutazione del sistema universitario: l'organizzazione attuale e metodologie di lavoro*.doc 6/97

Osservatorio. (1998). *Il riparto della quota di riequilibrio del fondo per il finanziamento ordinario delle università proposte per il triennio 1998-2000*.

Osservatorio. (1999). *Proposta per un programma di valutazione della produzione scientifica nelle università. Programma VPS*.

Padgett, J. F., & Powell, W. W. (2012a). *The emergence of organizations and markets.* Princeton University Press.

Padgett, J. F., & Powell, W. W. (2012b). The Problem of Emergence. In *The Emergence of Organizations and Markets* (pp. 1–66). https://doi.org/10.23943/princeton/9780691148670.003.0001

Palfreyman, D., & Tapper, T. (2014). *Reshaping the university: The rise of the regulated market in higher education*. Oxford university Press.

Paolini, A., & Soverchia, M. (2014). Public Italian Universities Towards Accrual Accounting and Management Control. In *Information Systems, Management, Organization and Control* (pp. 29-45). Springer, Cham.

Paolini, A., & Soverchia, M. (2015). La programmazione delle università italiane si rinnova: riflessioni e primi riscontri empirici. *Azienda Pubblica*, *3*, 287-308.

Paolini, A., & Soverchia, M. (2017). I sistemi informativi per il controllo e la valutazione delle performance nelle università italiane. *I sistemi informativi per il controllo e la valutazione delle performance nelle università italiane*, 5-14.

Papadopoulos, G. S. (1994). *Education 1960-1990: the OECD perspective. OECD.*

Paradeise, C., Reale, E., Bleiklie, I., & Ferlie, E. (Ed.). (2009). *University governance*. Springer Netherlands.

Peck, J., & Tickell, A. (2002). Neoliberalizing space. *Antipode*, *34*(3), 380-404.v

Perellon, J.F. (2001). *The development of quality assurance policy in higher education. A comparative analysis of England, the Netherlands, Spain, and Switzerland*. University of London - Institute of Education.

Phillimore, A. J. (1989). University research performance indicators in practice: The University Grants Committee 's evaluation of British universities. *Research Policy*, *18*, 255–271.

Pinto, V. (2009). Sulla Valutazione. In *Nuovi Saperi e nuova didattica nell'Universita del nuovo millennio: Dieci anni dopo il D.M.509.*

Pinto, V. (2012). *Valutare e punire: una critica della cultura della valutazione*. Cronopio.

Pollitt, C. (2013). The Evolving Narratives of Public Management Reform: 40 years of reform white papers in the UK. *Public Management Review*, *15*(6), 899–922. https://doi.org/10.1080/14719037.2012.725761

Pollitt, C., Birchall, J., & Putman, K. (2016). *Decentralising public service management*. Macmillan International Higher Education.

Pollitt, C. & Bouckaert, G., 2004, *Public Management Reform: A comparative analysis*. New York. Oxford University Press.

Pollitt, C., & Dan, S. (2011). *The impacts of the New Public Management in Europe: A meta-analysis.* (No. Grant Agreement No. 266887).

Potì, B., & Reale, E. (2005). *New Tools for the Governance of the Academic Research in Italy: The Role of Research Evaluation*.

Poulantzas, N. (1969). The problem of the capitalist State. *New left review*, *58*(1), 67-78.

Poulantzas, N. (1973). On social classes. *New Left Review*, (78), 27.

Powell, W. W., & DiMaggio, P. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*.

Power, M. (1997). *The audit society: Rituals of verification*. Oxford: Oxford university Press.

Pritchard, R. M. O. (1993). Mergers and Linkages in British Higher Education. *Higher Education Quarterly*, *47*(2), 79–102. https://doi.org/10.1111/j.1468-2273.1993.tb01617.x

Pritchard, R. M. O. (1994). Government Power in British Higher Education. *Studies in Higher Education*, *19*(3), 253–265. https://doi.org/10.1080/03075079412331381860

Radcliffe, J. (1991). *The reorganisation of British central government*. Dartmouth Publishing Company.

Reale, E. (Ed.). (2008). *La Valutazione della ricerca pubblica. Un'analisi della valutazione triennale della ricerca*. Franco Angeli.

Reale, E., & Potì, B. (2009). Local policy legacy and moving to an 'in between'configuration. In C. Paradeise, E. Reale, I. Bleiklie, & E. Ferlie (Ed.), *University Governance: Western European Comparative Perspectives*. Springer.

Rebora, G. (2012). Venti anni dopo: il percorso della valutazione dell'università in Italia e alcune proposte per il futuro. *LIUC Papers, Serie Economia Aziendale*, *38*.

Rebora, G., Ruffini, R., & Turri, M. (2016). A Serious Game: Performance Management in Italian Ministries. *International Journal of Public Administration*. https://doi.org/10.1080/01900692.2016.1201103

Rebora, G., & Turri, M. (2008). La governance del sistema universitario in Italia: 1989-2008. *LIUC Papers, Serie Economia Aziendale*, *32*(221).

Rhoades, G., & Slaughter, S. (1997). Academic capitalism, managed professionals, and supply-side higher education. *Social Text*, (51), 9-38.

Rhoades, G., & Slaughter, S. (2004). Academic Capitalism and the New Economy: Markets, State, and Higher Education. JHU Press

Rizzi, D., & Silvestri, P. (2002). La valutazione del sistema universitario italiano: una storia recente. *CAPP - Centro Di Analisi Delle Politiche Pubbliche*.

Robbins, Lord. (1963). *Higher education: report of the Committee appointed by the Prime Minister under the Chairmanship of Lord Robbins 1961–63*. London.

Rudolph, J. L. (2002). From World War to Woods Hole: The use of wartime research models for curriculum reform. *Teachers College Record*, *104*(2), 212–241. https://doi.org/10.1111/1467-9620.00161

Rüegg, W. (Ed.).(2011).*A history of the university in Europe: Universities Since 1945 (Vol. 4)*. Cambridge University Press.

Ryan, D. (1998). The Thatcher government's attack on higher education in historical perspective. *New Left Review*, 3-32.

Saint-Martin, D. (1998). Management consultants, the State, and the politics of administrative reform in Britain and Canada. *Administration & Society*, *30*(5), 533–568.

Saint-Martin, D. (2000). *Building the New Managerialist State: Consultants and the Politics of Bureaucratic Reform in Britain, Canada, and France*. Oxford University Press.

Sanders, D. L. (1998). Issues management and the participation of large corporations in the public policy process. The Claremont Graduate University.

Senato della Repubblica, Disegno di legge, (Istituzione dell'Autorità per la valutazione del sistema delle università e della ricerca), n. 3773. XIV Legislatura, 8 febbraio 2006

Schmelzer, M. (2012).The crisis before the crisis: the 'Problems of modern society ' and the OECD , 1968 – 74. *European Review of History: Revue Europeenne d'histoire*, *19*(6), 999–1020. https://doi.org/10.1080/13507486.2012.739148

Schultz, T. W. (1963).The economic value of education. New York: Columbia University Press.

Shattock, M. (1982).How should british universities plan for the 1980s? *Higher Education*, *11*(2), 193–210. https://doi.org/10.1007/BF00139688

Shattock, M., & Rigby, G. (1983).*Resource Allocation in British Universities. SRHE Monograph 56. Society for Research into Higher Education, University of Surrey, Guildford, Surrey GU2 5XH.* (SRHE Monog). Society for Research into Higher Education.

Shattock, M. (1984).British higher education under pressure: politics, budgets and demography and the acceleration of ideas for change. *European Journal of Education*,19(2), 201-216.

Shattock, M. (1987).The last days of the University Grants Committee. *Minerva*, *25*(4), 471–485. https://doi.org/10.1007/BF01099443

Shattock, M. (2012). *Making policy in British higher education 1945-2011*. McGraw-Hill Education.

Shattock, M. (2015).*The Society for Research into Higher Education and the changing world of British higher education: a study of SRHE over its first 25 years*. Society for Research into Higher Education.

Schick, A. (2015). The road to PPB: The stages of budget reform (pp. 39-56). Routledge.

Silver, H. (2003). *Higher education and opinion making in twentieth-century England*. Psychology Press.

Sizer, J. (1979).Assessing Institutional Performance--An Overview. *International Journal of Institutional Management in Higher Education*, *3*(1), 49-75.

Sizer, J. (1981).Performance Assessment in Institutions of Higher Education under Conditions of Financial Stringency, Contraction and Changing Needs: A Management Accounting Perspective. *Accounting and Business Research*, *11*(43), 227–242. https://doi.org/10.1080/00014788.1981.9729704

Sizer, J. (1987). An Analysis of Trends Impacting Upon the UGC and British Universities: 1961-May 1979. *Financial Accountability & Management*, *31*(1), 9–28.

Sizer, J.(1987b). *An analysis of institutions' responses to constitutional recommendations in the Jarratt Report*. CVCP

Sizer, J. (1988). British Universities' responses to events leading to grant reductions announced in July 1981. *Financial Accountability & Management*, *4*(2).

Sizer, J. (1988). The politics of performance assessment: Lessons for higher education? A comment.

Sizer, J. (1989). A Critical Examination of the Events Leading up to the U. G. C.' s Grant Letters Dated 1st July 1981. *Higher Education*, *18*(6), 639–679.

Spalletti, S. (2008). The Economics of Education in Italy (1960-1975). *History of Economic Ideas, Special Issue: Economic Policy during the Planning Era in Italy: Theories, History and Documents*, *16*(1/2), 225–243.

Stufflebeam, D. L. (1966). A Depth Study of the Evaluation Requirement. *Theory Into Practice*, *5*(3), 121–133.

Stufflebeam, D. L. (1968a). Evaluation as enlightenment for decision-making. Ohio State University, Columbus Evaluation Center

Stufflebeam,D.L.(1968b).Toward a science of educational evaluation. *Educational Technology*,*8*(14), pp. 5-12.

Stufflebeam, D. L. (1983). The CIPP model for program evaluation. In Stufflebeam D.L., Madaus G.F., Kellaghan T. (Ed.). Evaluation Models (pp. 117-141). Springer, Dordrecht.

Stufflebeam D.L. (2000) Foundational Models for 21st Century Program Evaluation. In: Stufflebeam D.L., Madaus G.F., Kellaghan T. (Ed.).Evaluation Models. Evaluation in Education and Human Services, vol 49. Springer, Dordrecht.

Stufflebeam, D. L. (2000). The CIPP Model for Evaluation. In D. L. Stufflebeam, C. F. Madam, & T. Kellaghan (Ed.), *Evaluation models* (pp. 279–317). Boston: Kluwer Acdemic Publishers. https://doi.org/10.1007/978-94-010-0309-4_4

Sudama, T. (1977). PPBS and Theories of Decision-making, Bureaucracy and Politics. Political Science, 29(1), 39-56.

SVIMEZ. (1961). *Mutamenti della struttura professionale e ruolo della scuola. Previsioni per il prossimo quindicennio*. Roma: Giuffrè.

Taggart, G. J. (2004). A critical review of the role of the English Funding Body for Higher Education in the relationship between the State and Higher Education in the period 1945-2003.Dissertation. University of Bristol. retrieved at https://research-information.bris.ac.uk/ws/portalfiles/portal/34498494/398707.pdf

Tapper, T., & Salter, B. (1978). *Education and the political order: Changing patterns of class control*. Springer.

Taylor, Jim Johnes, J. (1992). The 1989 research selectivity exercise: a statistical analysis of differences in research rating between universities at the cost centre level. *Higher Education Quarterly*, *46*(1), 67–87.

Tosi, P. (2003). *Prima relazione sullo stato delle Università italiane*.

Tröhler, D. (2011). *Education systems in historical, cultural, and sociological perspectives. Vol. 1*. Sense Publishers.

Tröhler, D. (2012). *Languages of education: Protestant legacies, national identities, and global aspirations*. Routledge.

Tröhler, D. (2013). The OECD and Cold War culture: Thinking historically about PISA. *PISA, Power, and Policy. The Emergence of Global Educational Governance*, *23*(1), 141–161. Retrieved from files/617/OECD and Cold War Culture.pdf

Trow, M. (1994). Managerialism and the Academic Profession: The Case of England. *Higher Education Policy*, *7*(2), 11-18.

Turri, M. (2014). The new Italian agency for the evaluation of the university system (ANVUR ): a need for governance or legitimacy? *Quality in Higher Education*, *20*(1), 64–82. https://doi.org/10.1080/13538322.2014.889429

Turri, M., Rebora, G., & Minelli, E. (2007). Governo e gestione delle università: la centralità della valutazione. *La Misurazione Della Performance Negli Atenei: Logiche, Metodi, Esperienze*.

Tyles, M. (2008). Professor John Sizer 1938–2008. *The British Accounting Review*, *3*(40), 286–287.

UFC. (1992). Research Assesment Exercise 1992. Circular 5/92.

UNIRES. (2014). Attuazione della Legge 240/2010: i risultati della prima survey sugli Atenei statali. retrieved from http://unires.unibo.it/publication/attuazione-della-legge-2402010-i-risultati-della-prima-survey-sugli-atenei-statali/

Vaira, M. (2011). *La costruzione della riforma universitaria e dell'autonomia didattica: idee, norme, pratiche, attori*. LED Edizioni Universitarie.

Vaizey, J. (1961). Comparative Notes on Economic Growth and Social Change in Education. *Comparative Education Review*, *5*(1), 7–12. https://doi.org/10.1086/444858

Vaizey. J. (1968). Personal Impression, in Budgeting, programme analysis and cost-effectiveness in educational planning, *Directorate for Scientific Affairs, OECD*

Viesti, G. (Ed.).(2016). *Università in declino. Un'indagine sugli atenei da Nord a Sud*. Donzelli Editore.

Viesti, G. (2018). *La laurea negata: le politiche contro l'istruzione universitaria*. Laterza & Figli.

Villari, L. (1972). *Il capitalismo italiano del'900.* Laterza Universale (n.310)

Vitale, V. (2000). L'attività della Svimez dal 1946 al 1991. *Rivista economica del Mezzogiorno*, 14(2), 541-652.

White Paper. (1970). The Reorganization of Central Government. Cmnd. 4506

White Paper. (1972). Education: a framework for expansion. Cmnd. 5174

Wildavsky, A. (1961). Political Implications of Budgetary Reform. *Public Administration Review*, *21*(4), 183–190.

Wildavsky, A. (1964). *The Politics of the Budgetary Process*. Boston: Little, Brown and Company

Wildavsky, A. (1966). The political economy of efficiency: cost-benefit analysis, systems analysis, and program budgeting. Public Administration Review, 292-310.

Williams, G. (1982). The economics of education: Current debates and prospects. *British Journal of Educational Studies*, *30*(1), 97–107. https://doi.org/10.1080/00071005.1982.9973616

Williams, P. (1992). The UK Academic Audit Unit. In M. Frazer & A. Craft (Ed.). *Quality Assessment in Higher Education: Proceedings of an international conference in Hong Kong, 1991*.

Wilson, H. (1963). *Labour and the scientific revolution. In Report of the Sixty-Second Annual Conference of the Labour Party*.

Wood, E. M. (1981). The separation of the economic and political in capitalism. New Left Review,

(127), 66.

Wood, E.M. (1995). *Democracy against capitalism: Renewing historical materialism*. Aakar Books.

Wood, E. M. (1996). Capitalism, merchants, and bourgeois revolution: Reflections on the Brenner debate and its sequel. *International Review of Social History*, *41*(2), 209-232.

Wood, E. M. (2002). *The origin of capitalism: A longer view*. Verso.

Wyn-Jones, S. (2016). *Rethinking Early Cold War United States Foreign Policy: The Road to Militarisation* (Doctoral dissertation, University of Sussex).

Ydesen, C. (Ed.).(2019). *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*. springer nature.

Zacharewicz, T., Lepori, B., Reale, E., & Jonkers, K. (2019). Performance-based research funding in EU Member States—a comparative assessment. *Science and Public Policy*, *46*(1), 105-115.

Zarka, Y. C.(2020). La valutazione: un potere spacciato per sapere/ Evaluation: A Power Disguised as Knowledge.*Cartografie sociali. Rivista di sociologia e scienze umane*, *2*(8).

# APPENDICES

## A. The 'University Constituencies' (Davies & Melchiori,1982)



Figure 3
UNIVERSITY CONSTITUENCIES
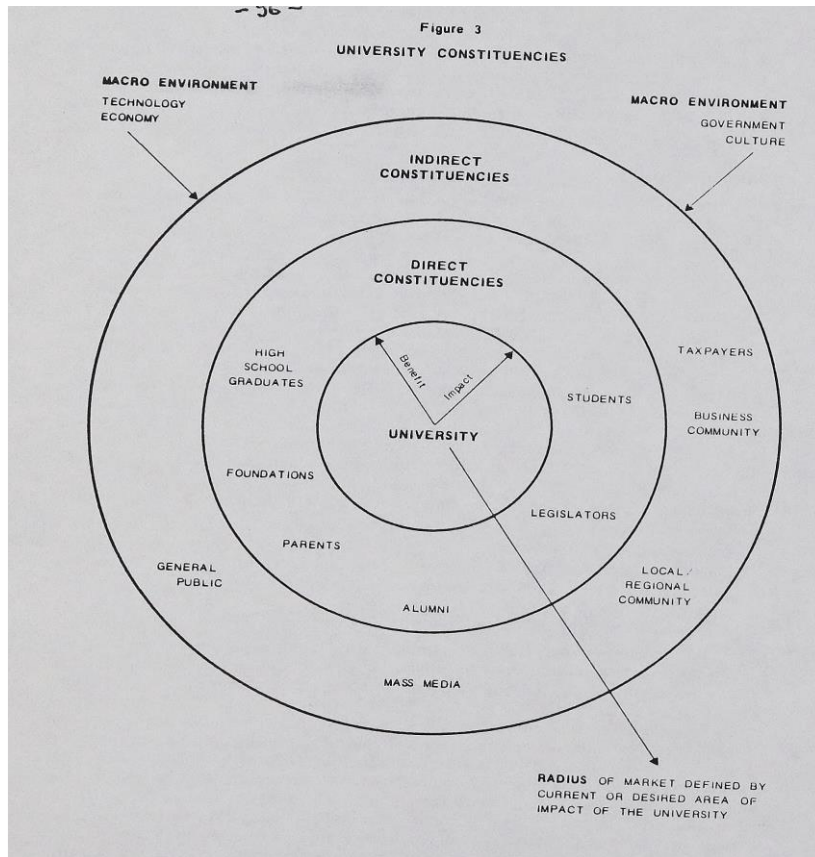
**B. Presidents, Vice-Presidents, and General Secretaries of the CRUI 1963 - 2002 (in order. Years with two names present President and General Secretaries only)**

- **1963- 1968:**
  - o Guido Ferro (University of Padova), Alessandro Faedo (University of Pisa), Giuseppe Lavaggi (University of Macerata)
- **1968 – 1972**
  - o Alessandro Faedo (University of Pisa), Tito Carnacini (University of Bologna), Mario Rolla (University of Pavia)
- **1972 – 1976**
  - o Tito Carnacini (University of Bologna), Carime Alfredo Romanzi (University of Genova), Francesco Carassa (Polytechnic of Milan)
- **1976 - 1984**
  - o Carmine Alfredo Romanzi (University of Genova), Luigi Dadda (Polytechnic of Milan)
- **1984 – 1987**
  - o Vincenzo Buonocore (University of Salerno), Gian Tommaso Scarascia Mugnozza (University of Tuscia, Viterbo)
- **1987 – 1994**
  - o Gian Tommaso Scarascia Mugnozza (University of Tuscia, Viterbo), Luigi Berlinguer (1(University of Siena)
- **1994 – 1998**
  - o Paolo Blasi (University of Florence), Luciano Modica (University of Pisa)
- **1998 – 2002**
  - o Luciano Modica[123] (University of Pisa), Fulvio Tessitore[124] (1999 and 2000)

---

[123] Member of the 7th permanent commission between 7 November 2002 and 27 April 2006
[124] Member of the 7th permanent commission between 22 June 2001 and 27 April 2006

## C. The life cycle of an educational activity (Lockwood & Davies, 1985)

| | DEVELOPMENT | INITIATION | GROWTH | CONSOLIDATION | STAGNATION | DECAY & TERMINATION |
|---|---|---|---|---|---|---|
| **OUTPUT** 0 | | | | | | |
| | | | | Optimum Return on Inputs | Diminishing Returns | |
| **CUMULATIVE 'INPUTS'** 0 | Investment for Future Viability | | | | | |
| – Time – Energy – Money | Determination of Need Specification Development Validation | Representation: Promotion of Activity – Appropriateness – Perceived Quality – Qualification – Support – Credibility – Vocational Value Etc. | Progressive Institutionalization – Resources Allocated – Procedures set up – Development of 'Norms' & Traditions – Growth of Support – Legitimacy Etc. | Adaptation to Changing Circumstances | Progressive Inertia – Inability to respond to Challenges & Criticism – Rigidity of Curriculum & Process – Loss of Clientele – Loss of Support & Legitimacy | Crisis – Enrolments fall below permissible minimum – Decision to wind-up Operation unavoidable |
| **CHARACTERISTICS** | | | | | | |

Figure 3.1
The Life Cycle of an Educational Activity.

### D. Members of the Jarratt Steering Committee for Efficiency Studies in Universities (CVCP, 1985a)

*Chairman*

1. Sir Alex Jarratt, CB, Chairman, Reed International PLC and Chancellor of Birmingham University.

*Members*

2. Mr. J. B. Butterworth, C.B.E., J.P., D.L., M.A., Vice-Chancellor, University of Warwick.
3. Sir Adrian Cadbury, M.A., Chairman, Cadbury Schweppes PLC and Chancellor of the University of Aston.
4. Professor F. H. Hinsley, O.B.E., M.A., F.B.A., Master of St. John's College, Cambridge.
5. *Sir Robin Ibbs, M.A., the Prime Minister's Adviser on Efficiency.
6. Dr. T. L. Johnston, M.A., Ph.D., Principal and Vice-Chancellor, Heriot-Watt University.
7. Dr. G. Lockwood, B.Sc., D.Phil., Registrar and Secretary, University of Sussex.
8. Mr. P. I. Marshall, FC.A., C.B.I.M., L.R.A.M., Director of Finance and Deputy Chief Executive, The Plessey Company PLC.
9. Professor P. G. Moore, T.D., Ph.D., FI.A., Principal, London Business School.
10. Professor M. H. Richmond, Sc.D., F.R,S., Vice-Chancellor, Victoria University of Manchester.
11. Professor Sir Peter Swinnerton-Dyer, Bt., M.A., F.R.S., Chairman of the University Grants Committee.
12. Mr. S. Thomson, F.C.C.A., Director of Finance and Executive Director, Ford Motor Company Ltd.
13. *Mr. I. Beesley, M.A., Head of the Efficiency Unit, with the agreement of the Committee, acted as alternate to Sir Robin Ibbs on occasions when he was unable to attend meetings.

## E. Membership of the Joint CVCP/UGC working group 'Performance Indicators Technical Committee' in 1986 (CVP/UGC, 1986)

*Chairman*

1. Prof. Sir Mark Richmond, FRS, Vice-Chancellor, University of Manchester

*Members*

*CVCP members*

2. G.R.Higginson, Vice-Chancellors, University of Southampton
3. Mr. A.D. Linfoot, Registrar and Secretary, University of Kent
4. Dr. E.S. Page, Vice-Chancellor, University of Reading
5. Prof. G.D. Sims, Vice-Chancellor, University of Sheffield

*UGC members*

6. Prof. M. Harris, Department of Romance Linguistics and Pro-Vice Chancellor, University of Salford
7. Prof. R. Layard, Centre for Labour Economics, The London School of Economics and Political Science, London
8. Prof. J. Sizer, Department of Management Studies, University of Technology, Loughborough

## F. Membership of the Joint CVCP/UGC working group 'Performance Indicators Steering Committee' in 1987 (CVCP/UGC, 1987a)

*Chairman*

1. Dr. E.S. Page, Vice-Chancellor, University of Reading

*Members*

*CVCP members*

2. Mr. J. Farrant, Planning Officer, University of Sussex
3. Mr. D. Hardie, Finance Officer, University of Birmingham
4. Prof. M.B. Harris, Vice-Chancellor, University of Essex
5. Mr.E.C. Wright, Registrar, University of Bristol

*UGC members*

6. Prof. E. Able, Professor of Inorganic Chemistry, University of Exeter
7. Mr.R.S. Johnson CBE, Director of Education, Leeds City Council
8. Prof. P.R.G. Layard, Professor of Economics, The London School of Economics and Political Science, London
9. Prof. J. Sizer, Professor of Financial Management, University of Technology, Loughborough

- 8 -

| Indicator Numbers | | Figure 1. University Management Information and Performance Indicators | Indicators shown by | |
|---|---|---|---|---|
| 1984/ 1985 A | 1985/ 1986 B | | Cost Centres | Instit- ution |
| | | **Table 1** | | |
| 1 | 1 | Expenditure per FTE student[x] | CC | |
| 2 | 2 | Expenditure per FTE academic staff | CC | |
| 3 | 3 | Expenditure on support staff per FTE academic staff | CC | |
| 4 | 4 | Expenditure on Equipment per FTE academic staff | CC | |
| 5 | 5 | Research income per FTE academic staff | CC | |
| | | **Table 2** | | |
| 6 | 6 | Research postgraduates as a % of FTE students | CC | |
| 7 | 7 | Taught postgraduates as a % of FTE students | CC | |
| 8 | 8 | All postgraduates as a % of FTE students | CC | |
| 9 | 9 | Ratio of FTE students to FTE academic staff | CC | |
| | | **Table 3** | | |
| 10 | 10 | Central admin expenditure as a % of Grand Total expenditure | | I |
| 11 | 11 | Central admin pay expenditure as a % of central administration expenditure | | I |
| 12 | 12 | Central admin expenditure per FTE student | | I |
| 13 | 13 | Central admin expenditure per FTE academic staff | | I |
| | | **Table 4** | | |
| 14 | 14 | Library expenditure as a % of total general expenditure | | I |
| 15 | 15 | Publications expenditure as a % of library expenditure | | I |
| 16 | 16 | Library pay expenditure as a % library expenditure | | I |
| 17 | 17 | Library expenditure per FTE student | | I |
| 18 | 18 | Library expenditure per FTE academic staff | | I |
| 19 | 19 | Book expenditure per FTE student | | I |
| 20 | 20 | Periodicals expenditure per FTE student | | I |
| | | **Table 5** | | |
| 21 | 21 | Computer services expenditure as a % of total general expenditure | | I |
| 22 | 22 | Computer services pay expenditure as a % of computer services expenditure | | I |
| 23 | 23 | Computer services expenditure per FTE student | | I |
| 24 | 24 | Computer services expenditure per FTE academic staff | | I |
| | | **Table 6** | | |
| 25 | 25 | Total premises expenditure as a % of total general expenditure | | I |
| 26 | 26 | Premises Pay expenditure as a % of premises expenditure | | I |
| 27 | 27 | Heat, water and electricity expenditure as a % of total general expenditure | | I |
| 28 | 28 | Cleaning and custodial services expenditure as a % of total general expenditure | | I |
| 29 | 29 | Repairs and maintenance as a % of total general expenditure | | I |
| 30 | 30 | Telephone expenditure as a % of total general expenditure | | I |
| 31 | 31 | Total premises expenditure per FTE student | | I |
| 32 | 32 | Premises pay expenditure per FTE student | | I |
| 33 | 33 | Heat, water and electricity expenditure per FTE student | | I |
| 34 | 34 | Cleaning and custodial services per FTE student | | I |
| 35 | 35 | Repairs and maintenance expenditure per FTE student | | I |
| 36 | 36 | Telephone expenditure per FTE student | | I |
| | | **Table 7** | | |
| 37 | 37 | Careers services expenditure per FTE student | | I |
| 38 | 38 | Student unions and societies expenditure per FTE student | | I |
| | | **Table 8** | | |
| 39 | 39 | Destinations of graduates after 6 months | | * |

[x] All references to FTE students relate to student load
* By academic subject group, and primary classifications

## H. List of the 37 cost-centres devised by the CVCP/UGC working group (CVCP/UGC, 1987a)

1. Clinical Medicine
2. Clinical Dentistry
3. Pre-clinical studies
4. Anatomy and physiology
5. Pharmacology
6. Pharmacy
7. Nursing
8. Other studies allied to medicine
9. Biochemistry
10. Psychology
11. Other biological sciences
12. Agricultural and forestry
13. Veterinary science
14. Chemistry
15. Physics
16. Other physical sciences
17. Mathematics
18. Computing
19. General engineering
20. Chemical engineering
21. Civil engineering
22. Electrical and electronic engineering
23. Mechanical, aero and production engineering
24. Mineral engineering
25. Metallurgy and materials
26. Architecture
27. Other technologies
28. Planning
29. Geography
30. Law
31. Other social studies
32. Business and management studies
33. Accountancy
34. Language-based studies
35. Humanities
36. Creative arts
37. Education

'A cost centre groups together operations having similar characteristics and broadly similar patterns of costs. Academic departments have been allocated to cost centres. Costs are shown in accordance with the student load and staff numbers recorded on returns to the USR. Small units servicing academic functions of particular cost centres, such as animal houses or engineering workshops, are included under the associated academic cost centres' (CVCP/UGC, 1987, p.89)

## I. Membership of the Research Sub-group of the Joint Performance Indicators Working Group (Sub-group of the JPWIG, 1993, §11)

*Secretariat*

1. Mr. David Price, HEFCE

*Observers*

2. Mr. Martin Markus, DFE
3. Dr. Angela Williams, OST

*Chairman*

4. Prof. Norman Gowar, Royal Holloway, University of London

*Members*

5. Prof. Mark Cross, University of Greenwich
6. Prof. Charles George, University of Southampton
7. Prof. Vera Gottlieb, Goldsmith's College, London
8. Prof. Chris Harris, University of Southampton
9. Prof. Hamish Keir, University of Aberdeen
10. Dr. Shekhar Nandy, HEFCE
11. Mr. Michael Powell, CVCP
12. Prof. Norma Rinsler, King's College, London
13. Mr. W.Solesbury, ESRC
14. Dr. Charles Suckling
15. Prof. Jim Taylor, University of Lancaster
16. Prof. Peter Townsend, HEFCE

*Assessors*

17. Ms. Morag Campbell, SHEFC
18. Mr. John Coote, DENI
19. Ms. Anne Hughes, HEFCW
20. Mr. Michael Sirby, HEFCE

**J.The composition of the ordinary and premial parts of the Fund for ordinary funding (FFO) in 1998 (Osservatorio, 1998)**

|  | Ordinary funding | Re-equilibrium funding (1.5% of FFO in 1995, 3.5% in 1996, 7% in 1997) |
|---|---|---|
| **Criteria** | 50% Teaching demand | 55% Teaching demand |
|  | 20 % research results | 25% Research results |
|  | 20 % teaching results | 20% Teaching results |
|  | 10% incentives |  |
|  | 100% | 100% |

## K. Members of the CRUI's 1991 evaluation committee (Allulli, 1995)

1. Lorenzo Bernardi (Padova university), Statistics
2. Alessandro Chiabrera (Genova university)
3. M.Rosa Gola (Genova university)
4. Giovanni Colucci (Siena university), Accounting Officer
5. De Bernardis and Ivana Greco (Basilicata university)
6. Giancarlo Filocano (Minister of the Treasury)
7. Gliberto Marselli (Naples University), sociology
8. Fabrzio Onida (Milan Bocconi University), Economics,
9. Dino Rizzi (University of Venice), Economics
10. Giorgio Allulli (Researcher at CENSIS)

## L. List of the 17 Macro-scientific disciplinary sectors following Ministerial Decree n.152 23/06/1997

- A   Mathematical Sciences
- B   Physical Sciences
- C   Chemical Sciences
- D   Earth Sciences
- E   Biological Sciences
- F   Medical Sciences
- G   Agrarian Sciences
- H   Civil Engineering and Architecture
- I   Industrial Engineering
- K   Computer science and engineering
- L   Ancient Sciences, Philological-literature sciences, and Historical-Artistic sciences
- M   Pedagogical, Psychological, Philosophical and Historical Sciences
- N   Juridical Sciences
- P   Economic Sciences
- Q   Political Sciences
- S   Statistical Sciences
- V   Veterinary Sciences

**M. Members of the CRUI's research quality indicators working group in 1999 (afferent discipline in brackets) (CRUI, 1999)**

*President*

1.  Prof Alberto Febbrajo (President),VC of the University of Macerata (Sociology)

*Members*

2.  Prof. Donatella Calabi, Architecture University Institute of Venice,(Architecture)
3.  Prof. Alberto Capozzi,Polytechnic of Bari, (Mathematical Analysis)
4.  Prof. Alessandro Chiabrera, University of Genova, (Electronic engineering)
5.  Prof. Mario Orefice, Turin Polytechnic, (Electronics and Electromagnetics)
6.  Prof. Massimo Pauri, University of Parma, (Physics and Earth Sciences)
7.  Prof. Gaetano Vacca, University of Basilicata, (Engineering)

# N. Staff weightings proposed by CRUI for research quality evaluation (CRUI, 1999, p.18)

| | PESI | |
|---|---|---|
| n° professori I^a e II^a fascia a tempo pieno[3]; | $n_1 \times 1$ | $N_1$ |
| n° professori I^a e II^a fascia a tempo definito[4]; | $n_2 \times 0,5$ | $N_2$ |
| n° ricercatori universitari, assistenti di ruolo e assimilati a tempo pieno, assegnisti[5]; | $n_3 \times 1$ | $N_3$ |
| n° ricercatori universitari, assistenti di ruolo e assimilati a tempo definito[6]; | $n_4 \times 0,5$ | $N_4$ |
| n° borsisti di Training and Mobility of Research della C.E. e n° borsisti post-dottorato normalizzati ai mesi in cui sono attivi, nell'anno, presso la struttura; | $n_5 \times 0,8$ | $N_5$ |
| n° dottorandi normalizzati ai mesi in cui sono attivi, nell'anno, presso la struttura; | $n_6 \times 0,7$ | $N_6$ |
| n° altri borsisti, normalizzati ai mesi in cui sono attivi, nell'anno, presso la struttura; | $n_7 \times 0,6$ | $N_7$ |
| n° tecnici di qualifica uguale o superiore al VII livello; | $n_8 \times 0,6$ | $N_8$ |
| n° altri collaboratori di ricerca con contratti di lavoro autonomo et similare, normalizzati ai mesi in cui sono attivi, nell'anno, presso la struttura; | $n_9 \times 0,5$ | $N_9$ |
| n° totale di unità di: personale amministrativo-contabile, tecnici e ausiliari di qualifica uguale o inferiore al VI livello. | $n_{10} \times 0,2$ | $N_{10}$ |
| **Totali** | $n$ | $N$ |

# O. Comparison between the CRUI's weightings for industrial engineering and economics (CRUI, 1999 pages 27-28, 31)

## Economics

| | | PESI | |
|---|---|---|---|
| Libro di ricerca originale | I | $p_1 \times 4$ | $P_1$ |
| | E | $p_2 \times 6$ | $P_2$ |
| Altri libri scientifici o di alta divulgazione | I | $p_3 \times 2$ | $P_3$ |
| | E | $p_4 \times 3$ | $P_4$ |
| Cura di libri, edizioni di testi con traduzione e introduzione | I | $p_5 \times 1,5$ | $P_5$ |
| | E | $p_6 \times 2$ | $P_6$ |
| Articoli e studi originali in riviste e volumi | I | $p_7 \times 1$ | $P_7$ |
| | E | $p_8 \times 1,5$ | $P_8$ |
| Relazioni in atti di congressi | I | $p_9 \times 0,5$ | $P_9$ |
| | E | $p_{10} \times 1$ | $P_{10}$ |
| Rassegne, recensioni critiche e interventi in atti di congresso | I | $p_{11} \times 0,2$ | $P_{11}$ |
| | E | $p_{12} \times 0,5$ | $P_{12}$ |
| Realizzazione di prodotti multimediali di interesse scientifico | | $p_{13} \times 0,5$ | $P_{13}$ |
| Pubblicazioni interne e rapporti di ricerca | | $p_{14} \times 0,2$ | $P_{14}$ |
| **Totali** | | $p$ | $P$ |

## Industrial Engineering

| | | PESI | |
|---|---|---|---|
| Riviste di livello elevato | | $p_1 \times 10$ | $P_1$ |
| Riviste di livello medio | | $p_2 \times 6$ | $P_2$ |
| Altre riviste pertinenti | | $p_3 \times 2$ | $P_3$ |
| Libri scientifici | | $p_4 \times 20$ | $P_4$ |
| Libri ad alta divulgazione | | $p_5 \times 10$ | $P_5$ |
| Libri didattici | | $p_6 \times 6$ | $P_6$ |
| Capitoli di libri scientifici e di alta didattica, monografie enciclopediche | | $p_7 \times 6$ | $P_7$ |
| Curatele di libri collettanei, atti convegni | | $p_8 \times 2$ | $P_8$ |
| Curatele di collane | | $p_9 \times 4$ | $P_9$ |
| Recensioni, prefazioni, voci enciclopediche, cataloghi, carte geologiche ufficiali, editorial board di riviste di livello elevato / medio | | $p_{10} \times 1$ | $P_{10}$ |
| Atti di convegni internazionali con revisori di livello elevato articolo su invito / articolo in extenso | | $p_{11} \times 5$ | $P_{11}$ |
| Atti di convegni internazionali con revisori di livello elevato articolo breve / poster | | $p_{12} \times 2$ | $P_{12}$ |
| Atti di convegni internazionali con revisori di livello medio articolo su invito / articolo in extenso | | $p_{13} \times 2$ | $P_{13}$ |
| Atti di convegni internazionali con revisori di livello medio articolo breve / poster | | $p_{14} \times 1$ | $P_{14}$ |
| Atti di convegni nazionali con revisori articolo su invito | | $p_{15} \times 3$ | $P_{15}$ |
| Atti di convegni nazionali con revisori articolo in extenso | | $p_{16} \times 2$ | $P_{16}$ |
| Atti di convegni nazionali con revisori articolo breve / poster | | $p_{17} \times 1$ | $P_{17}$ |
| n° di brevetti di cui è coautore almeno un dipendente della struttura | I | $p_{18} \times 3$ | $P_{18}$ |
| | E | $p_{19} \times 7$ | $P_{19}$ |
| Rapporti finali di ricerca pubblicati attraverso un ente e disponibili al pubblico | | $p_{20} \times 2$ | $P_{20}$ |
| **Totali** | | $p$ | $P$ |

**P. List of the 14 'Scientific disciplinary Areas' following Ministerial Decree 4/10/2000**

- 01  Mathematical and Computer Sciences
- 02  Physical Sciences
- 03  Chemical Sciences
- 04  Earth Sciences
- 05  Biological Sciences
- 06  Medical Sciences
- 07  Agrarian and Veterinary Sciences
- 08  Civil Engineering and Architecture
- 09  Industrial and information Engineering
- 10  Ancient Sciences, Philological-literature sciences, and Historical-Artistic sciences
- 11  Pedagogical, Psychological, Philosophical and Historical Sciences
- 12  Juridical Sciences
- 13  Economic and Statistical Sciences
- 14  Political and Social Sciences

**Q. Ranking Scale and associated weightings for evaluating research quality in the 2001-03 VTR** (author's tables from CIVR, 2004)

| | |
|---|---|
| **A** | At least 50% of research output is judged as excellent, the rest as good |
| **B** | At least 30% of the research outputs is excellent, the rest as good |
| **C** | At least 50% of the research output is judged as excellent or good, the rest as acceptable |
| **D** | At least 30% of the research outputs is judged excellent or good and less than 50% has limited value |
| **E** | At least 20% of research output is judged as excellent or good and less than 50% has limited value |
| **F** | Less than 20% of the research output is judged as excellent or good and less than 50% has limited value |
| **N/A** | Over 50% of the research products are of limited value |

| Judgment | Associated weight |
|---|---|
| Excellent | 1 |
| Good | 0.8 |
| Acceptable | 0.6 |
| Limited | 0.2 |