Gli effetti della valutazione della ricerca sulla research agenda dei ricercatori

Università degli Studi di Roma La Sapienza

Dipartimento di Scienze Sociali ed Economiche
Dottorato in Scienze Sociali Applicate

Valentina Carazzolo
Matricola 1566317

Relatore
Emanuela Reale

35^ Ciclo di Dottorato

ANNO ACCADEMICO 2022/2023

Table of contents

Introduction

In many Western countries, university has been experiencing a long-lasting period of transformation, due to a series of both internal and external factors. Indeed, since the '60s, the massification of the university access, the limited economic resources linked to the economic crisis during the 70s, let policymakers reflecting on what role should the universities play in this evolving socio-economic context (Paradeise et al, 2009). Governments' replies to this question have been several, and, among then, it is worth citing the performance-based allocation system and, generally, decreased available funds for research, the introduction of the concept of the third mission, as a means to let universities cooperate with private enterprises and local institutions, and, lastly, research evaluation. As a matter of fact, our research is mainly focussed on the latter element of the list, with specific reference to the Italian academic system. The Italian academic context, as a matter of fact, is peculiar for two main reasons: first, Italy has introduced a standardized national research evaluation system relatively later (namely, the VQR), in comparison to the other major European economies, where the process had already started in the '80s; second, the implementation of the VQR came together with the ASN, notably a major reform that regulated a new researchers' hiring mechanism.

The focus of our research is to understand whether the introduction of research evaluation has changed the way in which researchers working in the Italian universities conduct research, more specifically, the way in which researchers structure their research agenda setting. Hence, after a discussion on the historical and sociological basis for the introduction of research evaluation at both international and national level (first and second chapter), for the scope of our research we propose an operational definition of research agenda setting, and we define it as the short and medium-term actions with which the researcher selects the topics to research and the framework to address them (third chapter). The elements that compose the research agenda are the selection of topics to research, the motivation for researching, the selection of research practices, and the selection of the reference peers community.

In our opinion, the study on the eventual effects of the implementation of research evaluation of the way in which researchers select the topics to research is important for two main reasons. First, because of the importance of investigating how research evaluation has been affecting the overall "epistemological curve" of a scientific community. More concretely, it would be helpful to understand whether, for instance, researchers are selecting more mainstream topics to conduct research on, or, perhaps, their motivation and focus as scholars is evolving, ecc.

A second reason why understanding the effects of research evaluation on researchers' research agenda would be important is that it would be helpful to make a point on how the academic community is reacting, after 13 years (more precisely, with the Act of 2011 no. 111), to the implementation of VRQ.

The units of our analysis are the individual researchers, more specifically, social scientists (economists, statisticians, sociologists, political analysts), currently working in Italian universities (as full professors, associate professors, fixed-term researchers, full-term researchers). We opted for addressing the entire academic population, with the purpose of collecting the more diverse opinions possible on the issues at stake.

The research consisted of two phases, and a concurrent mixed-method approach (Creswell et al, 2003) was adopted. More specifically, a concurrent triangulation design was deemed appropriate in order to provide a thorough overview of the effects of the introduction of research evaluation and how it has been perceived by the Italian academic community, by overcoming the eventual weaknesses loss of information deriving from the use of only one type of method or data (quantitative or qualitative).

As for the first research phase, we submitted to the target population (more than 6000 people) a self-administered webmail survey, with the scope answering to our three research questions (namely, whether the introduction of VQR and ASN has somehow influenced researchers' research agenda setting, whether VQR and ASN have had a differentiated effect on researchers' research agenda setting, and, at last, whether some differences can be identified among scholars from different disciplines). Hence, the survey included a series of question on the eventual effects of VQR and ASN on researchers' research agenda setting, and a series of questions on the effects of other elements that,

according to the available literature, generally play a role in shaping researchers' research agenda setting (notably, scientific collaboration, getting peers' recognition, mainstream topics, the role of current event, research as benefit for society, and, at last, the effects of attendance to national or international calls, better illustrated in chapter 3 and 4).

For the second phase, we conducted informed interviews (Laudel, Gläser, 2007) with some experts on the Italian research evaluation system, who had previously participated to the survey, deepening the most relevant topics emerged from the survey free comments. As a matter of fact, with the choice of interviewing researchers who previously played an active role in the evaluation process, this qualitative phase was meant to propose a twofold point of view, namely the evaluator and the evaluated researcher's point of view. The interviews, that were conducted online, allowed us to have a quite thorough picture of researchers' opinions on the general functioning and effects of the Italian research evaluation system, from the consequences on the overall quality of knowledge production, to the effects on academics' recruitment (chapter 5).

The empirical phase of the study took about twenty months (from March 2021 to November 2022, from the drafting of the survey, to the content analysis of the interviews conducted with the experts.

The present work is part of the PRIN Project *"The effects of evaluation on academic research: knowledge production and methodological issues"*, coordinated by IRCRES-CNR.

Chapter 1
An overview on university and research evaluation on an international perspective

In many Western countries, the university has been increasingly called to play a more active role to help national governments to effectively face the upcoming societal challenges. Some new ways of conceiving the relation between universities, national governments and society at large emerged, thus subsequently paving the way to future changes in this relationship. The purpose of the present chapter is, therefore, to first describe the main different conceptions of the new university' role, that encouraged a series of reforms within the academic system of many countries, in the perspective of creating a more competitive higher education system. Second, we will define the concept of research evaluation, which is one of the most interesting elements of novelty that stemmed from this new conception of research and university, and on which our research is focussed on. In the third, and last, part of the present chapter we will present and discuss some of them major considerations from the academic community on the principles and functioning of research evaluation.

1.1 A changing higher education system

After the second World War, European higher education system (HES) has been experiencing a change in the general conception of research and academic activities, and, more precisely, a redefinition of the relationship between the academia and the nation state (Paradeise et al, 2009), which progressively led to a reorganization of its universities' administrative, research-related, and educational settings. As a matter of fact, the entire higher education system has been the subject of a remarkable rethinking of its role and relationship with its external stakeholders and society at large, in the wake of a more general reorganization and redistribution of power and functions among central government and the peripheral public facilities. This was especially the case of those countries whose academic traditions stemmed from the Humboldtian model and Napoleonic reforms (Shattock, 2014).

The academic environment was formerly conceived as an autonomous and self-regulated body, and researchers' community was organized in collegial bodies which established their own practices, norms, and values (Binswanger, 2014). Still, academic activity was rather idealized by society, and the state was solely expected to grant academia the "integrity" and autonomy (Merton, 1942), and the financial resources to let researchers to conduct their own activities.

Nevertheless, two major societal changes induced a shift in the perception of academics activities and their contribution to society at large: firstly, the massification of university access (encouraged by the massive students' protests during the late 60s) determined a shift of the conception of university access as a prerogative of the intellectual elite, and a remarkable increasing need to invest economic resources in higher education (HE); secondly, the economic crisis and the increasing unemployment rate urged, on the one hand, governments to rationalize the public sector investments, and, on the other hand, encouraged a debate on the role academics should play within society, by, for instance, finding and proposing solutions to the challenges governments were about to encounter.

Therefore, two new conceptions of the relationship between HE, nation state and society emerged (Paradeise et al, 2009), alongside the more traditional perspective.

According to the first conception, the state should have played a more incisive role in the management of university activities and mediating among the different interests of the academic community and external actors. The nation state was thus supposed to directly control and monitor the universities' activities, to better allocate and use the resources invested in a rapidly enlarging HE system. Indeed, there was the wide spreading idea that publicly invested universities were not efficient enough to foster social cohesion and economic growth (Enders et al, 2011), and that society should have had back from the academia a payoff for the huge amount of investment devolved in universities activities (Whitley et al, 2010, Martin 2011).

According to the second conception, HE activities were perceived more as commodities, and, as such, nation state would have to foster the quality of such commodity by introducing market-oriented policies. In this sense, the nation state should have shifted from a controlling and dirigiste HE

management, toward a more supervising and *steering-at-a-distance* approach. Consequently, the HE environment appeared no longer as an *ivory tower* far and detached from society, and, conversely, more often involved to address the needs of the several societal actors.

According to Paradeise et al. (2009), this changed conception of the role of academia and its relationship with society at large has led to the redefinition of the relationship between nation state and HE in three main possible models (both jurisdictional and theoretical), each observable with different degree of intensity in all European countries.

With the first model[1], the nation state started increasing its managerial control over HE facilities, and restructure universities' management with the explicit purpose of making HE sector more agile, efficient, and valuable. The control over academic activities was primarily exerted through a strengthened and vertical hierarchy, which had the mandate to stimulate the achievement of specific targets and increase the researchers' productivity. Alongside the strengthening of the university elites, new evaluation agencies and supervising bodies were instituted, and were in charge of framing the objectives to be pursued and controlling the universities' activities (especially with the ex-post evaluation). The rationale (or "narrative") underlying this redefinition of the nation state-university relation is the so-called *New public management*. With the well-known concept of New Public Management[2] we identify the series of reforms, inspired to the neoliberal doctrine, which marks the shift from central "traditional state control and academic collegial governance" toward an institutional managerial approach (Enders et al, 2011). The focus of the "evaluative state" (Neave, 1988) was thus the performance, rather than the process. The purpose was to shift from the traditional public administration, that was generally deemed inefficient, unproductive, and unfit for facing the major societal changes that countries were about to encounter, toward an agile, flexible, market and outcome-oriented public facility bureaucracy, with great focus on the efficiency and effectiveness of

[1] As we will refer to as a *stronger state management of the public sector*, Paradeise et al, 2009.

[2] The New Public Management concept was first introduced in 1991 by Christopher Hood (1992), and it refers to the wave of neoliberal-inspired reforms introduced in UK by the Thatcher government during the 80s.

the services provided. As previously mentioned, one of the consequences of the introduction of New Public Management was the implementation of accountability mechanisms and evaluation agencies, which could exert a strategic control over universities' activities and performances (according to a *steer-at-a-distance* strategy). The devolution of managerial tasks to the academic elites made them responsible for the achievement of the targets previously established by policymakers.

The "indicators" of the implementation of the New Public Management should be (Paradeise et al, 2009): 1) market-based reforms, 2) harder budget constraints, 3) focus on performance, 4) concentration of financial resources towards highly performative facilities, 5) vertical university management by setting explicit performance targets, 6) strengthened university hierarchy and governance, 7) introduction of managerial roles in university, and 8) growth of performance related pay.

Hence, some pioneer countries (such as UK, France, Spain, Netherlands, Finland, and Sweden) progressively moved towards the stronger state management of the academic sector by establishing evaluation bodies (like the *University Grants Committee* in UK, the *Comité National d'Évaluation* in France, the Spanish *Consejo de Universidad),* with which the central government could easily exert strategic control over universities' outputs. The rationale of the introduction of these first evaluation measures was then twofold (Neave, 1988): on the one hand, evaluation was meant to ensure control over the daily functioning of the university system; on the other hand, evaluation was "strategic" (Neave, 1988) for the achievement of the targets and objectives established by the legislators.

The second and the third conception of the redefinition of the state-HE relation are defined, respectively, the *Hollowing-out Thesis* and the *Democratic Revitalisation* by Paradeise et al (2009). In both these two perspectives, the academia is conceived more as a partner with which to cooperate and collaborate (with slight and indirect form of supervision), rather than a separate body whose activities need to be constantly controlled.

In the *Hollowing-Out Thesis* perspective, the nation state is losing control over academic activities, while international bodies (such as the European Union and the OECD) tend to gain momentum in

funding and framing research activities. Consequently, universities should include and deal with local, national and international actors in their research activities. More specifically, this process has particularly gained momentum with the signature of the Sorbonne Declaration in 1998 by 4 countries (France, Germany, Italy and UK), and, in 1999, the Bologna Declaration, to which adhered more 25 countries. The purpose of these two initiatives was, on the one hand, to encourage the modernization and the competitiveness of European higher education systems "by a powerful and more integrated knowledge-based economy" (Enders et al, 2011), and, on the other hand, to harmonize the European national HES, by at the same time keeping the respective local differences and specificities in research and teaching. The European Union thus became an important player in the framing the core measures that were supposed to help improving higher education systems and address the weaknesses. Indeed, the signature of the Bologna Declaration has deeply encouraged national countries to resettle their universities systems (Hicks, 2012) and opened the pathway to the constitution of a European Higher Education Area[3]. The institution of the European Higher Education Area (successively linked to the European Research Area) and the foundation of agencies like the European Network for Quality Assurance in Higher Education (ENQA), as an extension of the Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG), displays how international institutions have been playing an increasing role in framing the management of university activities. The influence of external institutions on universities' activities has assumed both a direct form, like for example with the development of the Framework Programmes by the European Commission, and a more indirect forms (such as, with the launch of the Lisbon Agenda[4], and the increasing role played by the OECD, which tended to warmly suggest universities a series of good practices).

[3] 49 countries have so far adhered to the European Higher Education Area (http://www.ehea.info/). The European Higher Education Area recognises as its founding principles the higher education systems commitment to "freedom of expression, autonomy for institutions, independent student unions, academic freedom, free movement of students and staff" and "increased staff and students' mobility and to facilitate employability".

[4] This objective was, specifically, a core target of the 2000 Lisbon Strategy, which stated the European Union commitment to make Europe the "the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion", https://www.europarl.europa.eu/summits/lis1_en.htm

According to the third form of redefinition of academia-nation state relationship, notably the *democratic revitalisation*, the academic governance was supposed to include both, on the one hand, the wide range of academic actors (such as professors, administrative staff and the students), and, on the other hand, the external stakeholders in the decision-making process underlying teaching and research activities. It is indeed in this specific context that the concept of the "Mode 2 knowledge production" has emerged (Limoges et al, 1994). According to the authors who first introduced the concept, research was already undergoing some changes, due to the increased relationships and exchanges among universities, industries, and the central state, and it was just a matter of time before research would have become something to be totally shared among the different actors at stake. The "Mode 2" knowledge production was thus conceived as a new practical approach to research, in which the scientific outputs were the results of shared research practices among the several research actors and external stakeholders (companies, national state), as opposed to the allegedly previous way of conducting research. The authors, still, assumed that a gradual transition from the traditional knowledge production (the "Mode 1") to the "Mode 2" knowledge production was then meant to gain momentum in the next future. The authors identify 5 features of research production that mainly summarize the differences between Mode 1 to Mode 2: in Mode 2, research is conducted in the *context of application* (namely, research is intended to be useful to society, industry or government); research is more *transdisciplinary* and *diverse in its organization* (the approach to conduct research tends to integrate more the role of different disciplines, with heterogeneous practices); research is, still, more *reflexive*, incorporating multiple views, and has new *quality control instruments*, which add up to the traditional peer reviewing process[5] (Hessels, Van Lente, 2008, Limoges et al, 1994).

[5] However, it is worth underling that, if on one hand the Mode 2 theorists claimed that the New Knowledge Production was meant to gradually spread to all disciplines, on the other hand some other authors maintained that it was not possible to identify a clearcut point which marked the shift from Mode 1 to Mode 2 (Knuuttila, 2012) that could prove the actual transition of research practices, especially for the wide diversity of practices and features of all the disciplines to be considered in the analysis (Hessels, Van Lente, 2008).

In this specific context, it is also worth mentioning the emergence of the universities' Third Mission, which identifies the re-orientation of universities' activities to foster and develop the social and economic context they are embedded in (notably, knowledge transfer to local industries, entrepreneurship education, engagement of stakeholders) for the development of a knowledge-based economy. The Third Mission is indeed a concept that appropriately represents the concrete collaboration of universities with local facilities, by including the good practices suggested by national and international institutions.

From the *hollowing-out* and the *democratic revitalisation* thesis, the *network governance narrative* has stemmed from. In this perspective, the academia is now perceived as just one of the several actors at stake. The participation, collaboration, and the encouragement of pursuing some targets shared by the multi-level actors, are now leading the knowledge production process, and the university governance at large. Even though there are no countries which have exhaustively embraced this narrative, in more contexts there are signs of an increasing presence of horizontal research management, rather than vertical (Paradeise et al, 2009). The most illustrative features of the network governance narrative are: 1) the development of networks between HE institutions and other social actors, which play a significant role in the university governance, 2) the self-steering and self-organizing capacity of these networks, 3) these networks present a joint problem solving and problem recognition approach, 4) external control is more indirect and less influential, with the Ministry of HE slightly intervening in universities' activities, 5) emphasis on soft management skills and networking capabilities, 6) human resources tend to reward team work rather than individuals.

1.2 Research evaluation policies, an overview

Research evaluation can be defined as a set of procedures aimed at evaluating the quality of research performances of individuals, research units (most frequently afferent to public-funded research institutes), or of research projects, conducted on a certain regular basis, by using bibliometric

measures and peer review, and carried on by appointed panels of experts. Since the bibliometric indicators are not applicable to all disciplines, peer review has become the most widely used evaluation method, and, when the panels evaluation activities are supported by all the available quantitative information on the research product to evaluate (such as publication and citation data), it is called *informed peer review* (Geuna, Martin, 2003).

One of the basic criteria with which distinguishing the different kinds of research evaluation is the moment in which the evaluation process is conducted: on one hand, with 'a priori' evaluation, we identify the whole process of the evaluation of a research project occurring before its eventual beginning, comprising also the overall monitoring activities conducted once the research activities have started; on the other hand, we identify a 'a posteriori' evaluation of the research product, all the research evaluation practices conducted on a finalized research product[6].

The evaluation of research products is generally performed by evaluating agencies, whose primary task is to evaluate the achievement of some previously established quality standards. Even though the evaluating agencies of the several national evaluation systems differ from each other for the tasks they are supposed to perform and the evaluation criteria they have to rely on (and that interestingly reflects the remarkable diversity among the several research evaluation systems[7]), it is possible to identify the four most commonly used output measures that are considered in the evaluation process (Geuna, Martin, 2003): volume, quality, impact and utility.

Since the first Research Assessment Exercise, which took place in the UK in 1986, the concept of research evaluation quickly spread to many other countries, till becoming a common practice in higher education policies. The purpose of the first evaluation exercise conducted in UK, and of the subsequent ones in 1989, 1992, 1996 and 2001, was to rank UK universities according to a logic based on academic quality, upon which to base research institutes' score, and subsequently

[6] More specifically, even though 'assessment' and 'evaluation' are mostly used interchangeably, their meaning is slightly different. Indeed, 'assessment' generally refers to an 'a priori' evaluation, whilst 'evaluation' indicates an 'a posteriori' evaluation (Hills, Dale, 1995).

[7] Geuna, Martin 2003, Ochsner et al, 2020.

determining the amount of public resources to invest on each research institutes (the funding institute, till 2018, was Higher Education Funding Council, which was created from the merge of the Universities Funding Council and the Polytechnics and Colleges Funding Council). The introduction of this funding allocation mechanism has marked the shift from a state support to research activities based on block grants to a mixed approach, which includes the allocation of resources based on universities' performances. The performance-based research funding system has become prevalent in the research finance policy of many countries, to the point that 18 of the EU28 countries in 2019 have adopted some forms[8] of performance-based funding system (Zacharewicz et al. 2019).

Alongside the performance-based funding system, the allocation of resources based on the approval of the financing of specific research projects has also contributed to the change of the management of public investment in research, and became preeminent, to the point to cover almost a quarter of the total amount of the public resources (Lepori et al, 2007).

Therefore, the combination of the introduction of research evaluation and the reshaped funding mechanism can be considered the core features of the changed relationship between the state and research institutes in the *new public management* sense: the state would thus be able and entitled to address research activities like a common buyer can choose a service or a service provider (Geuna, Martin, 2003, Martin, 2011), according to a *principal-agent* logic (Olssen, Peters, 2005).

The classification of the national research systems in different and well-defined clusters is a very difficult task, for three reasons (Ochsner et al, 2018, Ochsner et al, 2020).

First and foremost, the definition itself of the various procedures of which a whole research evaluation system is composed of, vary considerably among countries. The high level of variation of the definition of the research evaluation procedures is due to several reasons: the wide variety of national

[8] According to Hicks (2012), a performance-based research funding system should meet the following criteria: 1) research must be evaluated; 2) the evaluation must be ex post; 3) the output of research must be evaluated; 4) the distribution of funding from government must depend on the evaluation results; 5) the system must be national.

scholar legacy, different national languages, the diversified evolution over time of these policies with their non-linear path, the political institutions of each specific country from which these policies stem from, and, not least, the experts' conception and definition of these policies, who may define differently the various research evaluation components (Ochsner et al, 2018, Ochsner et al, 2020).

A second factor which deepens the complexity of the task is that in some cases there is not an explicit rationale of the research evaluation policies introduction, thus in many cases determining a policy implementation that has led to the juxtaposition of different and contrasting procedures with the purpose of mutually counterbalancing the eventual negative effects if the policies' introduction (Ochsner et al, 2020).

A third factor is linked to the changed academic governance and management, that enabled a redistribution of power and tasks among the players (as we will better see later in Whitley's work), that may have occurred differently in the various countries. Hence, the plurality of the policies is to be added to the plurality of the relations existing among the different actors at issue (Hicks, 2012, Ochsner et al, 2018).

Lastly, the complexity of defining national RES is additionally due to fact that in some countries research evaluation output is linked to the funding allocation system (both performance-based research funds and the funding of research projects)[9], and there is a relevant degree of variability of this aspect among different countries.

As we are going to see, the classifications available in literature are generally based on the combination of at least two of the following criteria (Ochsner et al, 2020):

1) the ex-ante evaluation vs the ex-post evaluation,
2) link between research evaluation outcome and funding (summative vs formative evaluation), and
3) the evaluation methods (metric vs peer review)[10].

[9] Hicks, 2012, Lepori, Reale & Spinello, 2018.
[10] (Ochsner et al, 2020).

A first contribution to the classification of national research systems was provided by the research conducted by Coryn, Hattie, Scriven and Hartmann (Coryn et al, 2007).

According to the authors, the research evaluation models of the countries analysed[11] can be clustered according to two main variables, namely, the allocation of resources for research[12], and countries general approach on research evaluation[13]. Interestingly, the study shows that few countries at that time presented a systematic and consistent approach (Type A), except for Australia, Hong Kong, the Netherlands, New Zealand, the UK and the US. Indeed, almost all countries show high level of internal inconsistency within their same research evaluation system, while the block grant seems to be the more prevalent funding mechanism. However, the authors suggested not to consider their work as definitive[14], for the complexity linked to the classification of national research evaluation systems.

Another worth analysing contribution to the categorization of the different types of national research systems has been proposed by Whitley (2010). More specifically, the author focussed on the eventual university governance arrangements, and how various public science systems have enabled different national research systems (Whitley, Glaser, 2007)[15].

According to the author, because of the 6 major changes that the higher education systems had to deal with[16], it is possible to identify 6 ideal types of relations existing among the three main higher

[11] USA, Japan, Hong Kong, Australia, Belgium, Czech Republic, Finland, France, Germany, Hungary, Ireland, the Netherlands, New Zealand, Poland, Sweden and the UK.

[12] 1) Type I, L*arge-scale performance/judgement-based exercises*; 2) Type II, *Bulk funding models*; 3) Type III, *Indicator-driven model*

[13] Type A, standing for "an approach which is systematic and consistent", and Type B, "pluralized approaches that can be characterized by a high degree of situation-specific variability in terms of their conceptions, methods and applications"), Coryn et al, 2007.

[14] Interestingly, the authors remarked the relative discontent showed by the experts surveyed on the efficacy of the research methods.

[15] As we will see in the next paragraph, Whitley's contribution is also focussed on the effects of the combination of different public science systems and national research evaluation systems on the knowledge production process.

[16] The author identifies the 6 major changes of higher education system as follows: 1) remarkable increase of human and financial resources invested in higher education in the second post-war period followed by a huge reduction of resources invested; 2) the shit from the block grant funding system to the project-based funding system; 3) the steering activity by the government over research and teaching activities; 4) delegation of monitoring universities activities to managerial figures; 5) the process of evaluating the quality of research, and 6) the increase of relationship between university and private enterprises.

education players, namely the university, the state agencies, and the employing organizations. The author thus identifies the 3 ideal types of authorities relation archetypes (that is, *state-dominated*, *state-delegated*, and *employer-dominated*), from which he has derived 6 possible public science systems: 1) the *state-centred* public science system, in which the state can directly exerting a relatively strong control over universities' activities, by addressing academics' research, and establishing the rationale with which assigning the financial resources and rewards to research institutes; 2) the *state-shared* public science system, in which the academic elites can be relatively independent from the state, and to organize their short and medium term activities by themselves; 3) in *state-delegated competitive* and in 4) *state-delegated discretionary* public science systems, researchers are employees of the universities (which is still largely funded by the central state), and have to compete with the colleagues in order to get the financial resources to conduct research (in the former), or are actually more independent in conducting research activities; 5) lastly, in *employer-competitive* and in *employer-centred* public science systems, researchers' employers play a significant role in addressing researchers' activities and in funding research.

Therefore, it is starting from the reflection on the possible relations among the higher education actors and the related governance, that Whitley has proposed a categorization of the national research evaluation systems.

According to Whitley (2007), the distinguishing criteria which may be used to classify the several national research evaluation systems are: the *frequency* with which the research evaluation exercises are conducted; the *formalisation* (referring to whether research evaluation is organized according to formally or informally specified conditions); the *standardisation*, notably "the use of common evaluation procedure and practices across the sciences and over time", and, at last, *transparency*, referring to the public availability of the various procedures and measures of research evaluation (evaluation outcomes, how panels experts are appointed, and so on).

From the different combination of these parameters, the author has further proposed a classification of the research evaluation systems in *strong* or *weak*. On the one hand, in strong research evaluation

systems, evaluation procedures are conducted on a regular basis, and conducted according to highly formalised published criteria. The evaluation outcomes (which are also published) often determine a considerable number of resources to be allocated. On the other hand, in weak research evaluations systems, evaluations exercises are conducted rarely or on an irregular basis, and may also be organized informally by the universities themselves. Both the evaluation procedure and outcomes are rarely published, and these results are rarely ranked according to an international scale of excellence. It is also worth mentioning that Whitley's contribution to the understanding and the classification of national research evaluation system includes some considerations on the research funding allocation regime as variable, and how relevant it is in the knowledge production process (Whitley, 2007, Whitley et al, 2018). Indeed, according to the author, it is possible to classify various national research evaluation systems (especially the ones considered strong) by considering the different national funding regimes. With the introduction of research evaluation policies and the need to rationalize resources to invest in research, the allocation of funding on a performance base has become an important leverage for governments to incentivize academic excellence (Martin, 2011, Hicks, 2011) and framing research goals (Whitley et al, 2018). On the one hand, in the case of strong research evaluation systems, nation states tend to allocate funding considering the evaluation outcomes; on the other hand, since weak research evaluation systems' primary goal is generally to foster the research institute performances, by focussing on the possible weaknesses and encouraging future organisational improvements, they would tend not to use the resources allocation as a leverage on universities and researchers' activities.

Ochsner et al. (2018) proposed a further set of dimensions aimed to categorize national evaluation systems (*institutional evaluation*, *project funding* and *career promotion*), with which they have created five ideal clusters:

1) No national database/non SSH-specific:
2) Non-metric/SSH-specific;
3) Performance-based funding/non-metric;

4) Performance-based funding/metric, and

5) Metric/push for English (Ochsner et al., 2020).

From this study[17], actual national research evaluation systems represent traits from almost all the five ideal types, but the authors chose to cluster each country in the ideal type that appeared to be the closest to the actual policy setting (see Table 1 below).

In the first group, which is characterized by "national research that does not have a national publication database, are not based on metrics, are not linked to funding and do not have SSH-specific procedures", Iceland seems to be the most representative country (while the others tend to deviate for one or two variables). As for the second group features ("in which research evaluation systems do not have national databases, don't use metric as primary evaluation method, do not incentivise publication in English and have dedicated funding programs for SSH research"), Switzerland is the country that best embed those characteristics. Moreover, the country that best represent the third group features (*Performance-based funding, non-metric*) is Norway, while Denmark is the best representative of the fourth group features. At last, Estonia is the best representative of the fifth group features.

[17] The study, conducted by Ochsner, Kulczycki and Gedutis (2018), is part of a series of works commissioned by the *European Network for Research Evaluation in the Social Sciences and Humanities* (ENRESSH, https://enressh.eu/about-2/), with the scope of studying the research evaluation policies and fostering the best practices. The ENRESSH main activities are articulated on the following topics: *Conceptual frameworks for SSH research evaluation* (Working Group 1), *Societal impact and relevance of the SSH research* (Working Group 2), *Databases and uses of data for understanding SSH research* (Working Group 3).

Ideal Type	*Countries closest to the ideal type*	*Countries difficult to classify, closes type chosen*
No national database, Non SSH-specific	Cyprus, France, Iceland, Macedonia, Malta, Montenegro, Portugal, Spain	Bulgaria, Italy
Non-metric, SSH-specific	Austria, Germany, Ireland, the Netherlands, Serbia, Switzerland	
Performance-based funding, non-metric	Lithuania, Norway, South Africa	Denmark, Israel
Performance-based funding, metric	Czech Republic, Croatia, Poland, Denmark	Finland
Metric, push for English	Bosnia-Herzegovina, Estonia, Hungary, Slovakia, Slovenia, Romania	Latvia

Table 1: Five ideal types of national evaluation systems and classification of countries, from Ochsner et al, 2018

Interestingly, some countries (right column of Table 1) seem to be very difficult to be assigned to a specific ideal type.

According to the authors, one of the most interesting aspects that emerged from the study is that the variability among national research evaluation systems is huge, since the research evaluation policies frameworks are conceived to address some determined policy issues, which, as a matter of fact, are specific to each country. Nonetheless, it is also worth remarking that there seem to be similarities among neighbour countries, thus showing how the various research evaluation systems may reflect the political and institutional legacy of the different countries. Indeed, Southern-European countries, German-speaking countries and Nordic countries can be roughly clustered together (Ochsner et al, 2020). Interestingly, as final remark, highly performative countries (such as Switzerland, Germany and the Netherlands) seem to have the most adaptive and less focussed on metric research evaluation approach, while, on the other hand, some other countries tend to use rankings and push for English to improve academic excellence.

1.3 Research evaluation: critics and some considerations on its effects

Since the introduction of research evaluation practices, the academic community has widely discussed the effects of these new policies. Particularly, the introduction of institutional research evaluation,

alongside the pre-existing peers evaluation, has let some scholar wondering on the goodness of the matching of the two "kinds" of research evaluation.

As a matter of fact, the referee system has marked the transition to the modern science system, in which the shift from the secrecy of scientific works (Zuckerman, Merton, 1971) towards the publishing of the manuscripts on (the few) accredited scientific journals (Zuckerman, Merton, 1971) occurred. In the referee system, the editors and the referees assume the role of "significant status-judges" (Zuckerman, Merton, 1971), since they oversee the assessment and the publication of the scientific works, which certifies the scientific value of the researcher's work. Hence, the real power of editors and referees mostly lies on the citation mechanism, by which the priority of a discovery and the ownership of determined ideas is attributed to a researcher (*priority rule*, for which scientists compete with each other), who wins the symbolic reward for his discovery and the scientific recognition of his peers (*winner takes all*).[18] Moreover, the recognition gained through the discovery will grant the researchers a positive spill over throughout his academic career, and we will presumably receive more credit than other colleagues (*Matthey effect*). In this specific context, the academia represented a quite self-regulated body, where the government hardly interfered in daily activities, and researcher's teaching and researching activities were unknown to most (Binswanger, 2014).

It is thus the change and re-balancing of the roles of academic actors and the newly external actors in the specific context of research evaluation, alongside the concrete consequences of these policies on universities and research funding organizations, that have encouraged researchers to reflect on the effects of institutional research assessment. Indeed, in these almost thirty years from the introduction of the first examples of research evaluation policies, many scholars have been studying how these reforms has affected the several aspects of academic activities.

[18] Small, 2004

Because of the awareness of the importance of addressing the biases eventually determined by the introduction of national research systems, there have been some examples of researchers' gatherings providing some guidance for the improvement of the current research assessment practices.

The first example is the Declaration on Research Assessment[19] (also called the DORA Declaration), developed in 2012 during the Annual Meeting of the American Society for Cell Biology in San Francisco. The DORA Declaration, which recorded 12,000 individual signers and almost 500 institutions, aimed at involving the several university and research actors (researchers, financing agencies, editors, and evaluation agencies), encouraging a worldwide debate, with the purpose of promoting a discussion on the issues and criticalities of research evaluation practices, shedding light on the good practices, and enhancing proposals across the different disciplines. Although the DORA Declaration has been criticised in some way, for instance, for its overemphasized focus on hard sciences (Gadd, 2015), it had the merit to open a wide discussion on the importance of monitoring the effectiveness of research evaluation practices.

Another worth quoting contribution is the Leiden Manifesto[20] (Hicks et al, 2015), which had a more comprehensive approach towards all the disciplines, and, according to its stated missions, it aimed at fostering research institutions' autonomy. Interestingly, the Leiden Manifesto was created from a discussion focussed on the consequences of the pervasive use (and misuse) of performance indicators. The Manifesto signers thus proposed the ten best practices in metrics-based research evaluation, with the primary purpose of making evaluation procedures more transparent for researchers, suggesting to avoid an overestimation of the goodness of the tool of the indicators.

As both the DORA Declaration and the Leiden Manifesto suggest, the evaluation tools and modalities with which research evaluation is conducted, namely the peer review process and indicators, are at

[19] https://sfdora.org/read/

[20] The Manifesto ten principles are the following: 1) the importance of matching qualitative and quantitative evaluation; 2) considering the social and local context alongside the performance indicators; 3) protecting excellence in locally relevant research; 4) transparency of the data collection and analytical process; 5) data transparency for the evaluated ones; 6) considering the publication and citation specificities of each field; 7) considering also the portfolio of the evaluated ones; 8) "avoid misplaced concreteness and false precision"; 9) "recognize the systemic effects of assessment and indicators; 10) "scrutinize indicators regularly and update them".

the centre of the debate on research evaluation at large, since they should be continuously adjusted to the needs of the specific context in which they are used.

We are going to analyse the main contributions on the effects of research evaluation as follows: in the following section we are going to address the critics to the concept itself of research evaluation and its most common tools and practices, also reflecting on the effects of research evaluation on researchers' practices, and, in the last section, we are going to report the effects of research evaluation according to an institutional perspective.

To better understand the rationale of the main critics of peer review and performance indicators, it is worth proposing some considerations on the concept of research quality itself and the functioning of the definition of quality among scholars, and then apply these considerations on the several examples of research evaluation practices.

The concept of assessing the quality of research contributions by a group of peers has been long debated among scholars. Many authors maintain that scientific knowledge is not objective but represents the result of a series of struggles among scientists (Bourdieu, 1988). Indeed, according to the field theory proposed by Bourdieu, the criteria that defines quality are not stable and eternal but do represent the values of the scientific elite which has social and cultural power in that specific moment, as in any other social arena. Academics fight with each other to win the struggle and, consequently, the prestige related to it, from which the possibility of defining the criteria of academic excellence derives. Therefore, the specific struggle in the scientific arena among different points of view pre-eminently lies not in the search of the truth, but in the attempt to make one's own representation true for the rest of the scientific community. It is thus for this reason that the evaluation of scientific works by someone who is defined an 'expert' by and external agent (namely, the scientific elite), could also enable an incommensurable encounter between the evaluator and the evaluated researcher's point of view, which represent two incompatible representations of the reality. It is then for the specific capability of those in power to define which representation of the reality is

true, that evaluation is not an aseptic picture of reality, but rather it reproduces and produces the reality, reflecting and crystallizing the power relations within the academic community (Bonaccorsi, 2015).

Taking into account these considerations on the dynamics of the academic community, it is quite understandable the reasons why the peer review process has been widely criticized, besides its use in the institutional evaluation process. The peer review process, in order to be fair, should aim at granting an objective evaluation of research works, and, to do so, the assessment is generally conducted according to a double-blind procedure. However, the process of evaluation that would take place into panels is deeply emotional and culturally embedded, and reviewers could easily assume a series of behaviours (such as an homophilic behaviour, or the so-called *Matthew effect*) that may bias panellists' judgements, without granting the transparency and the control over the evaluation procedures to the evaluated researchers (Lamonte, 2009).

Mathias Binswanger (2014) stressed that this same mechanism does not grant transparency to the submitters of scientific works, and that the acceptance and the rejection of papers happens according to unclear mechanisms, without adequately explaining the rationale of the decisions or, also, making very questionable decisions (Atkinson, 2001, Frey, 2003), like, for instance, rejecting very important works which would have won a Nobel Prize afterwards. By contrast, the overproduction of scientific works, matched with the research evaluation policies, in the last thirty years did not go hand in hand with an increased quality of research production (Salter, Martin, 2001, Butler, 2003), whilst some misconducts or strategic behaviours by researchers, like fraud and plagiarism, were detected (Binswanger, 2014). Indeed, the peer review system, matched to the scientific prestige and the real necessity for researchers to publish as much as possible (*publish or perish!*), seem to induce researchers to have an opportunistic conduct (Frey, 2003).

The most common researchers' strategic behaviours can be summarized as follows:

1) *strategic citing and praising*, meaning that researchers tend to adapt their research activity as much as possible to the editors and reviewers' ideas and to cite other authors' works strategically (Frey et al, 2009);
2) *no deviation from established or mainstream theories* (Atkinson, 2001, Frey, 2003), for the fear of seeing rejected by top journals their work;
3) *form is more important than content*, in order to make even simple ideas more attractive (Binswanger, 2014);
4) *undermining of anonymity by expert networks*: in determined disciplines and topics and with determined researchers who submit their papers, the anonymity and the double-blind mechanism may be lacking, and this could possibly affect in some way the evaluation process by reviewers (Binswanger, 2014);
5) *revenge of frustrated experts*, which, at last, identifies the tendency of engaging in evaluations that brings in itself a personal connotation, due to the reviewers' personal experience (Frey et al, 2009).

The bibliometric indicators are generally used in the research evaluation process and in the scientometric analysis to assess the impact of scientific papers on the academic community, by using statistical and mathematical analysis. The bibliometric quantitative analyses the impact of scientific papers, researchers and scientific institutions, by considering one or more of the following elements: the number of publications of an author, the number of citations of each publication (*citation index*), the frequency with which the average article in a journal has been cited in a particular year (*impact factor*, calculated annually by the American company Thomson Scientific), and the productivity of a single author (H-index, proposed in 2005 by the physicist Jorge Hirsch).

The considerations and debate on the use of metrics in academic evaluation are mostly based on the *priority rule* mechanisms highlighted by Merton (1942). According to the author, the citation is the tool by which a researcher is given credit and scientific recognition for his own discovery. In the

academic community, scholars do compete for the primacy in a determined scientific field or topic, and, when a scholar gets social recognition for his work, he will probably get an amount of scientific prestige disproportionately higher (*Matthey effect*), in comparison to his colleagues. Therefore, the "crudeness of this measure" (Small, 2004) induced Merton to think that the citation would be unsuitable for a research evaluation purpose, for the following reasons. On the one hand, because of the disproportionate social recognition that an author might receive in his career, it would enable an unfair judgement of the quality of his future works and of his colleagues' works; on the other hand, it may encourage some opportunistic behaviours to get relatively high score, as, for instance, reciprocal citations with the purpose of being active part of the academic debate and to get social credit as an extension of an another author's prestige (Small, 2004), and a series of other strategic practices like the *salami tactics* (namely cutting publishing many articles from a single idea)[21], strategic co-authorship and increased number of authors per article, forgery and fraud by universities and the academic community (Binswanger, 2014), and, even in this case, strategic or negative citations[22] to manipulate the indicator.

Interestingly, the concern of some scholars over the use of indicators as an instrument to guide and to determine the quality of research products pre-eminently relies on the use of the rough numbers for the standardization, the comparison and ranking of incomparable and incommensurable pieces of work (Kuhn, 1970).

Hence, taking these considerations on the effects of the use of the two main ways for evaluating research quality as a starting point, we can summarize the rationale of the two sides of the debate on the goodness of research evaluation.

On the one hand, the supporters of research evaluation (who, remarkably, are disproportionately fewer than critics) basically maintain that research evaluation is important in order to enable a reflexive discourse within the academic community with the purpose of monitoring and improving

[21] See Weingart, 2005.

[22] In this respect, Bonaccorsi (2015) maintains that a paper which has no impact at all would not be cited in any case.

the knowledge production process. Moreover, the evaluation culture, when institutionalized, represents a universal good because it let policymakers bring efficiency in the academic environment (Dahler-Larsen, 2012).

Bonaccorsi (2015) in his work explores the series of reasons why the assessment of other researchers' works represents the essence of the scientific method itself. First, getting the scientific recognition by the peers rests on the possibility of communicating the results of a research, which must follow the norms of the afferent community and that let the peers to mentally repeat the several steps of the research, and who then validate the research under examination. Secondly, the submission of a research paper to the peers community and getting a scientific reputation is fundamental for getting the access to the scientific community. Thirdly, the evaluation process of the colleagues' works is a constitutive and positive element for the scientific community itself (and has not to be considered as punishing), since it enables a reflexion on the functions and norms of the reference community, and, possibly, an improvement of the research practices. Fourthly, the formulation of a judgement by the peers (which is inevitably, imperfect, and fallible) tends to follow the democratic principles of the majority rule within the academic community, when a disagreement occurs. Indeed, the author claims that eventual disagreements within the academic community are not a good enough reason to deny the benefits of the institutional research evaluation system, and that wrong evaluations can always be corrected by the reference academic community. Hence, the institutionalization of the peers' judgement should occur by making the process and the evaluation criteria clear and transparent, to avoid, and eventually checking distortions and failures in the evaluation process. In the institutionalized research evaluation process, the academic community has then to agree on the definition of the evaluation criteria and how to define research quality, without necessarily agreeing on the evaluation of specific research works (in this case, then the democratic rule prevails).

On the other hand, alongside the several critics and aspects analysed previously on the peer review process and on the bibliometric indicators mechanisms, which represent a remarkable aspect in the

debate, a conspicuous number of critics of research evaluation maintains that institutionalized research evaluation is not suitable for evaluating and eventually fostering research quality.

Therefore, the main point that emerges from the works of the most critical authors is that academic organizational setting has been restructured by the implementation of research evaluation policies, that introduced a more formalized pyramidal hierarchical system (Elton, 2000), enabling a steering action on researchers' activities, thanks to the use of performance measurement tools and the establishment of research goals by central state,[23] which may remarkably steer researchers' freedom and induce them to adopt publication-oriented strategies in many ways (Borrelli, Stazio, 2018, Watermeyer, Olssen, 2016, Elton, 2000).

In the following section we are going to study the several voices on the effects of research evaluation (and the current discussion about it) as follows. Firstly, we are going to analyse many authors' contributions, who have tried to examine the effects and criticisms of research evaluation, primarily wondering whether the research evaluation policies have actually achieved the goal of enhancing the quality of research production, with a further reflection on the eventual changes in researchers' research practices induced by research evaluation; secondly, we are going to review the major contributions on the effects of research evaluation on the university as organization.

First of all, the *publish or perish culture*, and the need to publish as much as possible in top journals, has deeply affected the researchers' freedom in the study and discovery process that characterized the traditional researchers' approach to science (Binswanger, 2014), and the researchers' pursuit of excellence, defined as the discovery of something new or the introduction of a new research practice or approach in the scientific debate (Borrelli, Stazio, 2018). Then, the strong drive to publish could lead, in the first place, researchers to avoid topics or approaches that may need a lot of time for getting some valuable result, or, otherwise, to reject *tout-court* some research topics which seems to be very

[23]According to Gambardella, Grimaldi and Lumino (2019), the twofold action of technical instruments like the performance measurement tools and the establishment of research goals is part of the process of *governmental depoliticization* of the university system, which subtracts the processes and the objectives from the public scrutiny and debate, and subsequently attributes to the specific subject the aspect of necessity.

unproductive but from which some interesting works can have emerged afterwards (Borrelli, Stazio, 2018).

Secondly, some of the changes in researchers' behaviours can be ascribed to the gaming strategy. Indeed, performance measurement may generally induce individuals to adjust their own behaviour and adopt strategies to accomplish the fixed objectives thus enabling a series of negative unintended practices which may bias the performance assessment process (Lewis, 2015, Smith, 1995). Among these gaming strategies, we can find the *salami tactics* (as seen previously) and other strategic researchers' choices that may improve their performance indicators (as, for instance, to favour determined editorial outlets, like journal articles, at the expense of the monography or the book chapters, to give priority to international outlets)[24]. As for the former, research evaluation would have triggered a mechanism in which researchers are going to publish an increasing amount of low quality and low impact papers (Butler, 2003), that also makes research evaluation less cost-effective on the long term (Pinar, Horne, 2021). Interestingly, some authors maintain that the causal correlation between research evaluation policies and the increase of knowledge production is quite debatable (Osuna et al. 2010, Rijcke et al. 2015). According to Osuna et al. (2010), the increase of published papers in the Spanish higher education system would be attributable to the increased public investments in research, rather the introduction of research evaluation policies.

Thirdly, Hammarfelt and Rijcke (2015) maintain that the introduction of research evaluation policies has induced some changes in researchers' publications patterns. For instance, the study they conduct at Uppsala University showed that researchers give more importance to scientific articles, give priority to publish in English in international outlets, prioritising more international approach in their research agenda setting process. The authors maintain that the shift from the preference of the scientific paper rather than the monography may lead to an impoverishment of the cultural richness of scientific works.

[24] De Rijcke et al. 2015, Hammarfelt and de Rijcke, 2015, Binswanger, 2014, Elton, 2000.

Interestingly, the available literature present contradictory results on the effects of research evaluation on the interdisciplinary process (Rijcke et a., 2015). Indeed, if in some studies it seems to have discouraged transdisciplinary research (Elton, 2000), in other studies research evaluation seems to have fostered a multidisciplinary approach and a collective researchers' endeavour, also proved by the increasing number of transdisciplinary co-authorship, and scientific collaborations between researchers from research institutes and researchers working in private enterprises (Colarusso, Giancola, 2020, Van Den Besselaar et al, 2012, Henkel, 2005).

According to Whitley (2007), the effects of research evaluation on researchers strongly depend on the "strength" of the national evaluation system. Indeed, in strong systems researchers' research agenda seems to be more likely sensitive to external influences, and, moreover, that researchers from humanities and social sciences seem to imitate research practices of more "prestigious" disciplines (since they can get external funding more difficulty and are more sensitive to public funds).

According to some authors (De Philippis, 2015, Campbell, 2013), the overwhelming burden linked to the pressure to publish is likely to induce researchers to neglect a series of activities that constitute a relevant part of academic life (namely, teaching activities, writing referee reports or translating research outcomes for policy, to hold an institutional role, providing consultancy to policy makers, etc). Still, the excessive workload and this overemphasis on research performance seems to be one of the causes of a general increasing distrust, frustration, anxiety and sense of isolation of researchers (Rijcke et al, 2015, Moss, Kubacki, 2007).

According to Karlsson (2017), research evaluation seems to foster the researchers' focus on the quality of research production, thus "strengthening leadership, communication, and administrative order". Interestingly, the study shows that research evaluation has increased the quality of research production in the least successful teams (even though, according to the author, in some cases some strategic behaviours were detected), encouraging good practices to achieve the pre-established goals.

In the end, some contributions highlight that there seems to be an increasing a sense of distrust, anxiety, and powerlessness among the academic community (Watermeyer, Olssen, 2016, Burrows,

2012), partly due to the increasing emphasis on performance indicators of researchers' productivity, coupled to the high institutional pressure to publish.

The implementation of research evaluation policies has produced very differentiated effects on the university as an institution. Indeed, this process has not generally been either linear (except for the UK, Paradeise et al, 2009) or homogeneous across the different countries, with a preeminent path dependent approach.

According to the available literature, some major general effects on higher education institutions can be easily identified. First, in many cases research evaluation policies seem to have stimulated the strengthening the quality of university management (Minelli et al 2006, Elton, 2000, Musselin 2013), encouraging a collective endeavour by the academic community to achieve the required targets. More specifically, the introduction of research evaluation policies seems to have redefined the importance of the managing and administrative roles, which are now more entitled to frame medium and long-term strategies, adopting faster and easier ways of communicating with the other operative university areas (Minelli et al, 2006). Indeed, the introduction of research evaluation policies has encouraged an adjustment of universities hierarchies in a more vertical way, to exert higher control over academics' activities (Musselin, 2013, Campbell, 2013).

Some authors maintain though that the overemphasis on university management may enable a too strict managerial control which may limit researchers' freedom, without having real positive effects on the economic relevance of researchers' research production (Hicks, 2012, Watermeyer, Olssen, 2016, Capano, 2014).

Interestingly, the strengthened role played by administrative staff and academic elites (Musselin, 2013), paired with the high level of competition among different universities and the severe

consequences that the universities may undergo whether they cannot achieve the required targets,[25] is apparently inducing universities to assume market-inspired hiring policies (Colwell et al, 2012). Indeed, universities seem to considerably rely on individual performance indicators to decide whether to hire or not a researcher. Interestingly, it is also worth mentioning that the introduction of research evaluation policies has induced a fragmentation of the academic profession tasks, since academics should increasingly join several research networks, conduct more administrative tasks (Middlehurst, 2014, Musselin, 2013), thus enabling a process of redistribution of tasks and counterbalancing of power within the academic elite (Musselin, 2013, Reale, Marini, 2017).

The implementation of research evaluation policies and the pursuit of academic excellence seem to produce diversified effects on academic institutions which have different organizational setting. Indeed, Minelli et al (2006) have compared the impact of research and teaching evaluation in Italian (more recently implemented) and Dutch universities (more consolidated evaluation system).

Hence, the Italian case seems to be richer in consequences, for the relative novelty of the policies and procedures introduced, and, on the other hand, in the Dutch case the effects of evaluation seemed to be more in a phase of consolidation. Still, since in both case studies the implementation research evaluation policies have been gradual, the effects in the Italian case seem to be more fragile in comparison to the Dutch case, which present more long-term effects.

Moreover, Paradeise and Thoenig (2013) analyze the diversity of the organizational responses of universities with different core principles (*the top of the pile*, the *venerables*, the *missionaries* and the *wannabes*) to the external input such as the pursuing academic excellence and the introduction of research evaluation policies. More precisely, the four types of academic institutions are clustered according to their degree of attention to reputation and their attention to academic excellence (the *top* of the pile universities high attention for both, the *venerables* high attention to reputation but low

[25] According to Watermeyer and Olssen (2016), one of the most negative effects of research evaluation policies is the allocation of funding on a performance basis, which could enable a vicious circle and completely defund some universities in the long-term.

attention to excellence, the *missionaries* low attention for both categories, and the *wannabes* high attention to academic excellence but low attention to reputation). The authors first find out that the universities which both focus on academic excellence and on reputation have had to adapt more to the introduction of research evaluation policies, since they perceive more the institutional pressure. The authors also stress the importance of the local and traditional context of the higher education institutions at issue because they determine the differences of the application and the consequences of the research evaluation policies, and of the same definition of quality excellence.

Moreover, some authors identify diversified effects among different disciplines. First, Whitley (2007) considers both the "strength" and the "weakness" of the research evaluation systems and the peculiarities of the disciplines too. Indeed, the disciplines that seem to undergo a process of standardization and homogenization of procedures, practices and research agenda setting approach, due to the introduction of research evaluation policies, seem to be those who can difficulty find financial resources outside the academic context, namely the humanities and the social sciences with little application degree. This trend is even stressed in the strong research evaluation systems. Interestingly, the author maintains that in these fields researchers tend also to be less divergent in the research agenda setting, and that the topics and research practices are going toward a process of integration and coordination, even at national level. Still, the harmonization of the research topics is even intensified in those disciplines in which the intellectual elite generally agree on the research agenda setting process and, on the procedures, to adopt. On the other hand, when in a "research evaluation policies sensitive" discipline the intellectual elite present a more differentiated approach to the selection of research topics and procedures, research evaluation may stimulate intellectual competition among the different actors.

The disciplines who can easily find financial resources outside the academic context, by contrast, seem to be less sensitive to research evaluation, and can generally maintain high degree of autonomy in a broad sense.

Interestingly, Reale and Seeber's study (2013) seem to present similar conclusions. The study has indeed analysed the response to the introduction of REPs in different disciplines. Therefore, according to the data collected, in the Management departments[26] the introduction of research evaluation policies has stimulated a certain level of competition with other departments, "thus strengthening the trend towards fundamental research and international journals". By contrast, natural sciences seem to be less sensitive to the evaluation outcome *per se*, rather these departments seem to be sensitive to funding shortages, because of the high cost of the instruments necessary to their activities.

To conclude, another example of the diversified disciplines behaviour towards research evaluation is provided by Bonaccorsi (2015), who proposes a conceptual acceptance or rejection attitude assumed by the different disciplines in the Italian academic context. More specifically, the disciplines that can easily accept internal discussion and plurality, and that valorise empirical evidence, tend to have a more positive approach to research evaluation. On the other hand, disciplines that difficulty accept internal plurality and that do not value the empirical evidence, tend not to positively accept research evaluation practices.

1.4 Final remarks

In this chapter, we have proposed the historical and conceptual path that has led to the introduction of remarkable changes in the academic environment. Still, after a brief sketch of the main features that characterize research evaluation in general, we have attempted to provide a thorough overview of the main models that try to cluster the research evaluation systems of several countries. As it emerged, however, trying to cluster national research evaluation systems is very challenging, for the different purposes, functioning and definitions of the policies themselves, that each country propose. Interestingly, many criticisms on the principles and functioning of research evaluation were advanced by scholars. More specifically, some scholars' main considerations on the consequences of research

[26] University of Bologna and University of Modena-Reggio Emilia

evaluation on researchers' research practices (adaptation strategies, opportunistic research behaviours and publishing), and the goodness of both the peer review system and the bibliometric indicators to properly assess the quality of scientific works, were proposed, as conceptual background for the following chapters.

Chapter 2
Research evaluation and academic profession in the Italian context

In the present chapter, we are going to illustrate the specificities of the Italian academic system, with specific references to how research evaluation was introduced. With an halting and overdue process, with respect to other major European economies, that started in the late 80s, Italy finally endows itself with a standardized, transparent and nation-wide research system in 2010, with the introduction of the VQR (*Valutazione della Qualità della Ricerca*). In this chapter we are first going to retrace the path that led to the gradual introduction of research evaluation, with a particular focus on the major public debates and reflections that accompanied the implementation of the various laws, till the introduction of the VQR. Then, in the last part of the chapter, we are going to describe and analyse the different functioning and perspectives of VQR and ASN (the latter, aimed at framing the academics' hiring process).

2.1 Research evaluation in Italy: historical overview of the Italian academic system

The process that has led to the introduction of research evaluation instruments in Italy has started relatively later, namely in the early 2000s, in comparison to the first European countries that adopted these policies. The main reasons of this relative delay are basically linked to the conception of the utility and conception of university in the Italian socio-economic context (university was deemed as the education and training devoted to the social elites and to train it to the civil service), and the types of skills and professional figures required by the then-labour-market, who were not supposed to be highly specialized (Moscati, 2001).

As for what is concerns the Italian case, the introduction of research evaluation and the linked reorganization of the higher education system has been stimulated by a double pressure exerted by both national and international factors (Colarusso, Giancola, 2020). Indeed, on the one hand, the Italian higher education system was experiencing, like in many other countries, both the massification

of the university access (for which more public resources were needed) and the necessity of reducing the public financial investment in higher education, which determined the necessity of restructuring and reorganizing the university management. On the other hand, the incentive to harmonize the European higher education systems by an intergovernmental initiative (the so-called Bologna process, as already discussed in the previous chapter), and the consequences of the involvement of international organizations in research and universities activities, such as the European Commission, with the institution of the Framework Programmes (Paradeise et al, 2009), further encouraged the rethinking of the relationship between central state and intellectual elite, and the management of universities at large.

Before starting, it is worth contextualising the historical reconstruction of the path that led to the implement of the Italian research evaluation system, by briefly sketching some of the peculiarities of the Italian higher education system.

Firstly, Italy is a major industrialised country, but, in comparison with the other European countries, research seems to play a marginal role for the development of the economy. Italy indeed invests in R&D only 1.49%, which is far below the average European R&D expenditure (2.3% of GDP, Eurostat, 2017), and its number of researchers per inhabitants is lower than that of the other European largest economies (Italy 2.307, Germany 5.212, France 4.715, Austria 5.733, Spain 3.001[27]). Furthermore, even private companies' investment in R&D is low, and partially attributable to the same average features of Italian private enterprises: medium and small enterprises represent almost the 95% of the industrial population, which, comprehensibly, have a limited economic capacity to invest in R&D, the high-tech industries play a relatively marginal role (Potì, Reale, 2009), and only a very limited number of researchers are involved in applied research and research projects geared to enterprises' technological development (Rostan, 2011).

[27] https://statnano.com/report/s90

Secondly, the Italian cultural context seems not to encourage the involvement of graduated people in the productive system:

1) employment rate of graduates is 4,3 points lower than the European average (82.1%, with a far worse employment rate in Southern regions, Istat, 2022),
2) the ruling class itself seems to discourage students from attending the university,
3) and the same entrepreneurs have a very low average education level[28].

Thirdly, Italian higher education system has proven quite ineffective in reducing social and economic inequalities so far. Indeed, on the one hand, parents' educational level seems to be a quite good indicator of children's future educational level (Borrelli, Stazio, 2018), with concrete consequences on their future economic and social conditions. On the other hand, in the academic context the gender gap is still a remarkable issue, with only 35.1% of female researchers (Eurostat, 2022, Marini, Meschitti, 2018).

It is therefore possible to identify two main phases of the process of introduction of research evaluation measures:

1) the first phase, whose beginning can be identified with Act of 1989 no. 168, is mainly characterized by, on the one hand, the absence of structured and nationwide research evaluation exercises, but on the other hand, by a considerable degree of experimentation about the research evaluation;
2) the second phase sees the creation of the ANVUR (National Agency for the Evaluation of the Academic System and Research) and the institutionalization of nationwide research evaluation exercises.

It is worth remarking that both phases are considerably characterised by a trial-and-error approach, because the research assessment exercises and institutions implemented during these two phases

[28] The percentage of entrepreneurs with a degree is very low, around 14% (Borrelli, Stazio, 2018).

underwent a series of recurring structural changes, even in the final period, when the research evaluation exercises and institutions started to assume a more consistent pattern (Abramo, D'Angelo, 2015). This trial-and-error approach is also linked to the difficulties to consolidate a "genuine evaluation culture" within the Italian academic context, centred on an effective and efficient accountability system (Wang, 2022).

Hence, the beginning of the first phase coincides with the promulgation of three bills: firstly, Act of 1989 no. 168, that established the introduction of universities' autonomy[29]; secondly, Act of 1993 no. 537[30] that established the reformation of universities' funding system and introduced universities' internal evaluation units (*Nuclei di valutazione*); thirdly, Act of 1997[31] no. 57 laid the foundation of the rationalisation of public sector facilities (and, among them, university) and the institution the evaluation and monitoring instruments.

Firstly, Act no. 168 established that, in compliance with the article n. 33 of the Italian Constitution[32], universities have legal personality. For this reason, universities started to be financially and organisationally autonomous, and, additionally, independent for what it concerned the teaching and research activities[33]. Universities were also in charge of the maintenance and the functioning of university facilities and the organizational settings, and they could operate with broad discretion for the management of human resources and the recruitment process. As the result of the passage of the law, the Ministry of Higher Education was entitled to provide the framework and coordination of research and technological activities[34], within the limits foreseen by the Article 33 of Italian Constitution. As a matter of fact, universities' autonomy entailed the opportunity to enact their

[29] Legge 9 maggio 1989, n. 168: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1989-05-09;168#:~:text=%2D%20L'arte%20e%20la%20scienza,senza%20oneri%20per%20lo%20Stato.http://www.miur.it/0006menu_c/0012docume/0098normat/1668istitu.htm

[30] Legge 24 dicembre 1993, n. 537: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1993-12-24;537,%20art.%2022%20e%2023

[31] Legge 15 marzo 1997, n. 59: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1997-03-15;59!vig=

[32] "Arts and sciences are free, and free is their teaching." https://www.senato.it/istituzione/la-costituzione/parte-i/titolo-ii/articolo-33. Since that moment, as a matter of fact, universities were only partially autonomous.

[33] As we will se, within the limits posed by the Minister of higher education.

[34] Article n. 2 of L. 168/1989.

internal regulation bills (under approval of the qualified majority of universities governing bodies), that had to be transmitted to the Ministry of Higher Education, who should verify the legitimacy of the statutes. In legislators' mind, this reform was meant to pave the way to the resettling of the previous university management, alongside the reforms of the didactic systems (*ordinamenti didattici*) and the institution of doctoral course of studies. According to the recording (related to the 29th March joined session, Seduta di Mercoledì 29 1989) of the joined Parliamentary Committees sessions, which were working on this draft law, the legislators aimed at underlining the importance of the detachment of Higher Education affairs from the Minister of Education, by creating the Minister of Higher Education[35]. The newly created Minister was supposed to have a more flexible and less bureaucratic internal structure and to engage the academic collegial bodies (like the *Consiglio Universitario Nazionale*, the *Consiglio Nazionale della Scienza e della Tecnologia*, *Consiglio nazionale geofisico*, *Consiglio per le ricerche astronomiche*) in the law-making process.

Most importantly, Act no. 168 first introduced the concept of the evaluation of universities' activities, by laying the foundation of the debate on research evaluation, that is the core of the measures that would have been implemented afterwards. Indeed, both the recordings of the joined sessions of 29th March 1989 and 5th April 1989 (*Seduta di Mercoledì 5 Aprile 1989*), report some speeches of the Members of the Parliament who solicited the then-Ministry and the Committees to consider the reform of higher education system as an open issue, with some important questions to be deepened afterwards. Indeed, on the one hand, some Members of the Parliament recommended the Government to consider the creation of an institution for the evaluation of universities' research activities, to monitor the effects of universities' autonomy on researchers' performance. On the other hand, some politicians discussed (especially in the joined session of 5th April 1989) about the necessity of better defining the limits of universities' autonomy, to avoid to different kinds of problems: firstly, an

[35] The Minister of Higher Education and the Minister of Minister of Education were re-merged, according to Legislative Decree of 30th August 1999 n. 300 (https://www.parlamento.it/parlam/leggi/deleghe/99300dl.htm), with the purpose of rationalizing the public administration.

unregulated process of autonomy could have led to an excessive attribution of power to universities; secondly, without a proper regulation, a pseudo universities' autonomy was also possible, that would have been basically in continuity with the prior academic institutional setting.

It is wort noting, though, that the debate about the introduction of the principle of universities' autonomy, and its subsequently practical application, did not directly involve the academic community (Moscati, 2001).

Secondly, with Act of 1993 no. 537 the public higher education funding system (*Fondo per il Finanziamento ordinario, FFO*) was reorganised, and the institution of universities' internal evaluation units (*Nuclei di valutazione*) was made mandatory. In the years preceding the Act of 1993 no. law 537, some universities had already provided themselves with internal evaluation units, and some of these adopted the suggestions from the working group on research evaluation created by the Conference of Rectors[36] (*CRUI*). Therefore, with the implementation of Act n. 537, all the universities were supposed to create their internal evaluation units, and in many cases the Conference of Rectors' guidelines were helpful for this specific purpose.

The internal evaluation units were thus supposed to submit to the Ministry of Higher Education and to the national university council[37] (*CUN*) an annual report to allow an evaluation of the efficacy and the efficiency of universities' teaching and research programs.

Thirdly, Act of 1997 no. 57 (also called *Legge Bassanini*, from the name of the then-Ministry for the civil service), which was inspired by neo-liberal principles and aimed at substantially rationalize the entire Italian public sector system, introduced three major reforms:

1) the streamlining of the administration and organization of public sector facilities,

[36] Incontro Nazionale sulla Valutazione del Sistema Universitario, Osservatorio per la valutazione del sistema universitario.

[37] Introduced in 1979, the national university council has then been reformed in some occasions. The law n. 18/2006 has qualified the CUN as an elective body which represents university community's interests, and that mainly performs a consultative task within the Ministry of higher education, concerning the functioning of the university system in its several areas (https://www.cun.it/cun/).

2) the provision of administrative autonomy to regional and local institutions (the text of the law also identifies the caveats of the application of the autonomy);
3) the introduction of mechanisms to evaluate and monitor public facilities' performances.

Act no. 57, which is the first of a complex of four laws focussing on the reform of public administration, set the juridical foundations of the devolution of tasks and management responsibilities, in compliance with the principle of subsidiarity[38]. However, the subparagraph 3 of art. 1 of the law lists the subjects and areas to which the autonomy of local and regional institutions was not supposed to apply, and that, by contrast, were under the control of the related ministers. On school and university, as a matter of fact, regions' autonomy did not apply.

Indeed, Act no. 57 explicitly entrusted the Minister of Higher Education to deal with a series of tasks concerning the restructuring of higher education, among which:

1) overseeing the simplification and the reordering of the procedures in support of scientific research and didactic activities of universities,
2) clearing up and streamlining the functions of the minister's consultive bodies, in order to avoid the overlapping functions among the several bodies
3) defining the boundaries of the autonomy of universities, and
4) creating an inventory of the actors and bodies working in the higher education sector.

It is worth underlining that during the discussion that preceded the publication of the bill, some politicians expressed their concerns over the alleged over-centralization of tasks and control in the hands of the Minister of Higher Education (*Resoconto della I Commissione Permanente*, recording of 18th December 1996). As a matter of fact, these concerns were mainly related to the possibility of hindering that process of universities' autonomy that would have enabled the reorganization of higher

[38] Subparagraph 1 and 2, art. 1, law 57/1997: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1997-03-15;59!vig=

education facilities, according to the local and social specificities of the context the universities were embedded in.

On the other hand, other Member Parliaments, who participated to the draft of the bill, deemed positive the frequent references to the gradual convergence of Italian higher education towards international "good practices" (*Resoconto della I Commissione Permanente*, recording of 21st January 1997), and suggested that the process towards universities' autonomy and research evaluation had been positively encouraged by the international pressure.

According to the available recordings of the decision-making process of law no. 57, it emerges a twofold approach by the legislators towards research activity and its evaluation. On the one hand, it emerges consensus on the peculiarities of research and of research evaluation, in comparison to other public sector areas subject to evaluation mechanisms. For this reason, there was consensus on the goodness of not establishing research evaluation procedures according to criteria solely based on costs, since some features of research activity cannot be judged quantitatively (Reale, Pennisi, 2013), or on an economic return (*Resoconto della I Commissione Permanente*, recording of 16th January 1997). For this reason, the Ministry of Higher Education was entrusted to oversee the definition of the evaluation procedures, together with the support and consultancy of the evaluation bodies.

On the other hand, from Act no. 57 it deeply emerged the concept of scientific research and technology (and jointly also research evaluation) as an asset for society, the industrial sector and the administrative system at large (*Resoconto della I Commissione Permanente*, recording of 14th January 1997).

It is also worth underling that the process of autonomy and devolution of responsibilities to universities is only one part of a series of reforms that the Minister of Higher Education aimed at adopting to increase the competitiveness of Italian higher education. For this specific goal the Minister of Higher Education adopted a "mosaic strategy", to implement the series of reforms in a reasonable time span, attempting not to trigger much resistance from the academic community (Moscati, 2001). The series of reforms basically aimed at strengthening universities' productivity, by

implementing an "effective diversification of higher education into three levels of courses leading to degrees, the introduction of a binary postsecondary system, and the introduction of a credit system into all curricula" (Moscati, 2001).

Alongside the promulgation of the previously cited laws, during this first phase some national research evaluation institutions were created. Some of the most important tasks of these newly-born institutions were to consult the Ministry of Higher Education on the creation and testing of new research evaluation procedures, and to monitor and collaborate with the universities' internal evaluation units.

Therefore, research evaluation in Italy took shape with the institution of the *Osservatorio per la valutazione del Sistema Universitario*, instituted in 1996 and then converted in 1999 into the National Committee for the evaluation of the university system[39] (*CNVSU*), and the institution of the Committee for the Evaluation of Research (*CIVR*) in 1998.

The *Osservatorio per la valutazione del Sistema Universitario* had the primary task to work jointly with the universities' internal evaluation units, to evaluate the efficiency and efficacy of the research, administrative and teaching universities' outcomes. Indeed, according to Act of 1996[40] no. 261, issued by the Minister of Higher Education, the *Osservatorio* had to submit an annual report on the state of the Italian academic system, based on the reports submitted by the internal evaluation units of each university.

The CNSVU (which was part of the Minister of higher education and substituted the *Osservatorio*) was composed by nine members appointed by the Italian Government, who had proven expertise on research evaluation and university management. The CNSVU had several important tasks, including drafting an annual report on the evaluation of the university system, coordinate and elaborate the

[39] The National Committee for the evaluation of the university system was established with law 370/1999, and in 2000 it became operational. With the same law, the Italian Parliament reaffirmed universities' obligation to adopt internal evaluation units (*Nuclei di valutazione*), which had the task to record and evaluate the universities' organizational, teaching and research outputs, and to monitor the correct expenditure of the public resources available. http://www.miur.it/0006menu_c/0012docume/0098normat/1568dispos.htm#:~:text=Le%20universit%C3%A0%20assicurano%20ai%20nuclei,normativa%20a%20tutela%20della%20riservatezza.

[40] Ministerial Decree of 22nd February 1996, n. 261: https://www.gazzettaufficiale.it/eli/id/1996/05/16/096G0274/sg

information and reports from the universities' internal evaluation units, promoting the adoption and testing of evaluation procedures, projecting an annual universities' research evaluation system, consulting the Minister of higher education on the allocation of public funding to universities, and, to conclude, defining the standards and good procedures for universities' teaching and research activities.

This period of testing and experimentation of research evaluation procedures reaches its peak with the institution of the CIVR, which was instituted with Act of 1998[41] no. 204, and was composed by seven experts appointed by the Government. Interestingly, the CIVR had its seat at the Minister of higher education but was afferent to the Prime Minister cabinet (Reale, Pennisi, 2013). The CIVR was primarily conceived to conduct research evaluation on all research institutes' activities, including universities, public research institutes, the private research institutes which applied for being evaluated, and research projects, clustered by thematic evaluation panels, through both *ex-ante* and *ex-post* evaluation.

The CIVR was thus established to perform a series of tasks, such as: promoting the development of the culture of research evaluation, set the foundation of a nationwide research evaluation process, and consulting the Minister of higher education and universities on the best evaluation procedures to adopt (*Linee Guida per la Valutazione della Ricerca*, *CIVR*), and "improving the institutional link between evaluation and resource allocation" (Reale et al, 2007).

Still, the CIVR was entrusted to conduct the first national research evaluation exercise, the so-called *VTR* (Three-Year Research Evaluation Exercise), that started in December 2003. The VTR was a research evaluation exercise that aimed at evaluating the research production from 2001 to 2003 and was directed to both public and private research institutes, which could participate on a voluntary basis (Reale, Pennisi, 2013). One of its scopes was to encourage individual researchers and universities' self-evaluation, whilst other concrete uses of the evaluation outcome were not

[41] Legislative Decree of 5th June 1998, n. 204: https://www.gazzettaufficiale.it/eli/id/1998/07/01/098G0257/sg

completely clear (Minelli et al, 2008). The VTR was based on fully peer review *ex-post* assessment, and were evaluated 102 structures, among which 77 universities, 12 public research institutes and 13 private research agencies, which comprehensively submitted 18,500 research products, which could be assigned to one of the 20 disciplinary sectors identified by the CIVR, for each of whom panels of experts were set up (Franceschet, Costantini, 2010, Minelli et al, 2008). The VTR was articulated in three main phases. During the first phase, research institutions had to submit the products[42] to be evaluated and the data about human resources, research, and university management. It is worth citing that universities were not asked to provide products for each department, but the departments that did not provide any research product were considered "scientifically" inactive (Reale, Pennisi, 2013). In the second phase, panellists assigned research products to external referees for the assessment procedure of the research products (each product was assigned to at list two referees). At last, during the third phase, the CIVR had to produce a report on the evaluation outcome, integrating the panellists' relations with the information of each research institute (Franceschet, Costantini, 2010). The final report was issued in 2007, and it comprised both a summary evaluation of the structures which participated to VTR, and an evaluation of the national research system. According to Minelli et al. (2008), VTR presented some interesting features, because of its character of innovation in the Italian context. Indeed, even though not particularly appropriate for thoroughly evaluating the state of universities' research production (since not all universities participated and the selection of research products only reflected the excellence of the research institutes which applied), VTR presented some interesting features: it entailed a massive use of peer review, which was innovative for the Italian context; it contributed to the creation of an evaluation culture and, consequently, to the acceptance of an institutional research evaluation system; for the first time, a public rating list of the research institutes that participated to the evaluation exercise was created, which seemed to be quite courageous in the then-Italian academic context.

[42] Books, chapters of books, scientific paper, patents, exhibitions, art operas, designs.

However, the main limits identified by the authors were, in summary, the fact that the participation to VTR was non-compulsory for publicly funded universities, the questionable nature of the criteria for the drafting of the ranking lists of universities, and, in the end, that CIVR paid little attention to the aspects of transparency of the evaluation procedures.

According to Act of 2010[43], the CIVR was supposed to conduct a second national research evaluation exercise for the four-year period 2004-2008. However, the process was sharply interrupted by the institution of the *ANVUR* (National Agency for the Evaluation of the Academic System and Research), established by Act of 2006[44] no. 286, and then regulated four years later by the Decree of the President of the Republic n. 76/2010[45]. The ANVUR was supposed to merge the CNVSU and the CIVR.

With the institution of the ANVUR, the phase of institutionalization of the Italian research evaluation system begins. However, the overlapping of the time frames (namely, from the approval of the MD 8/2010, which regulated the second VTR cycle, and the institution of ANVUR, which was supposed to substitute the CIVR in conducting evaluation activities) triggered some concerns in the academic community (Reale, Pennisi, 2013), and the second phase of the Italian research evaluation system seemed to start in a rather chaotic way.

The chaos around the establishment of a straightforward research evaluation mechanism was also apparent in the political discussion of the bill. Indeed, the parliamentary discussions of Act no. 286 were rather troubled, and, for instance, the Member Parliaments from opposition parties maintained that the institution of the ANVUR was not compliant to objective criteria, and that it would have not fixed the well-known Italian academic system problems (*Resoconto della VII Commissione Permanente* - Recording of 12th October 2006). By contrast, the Members of the Parliament from the governing parties replied that the institution of the ANVUR was necessary in order to give order and

[43] Ministerial Decree 8th July 2010: https://www.gazzettaufficiale.it/eli/id/2010/10/01/10A11544/sg
[44] Law of 24th November 2006, n. 286: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2006-11-24;286
[45] Decree of the President of the Republic of 1st February 2010, n. 76: https://www.gazzettaufficiale.it/eli/id/2010/05/27/010G0098/sg

consistency to the research evaluation reforms that had occurred in the previous years. After a lively debate, the bill was then approved, with the explanatory statement that it was necessary for Italian academic system to converge towards other countries' research evaluation systems. Furthermore, the necessity to rationalize and reorganize the public investment in R&D was deemed even more urgent after the severe cuts to research investments made by the previous government (*Resoconto della VII Commissione Permanente* - Recording of 21st October 2006, Abramo, D'Angelo, 2015).

The internal structure and duties of ANVUR were then ruled by the Act of 2010 no. 76[46] and have remained unchanged so far. ANVUR's main roles are the President, the Executive Council and the Auditors Council. The term of office is of four years. ANVUR's duties are to be performed according to the principles of autonomy, impartiality, professionality, transparency, and public access to documents. ANVUR primary tasks are to: overseeing and supervising the national public system of evaluation of universities and public research institutes' activities; addressing the activities conducted by universities' internal evaluation units; evaluating the efficacy and efficiency of public funding programmes; promoting research activity and innovation; being in charge of indicating to the Minister the full and associate professors for creating the evaluation committees for the nomination of Italian researchers as full and associate professors, and creating the benchmark for the nomination (as we will see in the next paragraph); carrying out the function as national agency for the quality assurance and collaborating with both European Union bodies and international administrative agencies through exchange of information, in compliance with the principles and objectives of the European Higher education AREA.

As for what it specifically concerns evaluation activities, ANVUR main tasks are to:

1) evaluating the quality of processes and of outcome of managerial, teaching and research activities, with a specific focus on the transfer of knowledge to the local communities,

[46] Decreto del Presidente della Repubblica dell'1 Febbraio 2010, n. 76: https://www.gazzettaufficiale.it/eli/id/2010/05/27/010G0098/sg

2) defining the basic procedures, the minimum benchmark, and methods for the research evaluation of universities and research institutes, according to transparent and objective criteria, and promoting the self-evaluation practices,
3) setting the criteria for the allocation of public funds to universities, also considering the outcome of research evaluation exercises, and evaluating the efficacy and efficiency of public financing programmes.

Act of 2011 no. 111[47] established that ANVUR had to oversee the first VQR (*Valutazione della Qualità della Ricerca*, henceforth VQR1), namely the first nationwide mandatory research evaluation exercise, which evaluated the time span from 2004 to 2010.

The implementation of the VQR1 has entailed a series of new elements in the Italian research evaluation context, of which the most relevant are the following:

- the units of evaluation are universities, university departments, and publicly funded research institutes;
- participation to VQR is no longer on voluntary basis, and each research has to submit three research products published in the period at issue;
- VQR outcomes determine a remarkably higher amount of university funding, in comparison to VTR (less than 2% of FFO, Abramo, D'Angelo, 2015); still, the relevance of VQR in determining the resource allocation has increased over time (7% in 2009, 20% in 2015, 23% in 2016, Wang, 2022);
- universities have to submit the research products of the researchers working at the moment of the publication of the call for notice;
- the rating of the research products is more selective in the VQR than in the VTR, meaning that there is an increased difference in terms of score assigned to excellent and good products, and second-rate products (Reale, Pennisi, 2013);

[47] Legge 15 luglio 2011, n. 111: https://www.gazzettaufficiale.it/eli/id/2011/07/16/011G0153/sg

As we will better analyse in the next paragraph, the first VQR was considerably criticized by the wide academic community, and, with the subsequent two VQR, respectively concerning the time frame 2011-2014 (henceforth VQR2, established with the Act of 2015, no. 458[48]) and 2015-2019 (henceforth VQR3, established with the Act of 2019, no. 1110[49]), the ANVUR has partially tried to improve the methodologies adopted, in order to address the most relevant critical issues emerged.

2.2 VQR and ASN: two instruments, two perspectives

The purpose of evaluating research products represents a very challenging issue, since with the term *research* it is possible to identify a complex of human activities in which the dimensions of freedom and creativity of the researcher play a crucial role, and in which the process of studying that may conduct to important theories is not linear and straightforward (Cerroni, 2009). However, the increased social demand for establishing monitoring instruments and benchmark on research institutions' activities, has encouraged some scholars to reflect on the criteria for the scientific choices that are at the basis of the evaluation and the rating of scientific products. According to Weinberg (1963), the "scientific choice" underpinning the evaluation of research products, conducted by the peers or an evaluation agency, should follow internal criteria ("How well the science is done"), and external criteria ("Why pursue this particular science?"). More specifically, with internal criteria the author referred to the series of internal rules that a specific scientific community recognizes to understand whether a research product is scientifically valuable. With the external criteria, on the other hand, we can identify the compliance of a determined scientific product with a certain community's social and economic needs, technical progress, and ethical values (Cerroni, 2009).

[48] Ministerial Decree 27th June 2015, n. 458: https://www.anvur.it/wp-content/uploads/2015/07/dm_27_06_2015_vqr_11-14.pdf
[49] Ministerial Decree of 29th November 2019, n. 1110: https://www.miur.gov.it/documents/20182/484377/DM+n.1110_29.11.2019.pdf/a9191c4d-4e08-d008-a85b-3f975fce0919?version=1.0&t=1575040886555

Hence, we can say that a specific research evaluation system is defined by the choice and definition of the same dimensions it aims to evaluate (Fasanella, Martire, 2017), which, indeed, represent the *habitus* of a scientific and social community in that specific moment.

In this section, we are going to analyse the main features and dimensions of the two main research evaluation instruments introduced in Italy, namely:

- the already-mentioned VQR, which aims at evaluating research production on university and university department level, hence according to a *meso* perspective (Cerroni, 2009);
- and the ASN, namely the National Scientific Qualification, which, by contrast, focuses on a *micro* perspective since it evaluates individual researchers' scientific production.

2.2.1 VQR

As already mentioned in the previous paragraph, the VQR[50] is a nationwide research *ex post* assessment exercise. VQR assessment process is explicitly conceived for the evaluation of research institutions. VQR evaluation process is based on a hybrid approach, namely using bibliometric analysis for natural sciences (the so-called *bibliometric* areas), and peer review for humanities and social sciences (Franceschini, Maisano, 2017).

The primary goals of the implementation of VQR are:

1) providing the country with a rigorous and thorough snapshot of the research productivity of universities and publicly funded research institutes;
2) establishing a national ranking list of universities and university departments per disciplinary domain, to provide a rationale for the assignment of a large share of Ordinary Financing Fund (the FFO) and allocating the resources on a merit basis[51];

[50] For clarity purpose, when we say "VQR" we are referring to aspects and elements that generally pertain to all the VQRs (VQR1, VQR2, VQR3); as already mentioned in the previous paragraph, for discussing the eventual differences or similarities among the different VQRs, we are going to refer as VQR1 to the 2004-2020 VQR, as VQR2 to the 2011-2014 VQR, as VQR3 to the 2015-2019 VQR.

[51]

3) providing universities and departments with a snapshots of their research production, in order to understand how to allocate internal funds;
4) creating the preconditions for enabling a comparison on European and international level of Italian universities research production (Fondazione CRUI, 2015).

VQR is hence directed to public universities, legally recognised private universities, and public research institutes. The research products that can be submitted to the evaluation process are journal articles, book chapters, conference proceedings, critical reviews, commentaries, book translations, patents, samples, projects, plans, etc (ANVUR, 2011).

VQR overall involves three main actors: the evaluated-to-be institutes (universities and research institutes), ANVUR, and the GEV (*Gruppo di Esperti per la Valutazione*), namely the groups of experts appointed by ANVUR.

The VQR exercise is articulated in three major phases (Greco, Scarcello, 2013):

1) *the submission phase*, in which individual researchers are asked to submit their products published in the period specified by the call for notice; researchers upload their research products through the CINECA platform, ordering them in order of preference; research institutions should select the research products among the lists of products available; research institutions should also submit a report on some universities' information, regarding, for instance, the mobility of the personnel, the number of people attending advanced training (such as doctoral students, post-graduate school), and information about the institution's capability of attracting external funds;
2) *the evaluation phase*, in which GEVs play a fundamental role: each GEVs (with ANVUR) should define the evaluation criteria, select the peers, with which they manage (both directly and indirectly) the peer review of research products; GEVs draft the final report per disciplinary domain and turn it in to ANVUR;
3) ANVUR drafts and publishes the final report (*Relazione finale ANVUR*), basing on GEVs' reports, on the evaluation outcome per disciplinary domain, focussing on the results achieved by

research institutes and departments; ANVUR is also in charge of developing bibliometric indicators.

GEVs play a significant role in the VQR process since they are called to adapt the evaluation procedures to their specific discipline[52]. GEVs can chose whether to use, combined or singularly (where possible), the peer review (assigning research products to external independent referees) and bibliometric indicators. Peer review can also be requested for evaluating indexed articles that pertain to an emerging or multidisciplinary area. Bibliometric analysis of research products is used for articles indexed on Web of Science or Scopus and is based "on a quality score for each publication that derives from a combination of citations and journal metrics" (Abramo, D'Angelo, 2015, Abramo, D'Angelo 2016). Additionally, GEVs of bibliometric areas can freely decide the share of research products to be evaluated with bibliometric analysis, but at least the half plus one of the publications need to be evaluated with the peer review. The freedom left to GEVs is basically due to two issues: on the one hand, the necessity of reducing the costs of conducing a massive assessment exercise fully based on peer review; on the other hand, the need to comply with each disciplinary domain specificities, and to assure a fair evaluation process (Anfossi et al, 2016). During the evaluation process, GEVs can penalize research institutes for plagiarism, fraud, failure in the submission of research products and for the presence of inactive researchers (Wang, 2022, Franceschini, Maisano, 2017).

Since the introduction of the VQR1, some aspect of the setting of the evaluation exercise have been slighting changing, and we are going to briefly illustrate them henceforth. Indeed, a continuous reflection on VQR settings and methodologies has been widely solicited by ANVUR, with the purposes of, on the hand, to keep updated with technological advances and stakeholders' needs, and, on the other hand, to eventually adjust according to the suggestions by external experts.

[52] Even though, as we will better see afterwards, some scholars have criticized the GEV's interpretation of ANVUR directives (Fasanella, Martire, 2017, Baccini, 2014)

For instance, as can be seen in Table n. 2, the number of GEV areas and, consequently, the number of GEV experts, has increased over the three VQR editions. The GEV areas in bold in VQR2 and VQR3 represent the elements of novelty compared to the previous VQR cycles. The increase of the number of GEVs areas and experts was aligned to the logic of better adjusting the evaluation process to the specific publication and research patterns of each disciplinary area, and to emphasize the role of third mission (Wang, 2022).

VQR1		VQR2		VQR3	
GEV Areas	N.	GEV Areas	N.	GEV Areas	N.
1 Mathematics and Computer Sciences	25	1 Mathematics and Computer Sciences	22	1 Mathematics and Computer Sciences	29
2 Physics	18	2 Physics	33	2 Physics	43
3 Chemistry	23	3 Chemistry	22	3 Chemistry	31
4 Earth Sciences	9	4 Earth Sciences	15	4 Earth Sciences	19
5 Biology	38	5 Biology	33	5 Biology	49
6 Medicine	79	6 Medicine	58	6 Medicine	80
7 Agricultural and Veterinaly Sciences	24	7 Agricultural and Veterinaly Sciences	20	7 Agricultural and Veterinaly Sciences	36
8 Civi Engineering and Architecture	28	***8a Architecture***	14	8a Architecture	17
		8b Civil Engineering	9	8b Civil Engineering	15
9 Industrial and Information Engineering	40	9 Industrial and Information Engineering	33	9 Industrial and Information Engineering	59
10 Ancient History, Philology, Literature and Art History	42	10 Ancient History, Philology, Literature and Art History	36	10 Ancient History, Philology, Literature and Art History	56
11 History, Philosophy, Pedagoy and Psychology	38	***11a History, Philosophy, Pedagogy***	25	11a History, Philosophy, Pedagogy	29
		11b Psychology	6	11b Psychology	12
12 Law	37	12 Law	32	12 Law	39
13 Economics and Statistics	36	13 Economics and Statistics	31	***13a Economics and Statistics***	22
				13b Economics and Management	18
14 Political and Social Sciences	13	14 Political and Social Sciences	11	14 Political and Social Sciences	16
				Interdisciplinary: Impact and Third Mission	30
Total	450	Total	400	Total	600

Table 2: GEV areas per VQR cycle

Moreover, some features of the evaluation process have been modified as well. Firstly, the classes of merit for the evaluation of peer review have slightly changed (Table 3). Indeed, the VQR1 "Relevance" and "Originality/innovation" have been merged and substituted by "Originality" and "Methodological rigor" in VQR2 and VQR3; moreover, "Internationalization" has been substituted by "Impact" in VQR2 and VQR3. As a matter of fact, this issue was raised by the group of experts invited by ANVUR to provide some suggestions for improving VQR settings. According to the experts, the expression *potential or verified* might be very ambiguous, thus hindering a fair and as much objective as possible peer review process (Report of the Group of Expert charged by ANVUR to advice on the process "Valutazione della Qualità della Ricerca (VQR)", 2019).

VQR1	VQR2	VQR3
Relevance	*Originality*	*Originality*
Originality/Innovation	*Methodological rigor*	*Methodological rigor*
Internazionalization	*Potential or verified impact*	*Impact*
Transfer of technological development (for patents)		

Table 3: Classes of merit in the VQR cycles

Secondly, the scoring procedure has become less selective, if we compare VQR1 and VQ2. Indeed, the distance between products assigned to the highest scoring class (*Excellent*, which attributed 1 to the research product) and the lowest scoring class (*Not evaluable*, attributing -1 to the research product) was higher in VQR1, since in VQR2 and in VQR3 the point assigned to excellent products was 1 and to limited products 0. According to VQR1 call for notice, in case of fraud or plagiarism the penalty was of 2 points, while in VQR2 and VQR3 calls for notice this aspect is not specified. Still, according to VQR1 call for notice, in case of missing products with respect to the expected number of research products the institution must submit, the penalty for each missing product is -0.5, while in VQR2 and VQR3 research institutions will receive 0 point for each missing product.

Thirdly, the relevance of third mission in the evaluation process has increased in VQR3, in comparison to the previous VQR1 and VQR2 cycles, and this aspect is aligned with the suggestions of the group of experts. It is presumably for this reason that the VQR3 call for notice also established the creation of an interdisciplinary GEV, focussed on third mission, whilst in VQR2 the third mission was separately evaluated by ANVUR, according to the guidelines approved by the ANVUR Executive Board (ANVUR, 2015, Wang, 2022). Additionally, the group of experts encouraged ANVUR to urge universities to provide a "narrative" on universities' third mission-related initiatives.

Fourthly, VQR2 and VQR3 attributed more importance to research quality-related indicators (0.9) in comparison to the score assigned in VQR1 (0.5), which, indeed considered far more indicators[53], in

[53] *Staff Mobility 0.1*, *Internazionalization 0.1*, and *Self-owned resources 0.05*, for a total of seven indicators.

comparison to VQR2 and VQR3 (the IRAS1 indicator with score 0.75, plus the IRAS2 indicator with score 0.20, ANVUR, 2015).

And lastly, as we previously mentioned, the transition from VTR to VQR has determined an increased relevance of universities' research assessment on the allocation of resources from the FFO[54]. Interestingly, the outcome of VQR process have become even more important for the assignment of universities financing quotas in the subsequent VQR editions (Wang, 2022, Fondazione CRUI, 2015). Indeed, if on one hand since 2008 the total amount of resources for higher education has been considerably reducing, on the other hand, with Legislative Decree of 21st June 2013 n. 98, the Government established that the FFO share to be allocated according to the VQR outcome had to increase till 30% of FFO, with an annual increase of 2% (Fondazione CRUI).

The introduction of VQR has thus produced significant changes in Italian higher education system, and, for this reason, its implementation, setting and consequences have been at the centre of a lively debate among scholars.

In the first instance, the major effort of analysing such a huge number of research products and of collecting detailed information about research institutions' (analysing only during VQR1 almost 185,000 research products), has provided ANVUR and the government with the most precise portrait on Italian universities' productivity so far. Moreover, it has also increased the level of transparency over researchers' activities, opening the communication channels between research institutions and the community at large. For this reason, the implementation of VQR has significantly stimulated universities' efforts in reordering and improving managerial and accountability activities (Wang, 2022).

In the second instance, the introduction of the principle of competition among universities for resources allocation has been debated as well. Indeed, if on the one hand the competitive mechanisms may foster universities and researchers' performance to get access to as many resources as possible

[54] With VTR, the share of FFO that was allocated according to VTR outcomes was 20%.

(Ancaiani et a., 2015), on the other hand, the link between universities' research performances and resources allocation may induce a vicious circle, in which universities in less rich economic and social contexts might get a decreasing amount of resources over time, and thus making increasingly more difficult to succeed in future research assessment exercises cycles, deepening social and geographical disparities (Watermeyer, Olssen, 2016, Grisorio, Prota, 2020).

Moreover, some authors also criticize this mechanism for allocating resources, especially considering that some mechanisms of VQR could lead to biased results. Indeed, the fact that the selection of the research products to evaluate is not random but, as already analysed, is based on specific researchers and universities' choices, who rank select and rank the research products for the submission. For this reason, this mechanism may thus lead to an evaluation outcome that does not represent a fair estimate of the quality of researchers and universities' productivity but may be considerably influenced by researchers and universities' strategies during the process of selection of works to submit (Rossi, 2016). Moreover, Rossi underlines how the VQR setting does not consider the effects of not normalizing the huge variety of the dimensions of Italian universities.

Fourthly, some authors have discussed issues linked to both the role of GEVs and their composition. According to some authors, ANVUR should have explained more precisely the tasks GEVs should perform and how to manage the evaluation process. Indeed, with the purpose of letting GEVs partially free to develop an evaluation process which had to be suitable to their specific discipline necessity, ANVUR did not thoroughly specified the boundaries and features of GEVs' autonomy in this specific context (Fasanella, Martire, 2017). This produced a series of diversified situations in different disciplinary areas: for instance, if on the one hand GEV 14 seem to have just "copied and pasted" ANVUR's definitions of peer review classes of merit in their guidelines (Fasanella, Martire, 2017), on the other hand, GEV 14 adopted a too *sui generis* bibliometric analysis, namely conducting the evaluation process on a ranking list of journals that was drafted by the GEV 13 itself (Baccini, 2014). Additionally, the group of experts also recommended GEVs consistency of the evaluation procedure (since, in some cases, as we will better see afterwards, the evaluative mix has produced contradictory

results between peer review and bibliometric analysis), and of the citational databases for the bibliometric analysis (Group of experts, 2019).

The group of experts have also provided for some suggestions (raising the number of international experts, combining the election-appointment recruiting method, balancing the gender gap, fostering the transparency of the call for notice for hiring GEV experts) in order to foster the diversity and professionalism of GEVs composition, that, otherwise, could induce some homophilic behaviours (Baccini, 2014).

Fifthly, another key issue of the debate around the VQR setting is basically centred on the appropriateness and fairness of the evaluation methods, and the assessment indicator to adopt, which would be supposed to be supported by a relative strong scientific basis. Some concerns were raised indeed on the goodness of the adoption of the double evaluation procedure for the evaluation of research products pertaining to bibliometric disciplinary areas. For instance, some authors have noticed that the evaluation results obtained with the peer review and the bibliometric analysis do not match (Abramo et al, 2011), and that, considering the same GEV area, the evaluation outcome obtained with peer review are far more negative compared to the results obtained with the bibliometric analysis (Baccini, 2015, Baccini, De Nicolao, 2016). The authors hence reached the conclusion that ANVUR did not provide solid empirical evidence for the setting a fair evaluation procedure. Moreover, some authors also criticize ANVUR in the same way, for what it concerns the universities' performance indicators. Indeed, some authors underline that the assessment indicators proposed by ANVUR present some significant flaws, and that the experimentation phase of VQR procedure should not be considered closed, especially for what its outcome implies for universities' resources allocation (Abramo, D'Angelo, 2015).

2.2.2 ASN

The implementation of the National Scientific Habilitation in the Italian context (ASN, henceforth) complies with the logic of raising the level of academic recruitment and to subtract researchers' hiring

process to the nepotistic mechanisms that characterized to a certain extent the Italian higher education system in the previous decades (Marini, Meschitti, 2018, Abramo et a, 2014). Hence, established with Act of 2010, no. 240[55], the ASN defines the procedures with which researchers can accede to tenured positions, notably associate and full professors in Italian public and private universities. As we will better see afterwards, Act no. 240 introduces a recruiting system that is based on a committee of experts who evaluate the applicants through a selective examination which is supposed to evaluate applicants' research production quantity and quality (Musselin, 2004, Marini, 2017), before they get the possibility to accede to universities' national competitions (*concorsi*, henceforth).

The rationale of the ASN aims at overcoming the limits of the hiring mechanisms that were in place in the previous decades:

1) on the one hand, after the first long-lasting phase in which the universities' *baroni* had remarkably affected the academic recruitment process (Moscati, 2001)[56], the Minister of Education started a process of centralization of academic recruiting (Patti, 2019); indeed, with the Act of 1980 no. 382[57], the Minister of Education established two academic roles with tenure (*professori associati* e *professori ordinari*), for which it was necessary to win a national competition, in which the examination committees were composed by full professors appointed by the Minister of Education; the Minister of Education indeed made a considerable effort for counterbalancing the influence of *baroni*, by paying much attention to the process of appointing the members the evaluating committees that were supposed to evaluate applicants;

[55] Legge 30 Dicembre 2010, n. 240: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2010-12-30;240

[56] More specifically, the recruitment process before the 80s was substantially based on the direct appointment of young scholars by universities' full professors to all the several roles (*assistenti volontari*, *assistenti ordinari*, *professore incaricato*) that they had to cover before the tenure (Moscati, 2001).

[57] Decreto del Presidente della Repubblica dell' 11 Luglio 1980, n. 382: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.del.presidente.della.repubblica:1980-07-11;382!vig=

2) on the other hand, with Act of 1998 no. 210[58] the recruitment process is decentralized, and, universities themselves were supposed to hold *concorsi*; universities can then appoint full professors, associate professors and researchers.

Act no. 240 thus split and reorganizes the different disciplinary areas into "scientific disciplinary sector" (*settori scientifico-disciplinari*), by which researchers can position themselves to participate to national competitions. The applicants (both for the role of *associato* and *ordinario*) are evaluated by an examination Commmittee of 5 randomly selected full professors (who can work only for one Committee at a time), that are in charge for two years.

The ASN started in 2012 and its mechanism is articulated into two main phases: during the first phase, which is run by the Minister of Higher Education, the applicant's research products quality, and the impact of applicant's research production are evaluated by the members of the examination Committee (Patti, 2019); in the second phase, the university itself hires an associate or full professor by publishing a call for notice (the so-called *concorso*), for which are eligible the applicants who has previously gotten the National Scientific Habilitation[59]. For how ASN has been conceived, obtaining the scientific habilitation does not necessarily imply the appointment as associate or full professor in a specific university, and, furthermore, the habilitation lasts nine years. Until the year 2014 the applications could be submitted on a yearly basis, through the CINECA platform. Then, Act of 2014 no. 114[60] established that researchers can submit their application at any moment. If at the end of the evaluation process the applicant is ineligible for the habilitation, he should wait for one year before submitting another application (that is a sort of soft version of the "up-or-out" recruitment pattern, proposed by Musselin, 2004). Any applicant can apply for any scientific disciplinary sector, and both as associate and full professor. The evaluation process is conducted by the evaluation Committee

[58] Legge 3 Luglio 1998, n. 210: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1998-07-03;210
[59] As a matter of fact, the ASN process is not the only way to get the tenure. Another way, which is the direct calling (*chiamata diretta*) from the Minister of Higher Education of researchers who have determined prerequisites, according to Act of 2005, no. 230.
[60] Legge 18 Agosto 2014, n. 114: https://www.unioncamere.gov.it/sites/default/files/articoli/2022-03/Legge_114_2014.pdf

members, who are supposed to check whether applicants' research products fulfil the minimal thresholds, based on the medians of the total population. The thresholds are defined by ANVUR, as follows: for bibliometric sectors, the indicators are the number of articles, number of citations and H index; for non-bibliometric sectors the number of articles, number of articles in published in "A class" journal, and the number of books. However, the fulfilment of two indicators out of three does not necessarily imply that the applicant has automatically achieved the habilitation. The examination Committee also evaluates the applicant's resume and its research products published till the day of the submission.

As said before, the rationale of the introduction of the ASN mechanism for recruitment was to rationalize the academics' hiring process, by introducing the principles of evaluation and transparency, in order to partially tackle the *baroni*'s influence over universities. However, from the wide parliamentary debate (who involved also the scientific community) that led to the final draft of Act no. 240, a certain level of scepticism emerged on both its effectiveness and the rationale of the reform emerged.

On the one hand, some politicians maintained that the new recruitment process enabled with Act. No. 210 did not prevent the *baroni* from keeping a certain level of control and centralization over their afferent institutions, with the national *concorsi* being hold by the universities themselves and on which the *baroni* might exert a certain extent of power. Still, the universities based *concorsi* to hire researchers with the habilitation might also exacerbate the trend to provincialism that affects certain universities and disciplinary domains. Moreover, the speaker also maintained that the Government had to consider the to match the implementation of these reforms with a more careful planning of the resources to be invested in higher education, in order to avoid faults of the system, especially after the cut of the budget for university (*Commissione permanente in sede referente,16 Dicembre 2010*).

On the other hand, in another Committee session, some Members of the Parliament argued that this reform has represented a huge transformation for the academic recruiting process, and that the process would have deserved a more gradual approach in its application (*Commissione VII Cultura, 2010*).

Additionally, some Member of the Parliaments recommended caution for the introduction of the law, and to consider the possibility of implementing transitional arrangements, especially taking into account that with the ASN a large share of researchers with habilitation would have seen their habilitation expire without having been hired by any university (*Commissione VII Cultura, 2010*).

As a matter of fact, this was the case indeed of a large share of researchers with the habilitation: namely roughly the 40% of the entire population of researchers who got the habilitation (Patti, 2020). Even though the rationale of the introduction of ASN was to limit phenomenon like favouritism and to foster a transparent selection process, the data available so far and some recent facts have highlighted that the ASN system has some flaws. First of all, many researchers who had formerly applied for getting the habilitation, and that did not pass the examination, sued the examinations Committees and the Minister of Higher Education's choices. In mane case, the administrative court in charge of ensuring the transparency of national competition has challenged the examination Committees evaluations, for a series of reason: careless on the evaluation process of the applicants' research products, unfairness and flaws in the explanation of the rationale of the evaluation outcomes, and so on. Some recent studies have indeed highlighted that the selection system, for how it has been structured, is not efficient in preventing an homophilic behaviour from the examination Committees and in the *concorsi* at university level: researchers already working at universities as temporary researchers have more chances to get the habilitation, even with lower scientific performances, in comparison to applicants who do not work at universities, and, additionally, the number of researchers with the habilitation that are not hired by any university is supposed to increase, due to the lack of resources for recruitment at university level (Abramo et a., 2014, Abramo, D'Angelo, 2015).

To conclude, it is also worth adding that, if on the one hand ASN is accelerating the individual researchers' professional path, thus rewarding young researchers' scientific productivity (the medium age of researchers getting the tenure is decreasing, consolidating a long lasting trend, Rostan, 2011, Marini, 2016), on the other hand, the ASN mechanism has not been efficient yet in tackling the gender gap in the academic career advancement (Marini, Meschitti, 2018, De Paola et a., 2015).

2.3 Final remarks

According to the available literature and to the VQR calls for notice, we can thus affirm that Italy has endowed itself with a strong (with reference to the Whitley's conception of strong and weak research evaluation system), rather transparent and conducted on a regular basis research evaluation system. As also underlined by some authors, the *error-and-trial* approach, that is at the basis of the entire process of the introduction of the research evaluation culture in Italy, seems not to be over, since there are considerable differences among the three editions of the VQR.

As it will also emerges from our research and from the available literature, the introduction of VQR and ASN are much considered as a cutoff point in the Italian academic environment, even considering that they were both introduced in a period in which financial resources were already gradually decreasing.

Chapter 3
Research questions and methodological overview

In this chapter we are primarily going to present the research questions that are at the core of our research, and to illustrate the main methodological choices that we adopted in order to reply to our research questions and to test the related hypothesis we formulated. These concrete aspects are subsumed in the theoretical framework that let us operationally define the concept of research agenda setting. We also try to provide an explanation of the importance, in our opinion, of deepening a study on Italian researchers' research agenda setting, with the specific context of research evaluation.

3.1 The reasons for a study on research agenda setting

The emphasis on researchers' performance and productivity determined by the introduction of research evaluation instruments has induced many scholars to study the effects on researchers' research practices. As already discussed in the first chapter, the pressure to publish that research institutes exert on researchers (the so-called *publish or perish culture*) seems to have induced the same researchers to assume some adaptive behaviours. To recall them quickly, the most relevant strategic behaviours identified are those aimed at inflating the number of citations and publications, such as strategic citing (Frey et al, 2009), and *salami tactics*, with an increasing number of low level of impact research products (Butler, 2003), scientific collaborations and co-authorship (Colarusso, Giancola, 2020, Van Den Besselaar et al, 2012); some studies have shown that research evaluation policies seem to have widely encouraged the use of English language, and influenced the choice of determine outlets for publication (Hammarfelt and Rijcke, 2015), in order to get recognition by the reference scientific community; furthermore, the spillover of the overemphasis of research activities seems to have determined the decreasing importance of activities not directly linked to research, like teaching activities or administrative duties (De Philippis, 2015, Campbell, 2013),

The drive to publish seems also to be at the basis of an increasing sense of powerlessness and distrust among researchers (Watermeyer, Olssen, 2016, Burrows, 2012, Moss and Kubacki, 2007), worsened by the feeling of instability that is linked to a professional path that is apparently no longer linear, and with blurred tasks and features (Cantwell, 2011, Shattock, 2014, Musselin, 2014).

Indeed, these reforms are to be considered in the evolving specific working framework in which academics find themselves: the academic profession is indeed experiencing a long-lasting period of transformation of its specific tasks and features. More specifically, the relationship between the academia and society at large has been shifting from the *ivory tower* model, in which researchers were supposed to be a self-regulated body, with a relatively large extent of power in the university context, who conducted research in an autonomous and disinterested way, to a scenario in which research activities should be a common good for society, hence the various social actors are supposed to get involved in the knowledge production process (Henkel, 2005). As a result of this changed conception of the relationship between the academia and society at large, academic profession underwent a general process of restructuring of its functions and tasks in the European countries (Shattock, 2014, Musselin, 2005), notwithstanding the specificities of each country (Musselin, 2004). First of all, the scholars remarkably grew in number, as a result of the massification of university access, and relatively lost its elitist connotation. Alongside this fact, the academic profession has also seen a transformation of its power and is relatively less autonomous than previously (Musselin, 2005), since researchers were supposed to involve while conducting research activities other social actors, with which collaborating (Henkel, 2005). Additionally, since the increased number of tasks researchers are called to perform, their power is somehow fragmented across the various activities, evaluation boards and research councils they are supposed to attend (Musselin, 2014). At last, as a result of the "new managerialism" (Musselin, 2005), the control over academics' performance is not confined to the competitive recruitment process, but is extended throughout all the researchers' career path.

In this highly competitive and changeful academic context, we think that it would be interesting to investigate whether the introduction of research evaluation policies (REPs, henceforth) have somehow influenced the way in which researchers select the topics on which conducting their research on. Our interest in deepening the issue of the influence of REPs on how researchers select their research topics rests on two main reasons.

First, the process of selecting the topics to research is, as a matter of fact, a key process and factor in a researcher's professional life. Indeed, the activity of selecting the topics to research is fundamental in researchers' life (particularly for young scholars), since it enables the creation of personal and professional connections which enhance their own credibility cycle (Latour, Woolgar, 1979).

According to the authors, indeed, the attention paid by a researcher when selecting the topics on which to conduct his research activity is essential for his career, because with this process the researcher collocates himself in a specific spot of the scientific arena. As a matter of fact, the researcher can be rewarded by his peers and by the "market of information", if they are persuaded that the specific issue on which he conducted research is a fact and meaningful. The success of the researcher's career remarkably relies on the acceptance of its proposal, which constitutes the reward to his work in the market of information, increasing consequently his credibility as scientist in the scientific arena. The authors then describe the credibility as cumulative (for this reason, defined as *credibility cycle*), since the more a researcher get peers recognition, the more he will receive proposals for work in the future by the actors of the market of information, who need to invest in somebody who can in all likelihood make their investments profitable.

Secondly, we deem that an understanding of the eventual effects of the implementation of REPs on the process with which researchers select the topics on which to conduct research would be very interesting to further a thorough comprehension of the overall effects of REPs on researchers' behaviours.

As a matter of fact, if on the on hand many scholars have tried to understand the effects of the implementation of REPs on several researchers' research practices (as already analysed in the first

chapter), on the other, by considering the available literature, it is possible to notice that few important works have attempted to directly address the issue of the effects of REPs on researchers' selection process of research topics. As a matter of fact, Whitley (2007) has proposed a theoretical framework for describing the different levels of influence on research contents, according to the discipline features the researchers are afferent to. More specifically, the author maintains that there are several variables to consider when trying to understand the extent of influence on researchers' research contents. Firstly, a strong research evaluation system is likely to affect more researchers' topics to study, while, on the other hand, in weak research evaluation systems researchers are less likely affected by research evaluation exercises. More specifically, in strong research evaluation systems, researchers tend to adapt their research topics to the perceived necessity of their afferent research institutions, as for the high level of competition triggered by research evaluation; the high level of standardisation of procedure determined by the introduction of research evaluation tends also to standardize research topics across the different disciplines, according to the preferences of the academic elite that is in charge of conducting research products evaluation. Another important variable that is worth considering, according to the author, is the way in which the afferent research institution receives the research fundings: indeed, where most of research funding is allocated according to a block grant mechanism, the influence on researchers' selection of topics is weaker, while, on the other hand, when fundings are linked to research evaluation performance, the research content is generally more affected. Still, the author argues that researchers afferent to disciplines that can easily get fundings from external institutions tend to be more independent when deciding which topics to study (like, for instance researchers from hard sciences); on the other hand, social scientists tend to be more likely affected, as for their difficulty in accessing fundings from external institutions. Whitley, moreover, listed a series of other variables that may suggest that in those cases an influence on research content tends to be stronger: disciplines that are more coordinated nationally and internationally; the disciplines or university in which the academic elites are called to make judgements for contemporary issues; disciplines with high internal consistency and coordination.

Even though Whitley's theoretical framework is useful as analytic grid to understand the possibly researchers and universities' behavioural patterns, it is important to underline that it is not fully supported by an empirical basis (Gläser, Laudel, 2016).

As a matter of fact, studying the effects of the implementation of REPs on researchers' selection of research topics is actually a challenging task, for three main reasons (Gläser, Laudel, 2016). First of all, since in the same concept of REPs, which is very complex itself[61], as seen in the first chapter, many elements intertwine, a large-scale study or a comparison between different countries may be very tricky. Secondly, the fact that it is very difficult to isolate the research evaluation mechanism from other elements which are usually combined with (Gläser, Laudel, 2016). Some studies have indeed analysed many aspects connected to REPs (such as, for example, the effects of funding allocation on a merit basis[62], or the creation of a "negotiated space" created by the changed relations among academia and other external actors, in which the research contents may be shaped and discussed[63]) and its effects on research contents, but a clear relation between REPs implementation and research content has not been studied yet. Finally, Gläser and Laudel (2016) maintain that the lack of studies on this topic may also be explained by the impossibility of defining an impact of the implementation of REPs (since it has been impossible so far to compare the scenario in which there is research evaluation with its counterfactual), thus discouraging scholars to investigate the problem.

3.1.1 Research agenda setting, an operational definition

The concept of agenda setting was first introduced in the communication science field at the end of the '60s, to describe the role of mass media in framing the audience's political ideas[64]. Subsequent studies theorized that the agenda setting outcomes are the result of a dialectic process among the

[61] In the first chapter we speak about three main elements which make the comparison of REPs very difficult: the high level of variation of the same definitions of the policies among different countries, the scope and the complexity of the policy itself, the interrelation among the various actors at stake.
[62] Laudel, 2016.
[63] Henkel, 2005, Luukkonen, Thomas, 2016.
[64] The first research on mass media's agenda setting process and its consequences on voters' political opinion was conducted in 1968, by McCombs and Shaw (McCombs and Shaw, 1972).

several actors at issue: the way in which topics are framed is the outcome of a political and economic bargaining process among the political elites, the mass media and the economic stakeholders (Entman, 1989, Reese and Danielian, 1989).

The academic context has recently borrowed this concept from communication science, and, nowadays, many connotations can be attributed to the concept of research agenda setting (for instance, in some context it is the list of topics a researcher commits himself to study in the upcoming years, that has to be submitted to his afferent institution, the series of topics that researchers and private enterprises discuss about which will be the common ground on which to build a professional relation).

The first attempt to propose a thorough definition of research agenda setting was provided by Ertmer and Glazewski (2014), highlighting the essential role the process recovers in order for PhD students and young scholars to enhance their own academic career. According to the authors, the research agenda setting process is the sequence of short and long-term actions that a researcher takes to address his specific interests and work, following the framework that he decides to adopt to address a specific topic. Interestingly, Ertmer and Glazewski's definition of research agenda setting identifies its twofold nature: namely, the internal element of the research agenda, that reflect the researcher's "purpose and focus" (Ertmer, Glazewski, 2014), and the external component, that is related to the features and culture of the reference community of peers. From the combination of the internal and external aspects, the researcher makes a series of choices aimed at building his work and reputation within the academic community, by deciding which projects to adopt or turn down.

The choices that compose the researchers' research agenda setting are:

1) identifying his own research focus (that is, establishing the motivation of research and the audience the research is addressed to),
2) his own research interests,
3) locating his scholarly community, and, at last,

4) identifying his own research context (namely, establishing his own practices by the meeting with other researchers' practices).

The study of the factors that can contribute to the formulation of the researcher's research agenda setting is a rather challenging issue, since the available literature on this topic has shown that there are several determinants at stake that contribute to the definition of the process. Indeed, even though the researcher's research agenda is the result of his personal choices and knowledge (Polanyi, 2012), it is possible to identify two major groups of factors that contribute to the research agenda setting process: on the one hand, the group of factors referring to the researchers' personal aspects; on the other hand, the factors attributable to the working and social environment the researcher is embedded in (Santos, 2020). However, it is worth noting, as we will see afterwards, that an excessive clear-cut distinction between social and personal factors can be misleading, since some of the factors (as, for instance, peer recognition, the role of mentor, the passion for research) can be assigned to both the groups and denying the complexity of these factors would preclude a thorough comprehension of the research agenda setting process (Santos, 2020). For this reason, we will consider this distinction as a theoretical framework with the scope of clustering and streamlining the comprehension of the several factors at stake.

First of all, the conception of the purpose of academic profession and the researcher's interests, the researcher's study path and the role of the mentor can be ascribed to the personal factors that contribute to the formulation researchers' research agenda.

First, researchers' research agenda setting can be influenced by their perception of the meaning and purpose of being a scholar and conducting research activities (Åkerlind, 2008). According to the author, it is possible to identify four basic conceptions of being a researcher ("fulfilling academic requirements, with research experienced as an academic duty; establishing oneself in the field, with research experienced as a personal achievement; developing oneself personally, with research experienced as a route to personal understanding; and enabling broader change, with research

experienced as an impetus for change to benefit a larger community"). Academics belonging to the second, the third and the fourth category are more likely to carefully select the topics to study, as for the noteworthy involvement with the academic profession and the public utility of their job. Hence, for the academics of the third and fourth group, the passion for research and benefiting society at large are major determinants of their research agenda setting.

Moreover, the researcher's study background seems to play a relevant role in shaping research agenda setting. Indeed, the study path that the researcher conducted when he was a student (Karvalics, 2013, Pinheiro et al, 2014) and the benefits of getting specialized in a specific field in terms of expertise and peer recognition (Leahey, 2007, Latour, Woolgar, 1979), seem to address researcher's future research agenda setting quite often. Furthermore, the role of the researcher's mentor also plays an important role in shaping the researcher's research agenda, due to the relevant role that the researcher's mentor plays in the academic hierarchy (Bourdier, 1999).

Secondly, the factors that can be ascribed to the social environment are scientific collaborations, the recognition of the peers, the pressure of the afferent research institutions and stakeholders, and the availability of research funding.

The social environment seems to play a relevant role in shaping researchers' research agenda. Scientific collaborations with both fellows and non-researchers, and the research topics that may let researchers achieve good reputation (like, for instance, the mainstream topics), seem to contribute remarkably to researchers' agenda setting process (Santos, 2020, Latour, Woolgar, 1979, Bourdieu, 1999, McGinn, 2012).

Still, the changed relationship between universities, central state and external stakeholders has enables the creation of a "negotiated space" (Luukkonen, Thomas, 2016), in which universities and researchers are called in some cases to adjust their research and teaching activities to meet external stakeholders' needs. Therefore, in this specific context the research agenda setting process is the result of the bargaining process among researchers and the other actors' needs and ideas at stake (Luukkonen, Thomas, 2016, Henkel, 2005, Leisyte et al, 2008). According to the available studies on

this topic, in this bargaining process researchers generally attempt to keep the decision-making autonomy, but this is not always the case (Luukkonen, Thomas, 2016, Leisyte et al, 2008).
Moreover, the introduction of the rationalization of resources and accountability mechanisms in the universities-central state relationship have determined a certain level of pressure and stress on universities' activities and performances. For this reason, some authors have highlighted that universities may exert a pressure on the researchers' research agenda setting process, to ensure the research institution visibility and prestige (McGill and Settle, 2012), to comply with the benchmark established by the governing bodies (Shore and Wright, 2004), and to contributing to the overall national knowledge production and academic expertise (Vessuri, 2008). Additionally, a study conducted by Laudel (2006) showed how the availability of research funds can influence the content of researchers' research agenda.

A leading attempt to understand the basic determinants of researchers' agenda setting process is Horta and Santos' works (2016, 2018), with the creation of the multi-dimensional research agendas inventory. The multi-dimensional research agendas inventory (then proposed in a revised version, Santos, Horta, 2020), primarily tested on social scientists and then validated to researchers from all the disciplines, comprises eight dimensions and ten subdimensions (Santos et al, 2019), whose objective was to identify the endogenous factors underpinning the researchers' agenda setting attitude. The MDRAI dimensions are: *scientific ambition* (divided in two subdimensions *prestige* and *drive to publish*), which identifies the researchers' personal desire to get prestige and recognition by their peers; *convergence* (whose subdimensions are *mastery* and *stability*), which identifies the tendency to focus on one or few disciplines, *divergence* (whose subdimensions are *branching out* and *multidisciplinarity*), namely standing for a multidisciplinary approach, *discovery* and *conservative*, the former singling out a preference for emerging fields, meanwhile the latter a preference for carrying on well-established topics, *tolerance for low funding*, identifying the researchers' availability of

conducting research even in low economic resources conditions, *collaboration*, divided in the subdimensions *willing to collaborate* and *invited to collaborate*, and *mentor influence*.

The MDRAI model has enabled the identification of two main research agenda setting approaches (not mutually exclusive, and often coexisting), notably the cohesive research agenda and the trailblazing research agenda. Cohesive research agenda identifies a less-risky approach, opting for getting specialized on one or few specific fields, and gaining a remarkable expertise by a long-term effort. Cohesive agendas researchers generally show a lower degree of collaboration, also explained by the long-lasting influence of their mentor.

Trailblazing research agendas generally identifies a more exploring, risky, and multidisciplinary approach. Researchers who may be ascribed in the trailblazing agenda generally tend to conduct research selecting new approaches and topics, frequently seeking collaborations with their peers.

The MDRAI model, also implemented in its revised version (with the subdimensions *academia driven* and *society driven* instead of *convergence* and *conservative*), has been used in subsequent studies to further investigate research agenda setting processes.

Santos et al (2020) conducted a study to understand whether there may be gender based differences in research agenda approach. According to the study, women generally have a more conservative, less risky and more collaborative approach, and, moreover, for women it takes longer to reduce a mentor's influence on their own research activities. Men, instead, show a trend toward selecting more risky and disruptive research topics.

Interestingly, the association between researchers' conceptions of research (Santos, Horta, 2020) and thinking styles (Santos et al, 2020) has been studied with the MDRAI model. The authors have associated the eight dimensions model with the CoRI model, developed by which Meyer, Shanahan and Laugksch (200), listing the five possible conceptions of research: *research as the discovery of the truth*; *research as an insightful process*; *research viewed as a re-research*; *finding solutions to problems*; *misconceptions about research* (this model revokes Akerlind's study on the impact of the four possible conceptions of research, 2008). The findings of the study highlight the existence of

somewhat associations with five dimensions of the MDRAI (*scientific ambition*, *collaboration*, *discovery*, *academia-driven approach*, and *society-driven approach*): more precisely, the study identifies a positive correlation between a Type I thinking style (which is more creativity-generating) and the scientific ambition, multidisciplinarity, and collaboration dimensions. On the other hand, the Type II thinking style is associated with well-established topics, since this thinking style refers to a "norm-favouring tendency" (Santos et al, 2020), which generally denotes scarce creativity and initiative.

Furthermore, Horta and Santos' contribution to the understanding of research agenda setting mechanisms includes a study conducted to explore whether there is association between research agenda attitudes and researchers' propensity to adapt to organisational directives due to the implementation of top-down managerialist policies implemented at university level. According to the authors, the findings show that "organisational characteristics are associated with and can influence the research agendas of academics". More specifically, the authors maintain that leaving researchers freedom encourages the drafting of a more multidisciplinary and disruptive agenda. In addition, the freer researchers are free, the more ambitious their research agendas are and the more productive researchers are. By contrast, the implementation of top-down new public management inspired policies may constrain researchers' tendency towards conducting disruptive agendas and being active in gaining peers recognition, adopting a safer research agenda approach.

Hence, considering Ertmer and Glazewski's definition and the several factors that may concur to the shaping of the research agenda as starting point, we can hence propose an operative definition of the research agenda setting: we define the researcher's research agenda setting as the short and medium-terms actions with which the researcher selects the topics to research and the framework to address them. The elements that compose the research agenda are thus, as previously mentioned: 1) the selection of the topics to research; 2) the motivation for researching; 3) the selection of the research practices; 4) the selection of the peers community to address. Moreover, the factors that can influence

the research agenda setting process are scientific collaborations, the availability of funds, peer recognition, the state of the art and the mainstream topics.

3.2 Research questions, methodological choices, and hypotheses

3.2.1 Our research questions

Considering the relevant changes that the introduction of VQR and ASN has determined in the Italian academic context, and the pressure that research institutions have been increasingly exerting over researchers' activities, the purpose of our research is then to understand whether the introduction of VQR and ASN has changed or affected somehow the way in which Italian researchers structure their own research agenda setting.

Therefore, the research questions we aim at answering are:

1. Has researchers' research agenda setting been influenced by the introduction of research evaluation policies, namely VQR?
2. Are there differences between recruitment-oriented research evaluation policies (ASN) and universities-oriented research evaluation policies (VQR) in terms of effects on the researchers' research agenda setting process?
3. Is it possible to identify different effects across the various scientific discipline?

After reviewing the factors and determinants of research agenda setting from the available literature, and the main features of the implementation of VQR and ASN in the Italian context, we deem that it is possible to make some hypotheses with regard to our three research questions:

1. According to the available literature (Whitley, 2007), it is reasonable to say that we do expect that from our results it will emerge a certain change of researchers' research agenda setting, since social sciences tend to be more sensitive to research evaluation policies in strong research evaluation systems;
2. We assume that the ASN influence on research agenda setting process may be stronger in comparison to VQR, since the achievement of the scientific habilitation directly affects

researchers, as we have seen in chapter 2, while, on the other hand, VQR is not addressed to researchers' performance but rather to universities' performance;

3. We assume that the influence of ASN and VQR may be stronger in those disciplines that give value to empirical evidence (such as statistics, business economics, economics, applied sociology) rather than on those disciplines who rather give value to theoretical argumentation, as sociology, political analysis and political history (Bonaccorsi, 2015, Lamonte, 2009).

3.2.2 Methodology: the reasons for a mixed method approach

For the scope of our research, we opted for a mixed-method approach, namely combining a first phase of the research, in which a self-administered webmail survey was submitted to the target population, and a second phase, in which we conducted qualitative interviews with experts of research evaluation. The purpose of setting a mixed-method approach is twofold:

1) on the one hand, with the submission of a webmail survey to the target population, we aim at grasping the highest number possible of information, and answering the research questions we introduced in the previous paragraph;
2) on the other hand, conducting qualitative interviews may let emerge new topics and patterns of interpretation, due to the openness of the method used, since the topic at issue is rather unexplored and open (Creswell, 2014).

The mixed-method procedure that we aimed to adopt thus allowed us to add to the prior knowledge gained during the first (quantitative) phase, to answer to the research questions previously presented, the further information collected through the subsequent (qualitative) phase, to provide and corroborate a thorough overview of the various aspects connected to the introduction of research evaluation in Italy.

By referring to the mixed-method classification proposed by Creswell, Plano Clark, Gutmann and Hanson (2003), who distinguished between two major categories (sequential and concurrent mixed-

method design[65]), for the purpose of our research we thus aim to adopt as the concurrent triangulation design (Creswell, 2013), as illustrated by the figure below (borrowed from Creswell, 2013):

Figure 1: Concurrent triangulation research design

With the adoption of the concurrent triangulation design, the information collected and the analysis outcomes from both quantitative and qualitative approach aim at confirming or corroborating the findings within a study. Hence, there is not a logic or sampling-based sequence between the two approaches. It is important to specify that in our work we generally refer to the *first* and *second* phase to describe the chronological order with which the research was conducted (as it may happen in concurrent triangulation design), and not a logical succession between the two phases and approaches (Creswell, 2013). Hence, the second analytical phase does not represent a deepening of the features of the sample studied in the first phase. Indeed, the only link between the first and the second phase samples is that the interviewees were selected among the survey respondents who proposed themselves (through a dedicated tool bar) on a voluntary basis for the second research phases. Furthermore, it is worth specifying that the only conceptual link between the first and the second phase is that we used the most relevant information extracted from the survey free comments to build up a sketch of the set of questions to submit afterwards to the experts.

[65] According to the authors, in sequential designs, either qualitative or qualitative data are collected in a previous stage, followed by the collection of the other type of data not collected in the first stage; in concurrent mixed method designs, both types of data are collected at the same stage.

A full description of the methodology adopted follows. During the first phase, the survey was submitted to the target population, namely composed by the all sociologists, economists and political analysts currently working in both private and public Italian universities, as full or associate professor, temporary and fixed-term researchers. The choice of the self-administered webmail survey was deemed appropriate for a series of reasons. First, the studies conducted so far show that this topic should be addressed on a micro-level, namely directly reaching out the respondents, and conducting an analysis on the collection of secondary data (for instance, citations and publications) would not have been suitable for studying such kind of phenomenon, which is necessarily intertwined with other evaluation-related mechanisms and elements, to infer any causal relation (Glaser, Laudel, 2016). Secondly, the use of the survey is deemed appropriate for the possibility it provides to directly observe and collect as much information as possible over researchers' opinion, considering the relative high number of units of analysis our target population is composed of (Babbie, 2020). Thirdly, the inherent features of the target population suggests that the webmail survey is suitable to easily reach out to the largest possible number of respondents, by using their institutional email address (Corbetta, 2014). Indeed, academics generally access daily their institutional address for working reasons, especially after the massive shift to smart working due to the Covid 19 pandemic outbreak (Favale et al, 2020). Fourthly, as previously mentioned, if it is not feasible to estimate the impact of the introduction of research evaluation policies on researchers' research agenda setting, since the citational databases currently used (that is, Wos, Scopus and Google Scholar) in social sciences do not guarantee a good coverage of scientific works, and thus there is not the basis for estimating a counterfactual scenario to estimate the impact, on the other hand a webmail survey conducted on a remarkably high number of people, may provide us with a thorough overview of general Italian researchers' opinion on this specific matter. Fifthly, the use of the web survey is deemed the most appropriate method since it grants a good geographical coverage (Lombi, 2015) of the target population we address (our target population is distributed over almost all the Italian territory), and, still, because it may reduce social

desirability biases linked to the discussion of sensitive topics for the interviewees (Lombi, 2015, Tourangeau, Yan, 2007).

During the second phase, the qualitative interviews were addressed to the survey respondents who offered their availability for a further deepening of the topics addressed with the survey. Among the respondents who offered their availability, we chose the respondents who had proven expertise as both researchers and research evaluators (both at national and at university level). We opted for interviewing researchers with evaluation experience to benefit from their peculiar point of view (Laudel, 2006, Laudel, Gläser, 2007, Morawska, 2018), in order to grasp a thorough overview of the determinants of research agenda setting. The set of questions on which structuring the qualitative interviews were based on the most relevant topics emerged from the survey open-ended questions, as we will see in the following chapters. Both qualitative interviews and the open comments were conceived as an instrument to let researchers freely express on a relatively new topic, and to grasp insights for future research and reflections.

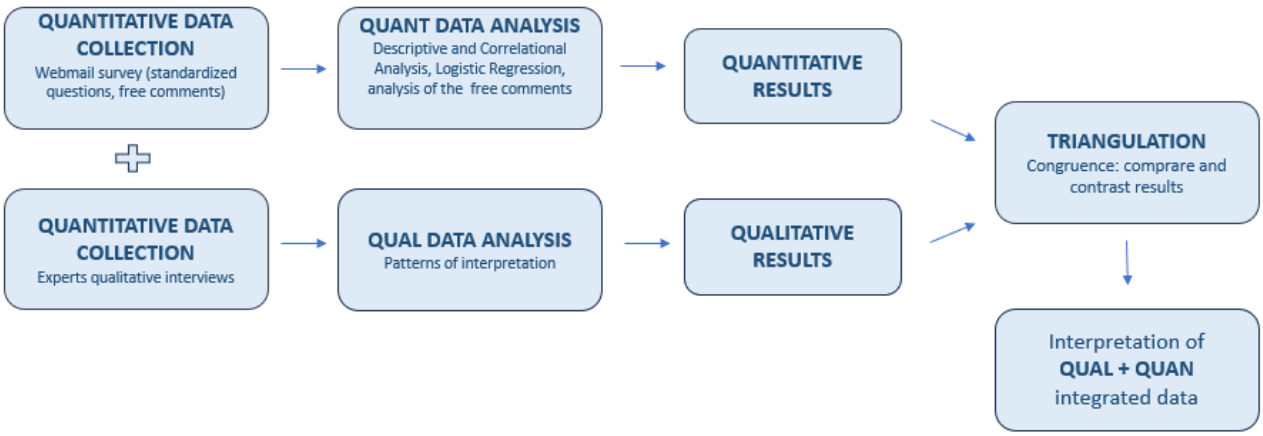


Figure 2: Figure from Vaughan Dickson et all, 2011

With Figure 2, borrowed from Vaughan Dickson et all (2011), we attempt to summarize the methodological process that we have described previously. With the integration of the interpretation of both quantitative and qualitative data, we thus aim to provide a thorough overview of the Italian

researchers' opinion on the effects of the implementation of VQR and ASN, with a specific focus on whether the "epistemological curve" of the Italian social scientists' community has been somehow affected, by influencing researchers' research agenda setting process. More specifically, the quantitative (and, chronologically speaking, first) phase is focussed on answering the three research questions listed previously (whether VQR and ASN have affected researchers' research agenda setting, and to eventually identify differences among disciplines). Besides, the qualitative phase is mainly focussed on understanding how researchers, who played an active role in the evaluation process, judge the Italian evaluation system, according to the most relevant topics emerged from the survey free comments.

3.2.3 Survey analytical dimensions

The questionnaire is composed of three main parts: the first section asks respondents for general information on their career status and other personal information; the second part is an extract of the MDRAI-R model, and the third part is dedicated to the understanding of the researchers' attitudes toward research agenda and research activities, based on the research agenda setting operational definition we have proposed above.

In the first part of the survey, we asked personal and career-related questions: we asked for researchers' afferent university, the dimension of their afferent university, the role currently filled in university, the length of experience in academia, gender, nationality, and marital status.

The central part of the survey has been structured by extracting from the MDRAI-R model the dimensions which mainly let distinguish the cohesive and trailblazing archetypes, according to the dimensions scores collected in Horta and Santos previous studies (Horta, Santos, 2018, Horta, Santos, 2020). The objective of this section was to identify which cluster may researchers belong to, since the cohesive research agenda approach would be associated to a more organisational compliant approach, rather than a riskier and discovery-oriented approach (Horta and Santos, 2020). The MDRAI and MDRAI-R dimensions that are deemed relevant for our work are *Convergence* (*Mastery*

and *Stability*), *Divergence* (*Branching out* and *Multidisciplinarity*[66]), *Tolerance for low funding*, and *Academia driven*. The questions are structured in a Likert scale, without the medium response item, to stimulate respondents to take a stand. The statements are ordered in a random way to avoid the response-set bias effect, with the poles of the Likert response categories alternated (Babbie, 2020). The *Collaboration* dimension, even though it contributes to identifying the trailblazing agenda, has not been included in this section, but rather in the subsequent one. The issue of collaboration among social scientists needed indeed to be separately analysed (in the third part of the survey, indeed), considering the specific Italian academic context where the survey has been submitted, since the study conducted by Horta and Santos does not take into consideration the countries-specific academic features (the study has involved 923 researchers from more than 65 countries). The survey was then structured to study the effects of both collaboration with peers from the same field, and collaboration with colleagues from other disciplines on the research agenda setting process, since the trend of co-authorship in the social sciences has consistently increased (Colarusso, Giancola, 2020, Bellotti et al, 2016).

Furthermore, the third part comprises the questions aimed at understanding respondents' opinions on the effects of potential determinants of the research agenda setting process (according to the operational definition provided), namely: collaboration, peers recognition, the role of mainstream topics, the role of current events, beneficial role of research activity on society, the availability of fundings for international or national competitive calls, the introduction of ASN, the introduction of VQR. The ASN and VQR questions asked the respondents to specify in which aspect of their research agenda setting process the introduction of these policies had influenced, notably, the choice of the topics to research, the motivation in conducting research, the choice of the research practices and the choice of the reference community of peers.

[66] The concept of *Multidisciplinarity* is to be conceived in the distinction of theory driven and application driven research (Bruce et al., 2004), in which multidisciplinary research in which disciplines "are juxtaposed with little cross-fertilization" (Bellotti et al, 2016).

As already mentioned, in the survey there were many open-ended questions, where the respondents could freely deepen their opinions. The open questions were added after the questions referring to the possible research agenda setting determinants, and one more at the end of the survey, to ask for general comments.

Some closed-ended questions were accompanied by descriptions or definitions associated with the construct we aim at understanding, since, for the newness of the topic, some more explanations were required to ensure the validity of the questions (Schlipphak, Isani, 2020). Some time frame references were inserted in the explanations (also specifying the meaning of research agenda setting itself, which is quite a new and blurred concept), to enable the process of the retrieval of information (Groves et al, 2011), since the interval time respondents are required to think of is quite extended over years.

The following table represents the operationalization of the concepts we wanted to explore, with the wording proposed.

Research agenda setting determinants	**Operation definition**	**Type of question**	**Measure**
Collaboration	*Indicate your opinion on this sentence: "I often search for colleagues with who I can collaborate on my publications"*	Likert scale	Quasi-cardinal
Collaboration	*How do you think that collaboration with peers from your disciplinary field has influences your research agenda setting?*	Closed-ended question	Quasi-cardinal
Collaboration	*How do you think that collaboration with peers from other disciplinary fields has influences your research agenda setting?*	Closed-ended question	Quasi-cardinal
Peers' recognition	*I have frequently oriented my research agenda to gain social recognition from my peers*	Likert scale	Quasi-cardinal
Mainstream topics	*Do you think that conducting research on a mainstream topic increases your chances that your work will be published?*	Likert scale	Quasi-cardinal
Current event	*Do you think that current events are relevant when programming your research agenda?*	Likert scale	Quasi-cardinal
Research as benefit for society	*Do you think that societal needs are relevant when programming your research agenda?*	Likert scale	Quasi-cardinal
Effects of attendance to calls for national or international calls	*Do you think that attending competitive calls has influenced your research agenda?*	Closed-ended	Nominal
Effects of attendance to calls for national or international calls	*How do you think that attending competitive calls has influenced you research agenda?*	Closed-ended	Nominal

ASN	*My research agenda has been influenced by the introduction of ASN*	Closed-ended	Dichotomic
ASN	*Has the introduction of ASN influenced the choice of topics to research?*	Likert scale	Quasi-cardinal
ASN	*Has the introduction of ASN influenced your motivation for researching?*	Likert scale	Quasi-cardinal
ASN	*Has the introduction of ASN influenced your research practices?*	Likert scale	Quasi-cardinal
ASN	*Has the introduction of ASN let you address a different peers community?*	Likert scale	Quasi-cardinal
VQR	*Has your research agenda been influenced by the VQR*	Closed-ended	Dichotomic
VQR	*Has the introduction of VQR influenced the choice of topics to research?*	Likert scale	Quasi-cardinal
VQR	*Has the introduction of VQR influenced your motivation for researching?*	Likert scale	Quasi-cardinal
VQR	*Has the introduction of VQR influenced your research practices?*	Likert scale	Quasi-cardinal
VQR	*Has the introduction of VQR let you address a different peers community?*	Likert scale	Quasi-cardinal

Table 4: Operationalization of determinants of research agenda setting

3.2.4 Target population

The target population is composed of 6297 individuals[67], comprising full professors (*professore ordinario*), associate professors (*professore associato*) and permanent researchers (*ricercatori a tempo indeterminato*), and temporary researchers (*ricercatori a tempo determinato*), currently working for both public and private Italian universities (the total number of universities is 99). The lists of academics are easily accessible on the Ministry of Education and Research website (CINECA). It is possible to download the lists from the CINECA website by selecting the 13/A (Economics), 13/B (Business Economics), 13/C (Economic History), 13/D (Statistics), 14/A (Political Theory), 14/B (Political History), 14/C (Sociology), and 14/D (Applied Sociology). We then merged the different lists of academics.

We decided to select as target population the sociologists, economists and political analysts in the first place because social scientists seem to be a particularly interesting population for our research

[67] As of the data collected in October 2021.

on the effects of research evaluation policies on researcher's research agenda setting, since they could be more sensitive to the implementation of research evaluation policies, in comparison with scholars from natural sciences (Whitley, 2007). Still, it was also interesting considering that economists, sociologists and political analysts do have similar research practices, in comparison to scholars afferent to humanities.

Moreover, since the writer's disciplinary field is sociology, investigating the influence of research evaluation policies on social scientists' research agenda setting is of remarkable scientific value in our opinion to understand the scientific practices of our afferent peers community.

As it is apparent from the map below (Fig. 3), the three Italian regions with the highest number of research are Lombardy, Lazio and Campania. More than a North-South divide (that emerges to a certain extent), it is possible to see that the regions with the biggest cities are also the regions with the highest frequencies of researchers. In these regions, there are indeed several universities, both private and public ones.

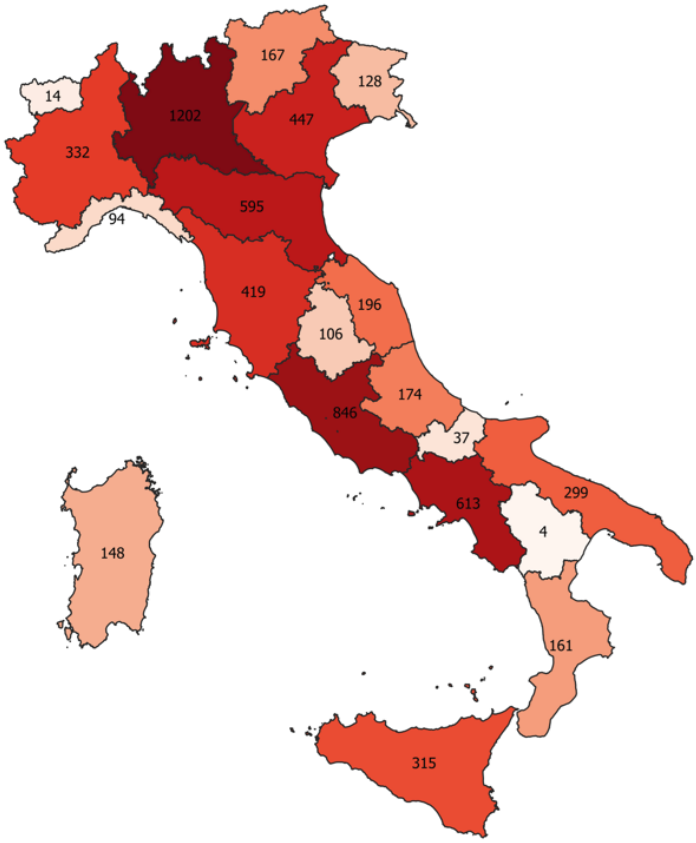


Figure 3: Target population distribution per region

Figure 4: Target population distribution by gender

As it is apparent from Fig. 4, the 60.98% of the target population are men (namely, 3840 people), in contrast to 2457 female researchers.

Moreover, the vast majority of our target population is afferent to economics sectors and statistics, counting for more than half of the total amount of people, as it is apparent from Fig. 5. On the other hand, political science (sector 14/A and 14/B) is the disciplinary domain with less researchers.

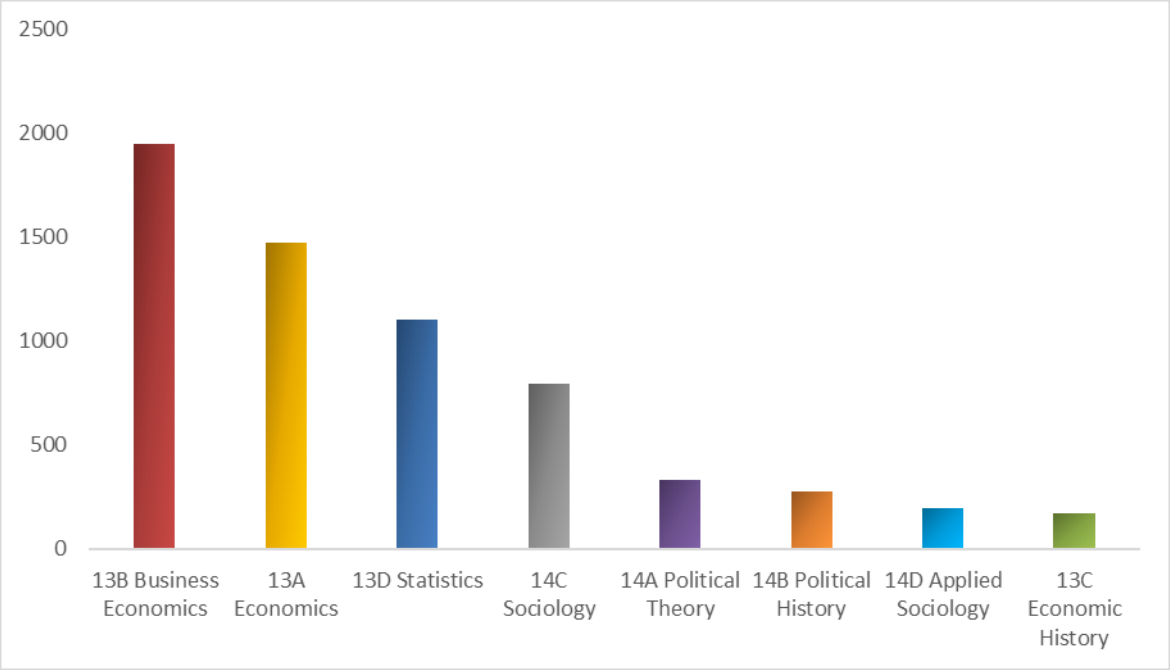


Figure 5: Target population distribution by disciplinary domain

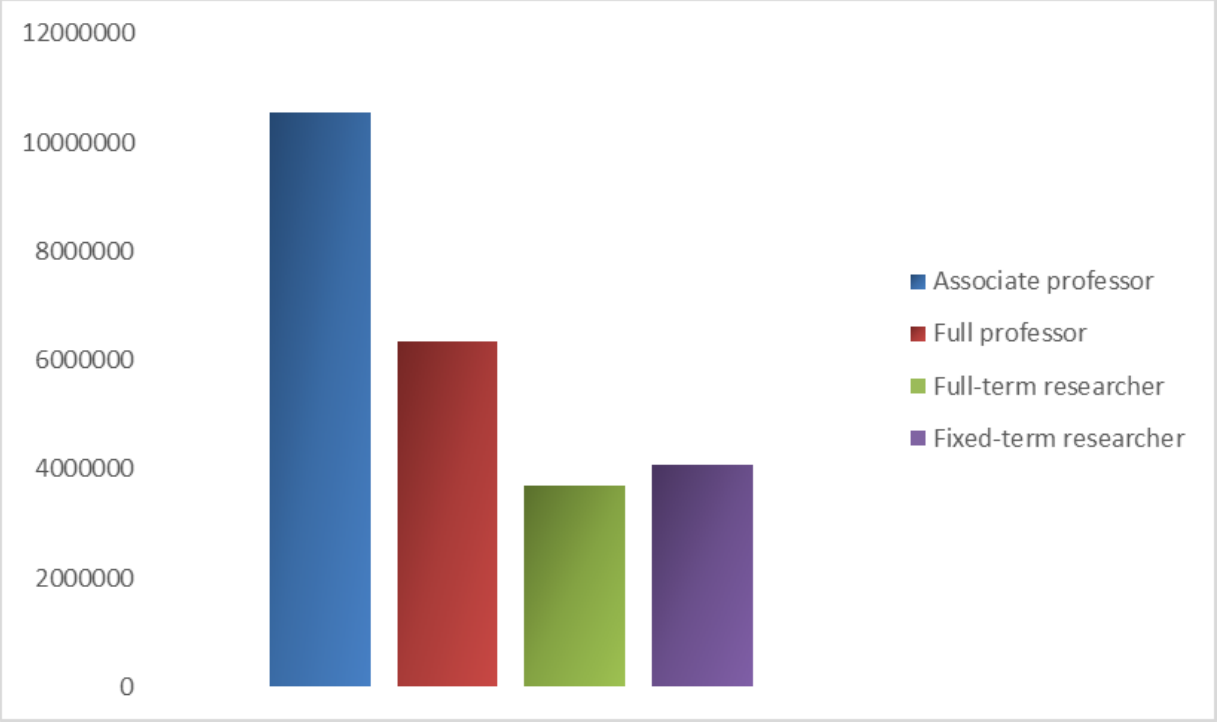


Figure 6: Target population per academic role

According to Fig. 6, it emerges that associate and full professors represent most of the entire population (almost the 75,2%), followed by full term researchers (13,9%) and temporary researchers (10,89%).

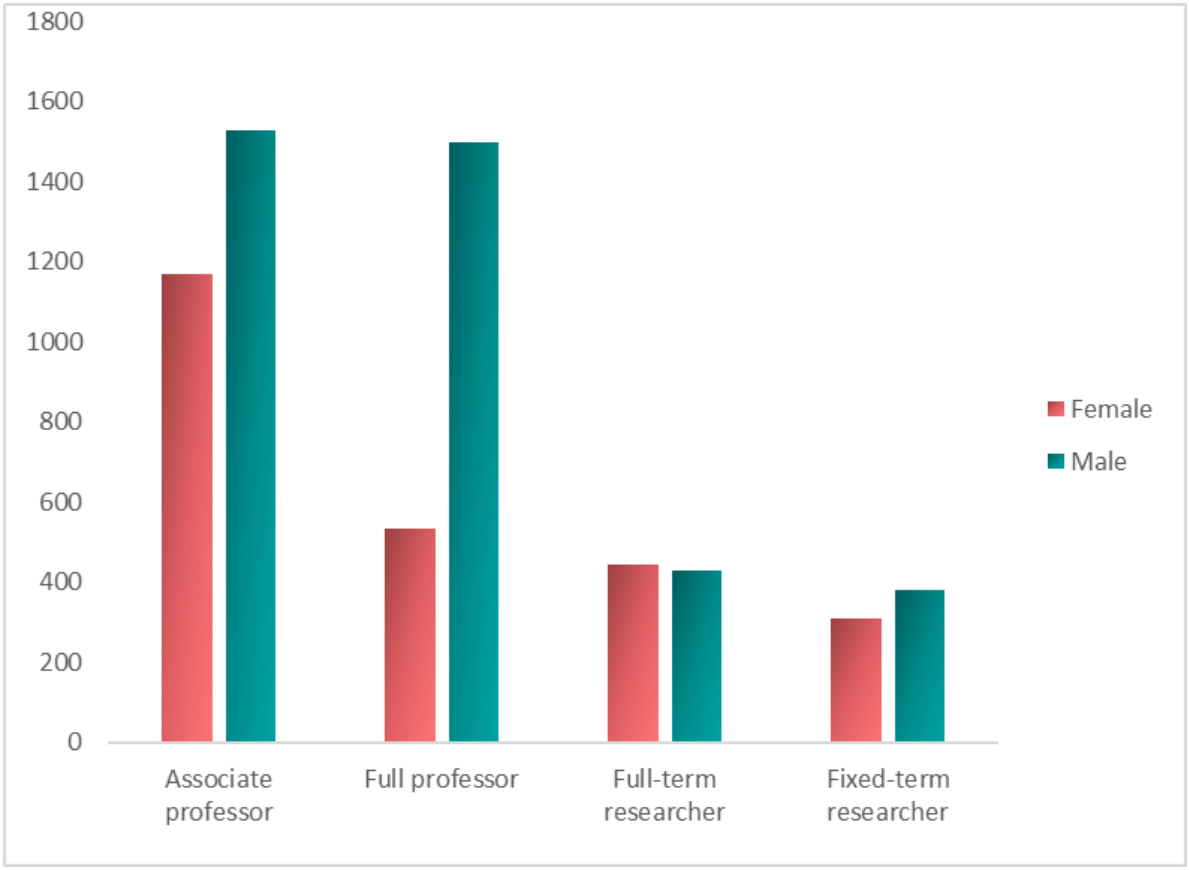


Figure 7: Researchers' population distribution per professional status and gender

To conclude, Fig. 7 shows that men,who represent the majority of the reference population, that is the 61%, outnumber women in most of the groups highlighted above, with the only exception of the full-term researchers group, were women slightly outnumber men.

3.3 Final remarks

The purpose of the present chapter was to illustrate the theoretical framework that was at the basis, on the one hand, of the operative definition of research agenda setting, and, on the other hand, of the drafting of the survey that we have submitted to our target population. As it emerged from the

literature review, this study would be actually very helpful for understanding whether and eventually how Italian researchers' aggregate epistemological curve has somehow been influenced by the introduction of VQR and ASN. This would be extremely interesting to even further our knowledge on the elements that contribute to the advancement of science in a broad sense.
The choice of including both a quantitative and a qualitative approach was hence due to two main considerations, as previously said: on the on hand, we wanted to combine the respective strength and counterbalancing the weaknesses of both methods, in order to grasp the greatest amount of information possible; on the other hand, this need for collecting as many information as possible is primarily linked to the relative newness of the topic at issue.

A final consideration is to be made on the target population. Indeed, we deem the target population very interesting, since it is composed of scholars that, even though they pertain to similar disciplinary fields, they tend to have a different theoretical and empirical approach (the economists have a more empirical-oriented approach, while sociologists and political analysts' research practice comprises an important component that mainly focussed on theoretical debate[68]). Hence, this co-existence of both similarities and differences among disciplines may be very interesting, in the wide context of the studies on knowledge production.

[68] Lamonte, 2009, Bonaccorsi, 2015, Whitley, 2007.

Chapter 4

A survey on the effects of research evaluation on research agenda

In this chapter, we are going to analyse the survey design, the management of the survey and the analysis of the most relevant survey outcomes. As already mentioned in the previous chapter, conducting a webmail self-administered survey on our target population is, in our opinion, the most suitable instrument for the purpose of our research, namely, achieving a better understanding of the effects of research evaluation (VQR) on researchers' research agenda setting. According to the purpose of our research, we can thus classify our survey as a case control study, since it aims at retrospectively understanding a current phenomenon (Kitchenham, Pfleeger, 2002).

First of all, conducting a webmail survey is an easy and cost-effective way to collect the major number possible of data from our target population (Kitchenham, Pfleeger, 2002). Indeed, thanks the available technologies for collecting researchers' email addresses, sending the questionnaires and collecting the responses, it made it feasible to reaching out and collecting a large amount of information with reduced costs. Secondly, including a webmail survey in our research design has proven relatively resilient to bias (Kitchenham, Pfleeger, 2002). Indeed, the intrinsic features of the target population selected (the Italian social scientists' community) contributed to reduce the *under-coverage* bias, since the target population at issue is characterized by a high level of digitalization (Carbonaro, Bainbridge, 2000, Bethlehem, 2010), with daily access to the Internet due to work-related reasons, even increased after the Covid 19 pandemic outbreak (Bethlehem, 2010, Carbonaro, Bainbridge, 2000, Favale et al, 2020). Furthermore, we intended to tackle the *self-selection* bias that can easily occur with online surveys (CAWI) by directly addressing to each recipient in our presentation email, and contacting out the most important academics' organizations, in order to solicit recipients to participate through an awareness campaign.

As already mentioned, this chapter is structured as follows: in the first part, we describe the dimensions of which the survey is composed of; in the second part of the chapter, we are going to

describe how the survey was conducted, from the collection of researchers' contacts to the data cleaning phase, comprising the direct interaction we had with survey respondents; at last, in the third part, we are going to present the most relevant analysis outcomes that will be more relevant for answering our research questions, comprising a content analysis of the questionnaire free comments.

4.1 Survey design

The survey we conducted was structured on the basis of the research agenda setting determinants that we furthered in the previous chapter, with the purpose of answering our three research questions, namely: 1) whether researchers' research agenda setting has been influenced by the introduction of research evaluation (VQR), 2), whether there have been differences between recruitment-oriented research evaluation (ASN) and universities-oriented research evaluation (VQR) in terms of effects on the researchers' research agenda setting process, and, 3) if it is possible to identify different effects of the VQR introduction across the various scientific discipline.

In this section, we will address the operationalization of the concepts underlying the several dimensions the survey is composed of, namely:

- the sociodemographic dimension,
- an extract of the Likert scale from the extended version of the MDRAI-R model (Horta, Santos, 2019),
- the influence of collaboration on researchers' research agenda setting,
- the influence of peers recognition, of topics linked to the news story, of the mainstream topics and the wish to contribute to society, on researchers' agenda setting,
- the influence of competitive calls for funding on researchers' research agenda setting,
- the influence of the introduction of ASN on researchers' research agenda setting,
- and, lastly, the influence of the introduction of VQR on researchers' research agenda setting.

Furthermore, in the second part of this section we will explain the methodological choices adopted regarding the survey drafting and the target population.

4.1.1 Survey dimensions: sociodemographics

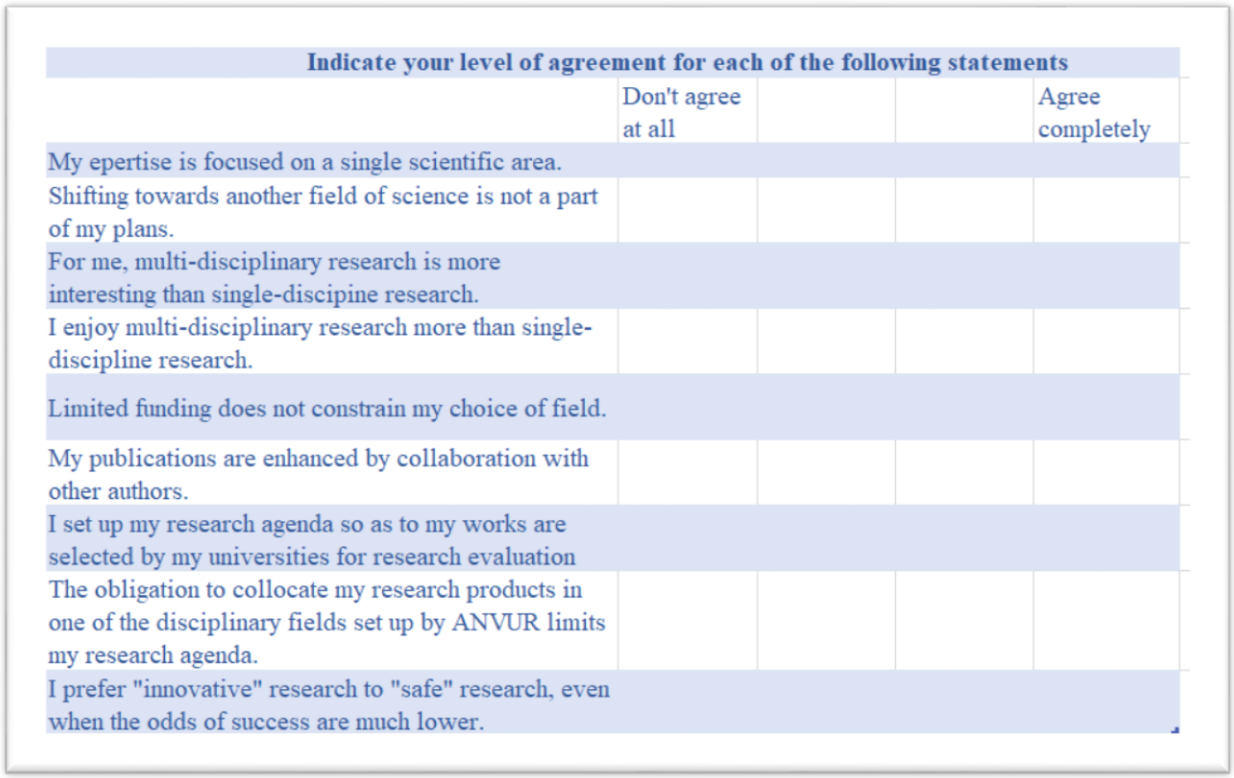
The survey starts with some questions regarding the respondents' biographical information (such as gender, age and nationality), and their working condition (their seniority, the academic position, the afferent disciplinary field). These questions were meant to work as a background for the subsequent analysis we aimed to conduct on the research agenda setting-related choices, especially considering the gender, the disciplinary field and the academic position.

We decided to put these questions at the beginning of the survey to put the respondents at ease (Babbie, 2020), since it does not seem to generally affect the drop-out rate (Rauhut et al, 2020).

4.1.2 Survey dimensions: an extract from MDRAI-R model

As already explained in the previous chapter, this section comprises an extract from the extended version of the MDRAI-R model and comprises the MDRAI-R dimensions that let best distinguish the cohesive and trailblazing archetypes (Fig. 8 below), according to the dimensions scores collected in Horta and Santos previous studies (Horta, Santos, 2018, Horta, Santos, 2020, Horta, Santos, 2019).

The MDRAI and MDRAI-R dimensions reported in the survey are *Convergence* (*Mastery* and *Stability*), *Divergence* (*Branching out* and *Multidisciplinarity*), *Tolerance for low funding*, and *Academia driven*. The two sentences related to the *Academia driven* dimension have been rephrased, to include a direct reference to Italian evaluation system. The questions are structured in a Likert scale, without the medium response item, to stimulate respondents to take a stand. However, respondents were given the possibility to press the *N/A* button for whatever reason, and to continue the questionnaire (even if only the totally fulfilled questionnaires were considered in the analysis). The statements are ordered in a random way to avoid the response-set bias effect, with the poles of the Likert response categories alternated (Babbie, 2020). It is worth citing that in the MDRAI-R model the same dimension is operationalized in more scale items, but we tended to avoid the repetition of items with the same concept, to make the questionnaire quicker.



Indicate your level of agreement for each of the following statements				
	Don't agree at all			Agree completely
My epertise is focused on a single scientific area.				
Shifting towards another field of science is not a part of my plans.				
For me, multi-disciplinary research is more interesting than single-discipine research.				
I enjoy multi-disciplinary research more than single-discipline research.				
Limited funding does not constrain my choice of field.				
My publications are enhanced by collaboration with other authors.				
I set up my research agenda so as to my works are selected by my universities for research evaluation				
The obligation to collocate my research products in one of the disciplinary fields set up by ANVUR limits my research agenda.				
I prefer "innovative" research to "safe" research, even when the odds of success are much lower.				

Figure 8: Extract of the revisited MDRAI-R model

4.1.3 Survey dimensions: collaboration

The issue of collaboration has been both analysed in the Likert scale section and in its dedicated section. This section opens with a filter question: if the respondent has affirmed not to seek for scientific collaborations, he will consequently be directed to the following section. If the respondent has replied to (at least partially) seek for scientific collaboration, he will be directed to the two following questions regarding the respondent's collaboration attitude: he is asked whether he seeks for collaborations with peers from the same or another disciplinary field. At the end of this section, we inserted a bar for typing a free comment, with which the respondent has the possibility to express his opinion on this topic (or on whatever topic he wants to).

4.1.4 Survey dimensions: peers recognition, topics linked to the news story, the mainstream topics and the wish to contribute to society

This section is composed of four Likert scale questions in which the respondent is asked whether one of these determinants have contributed to the formulation of his research agenda setting, as it has already been observed in the available literature (Åkerlind, 2008). These questions have the *I strongly disagree*, *I somewhat agree*, *I somewhat disagree* and *I strongly agree* as response options randomly distributed, to avoid the response-set effect. In this case, there is no the *N/A* response option.

4.1.5 Survey dimensions: competitive calls for funding

This section is composed of a double filter question. With the first filter question, we ask the respondent whether he has ever participated to a competitive call for funding: if he replies positively, we will be addressed to the second question of the section; otherwise, he will be directed to the following survey section. The second filter question asks the respondent whether the participation to a competitive call for funding has somehow influenced his research agenda setting: if he replies positively, we will be addressed to the third (and last) question of the section; otherwise, he will be directed to the following section. The last question is a closed-ended question, in which the respondent can express how the participation to competitive calls for funding has affected his research agenda setting.

The two options are:

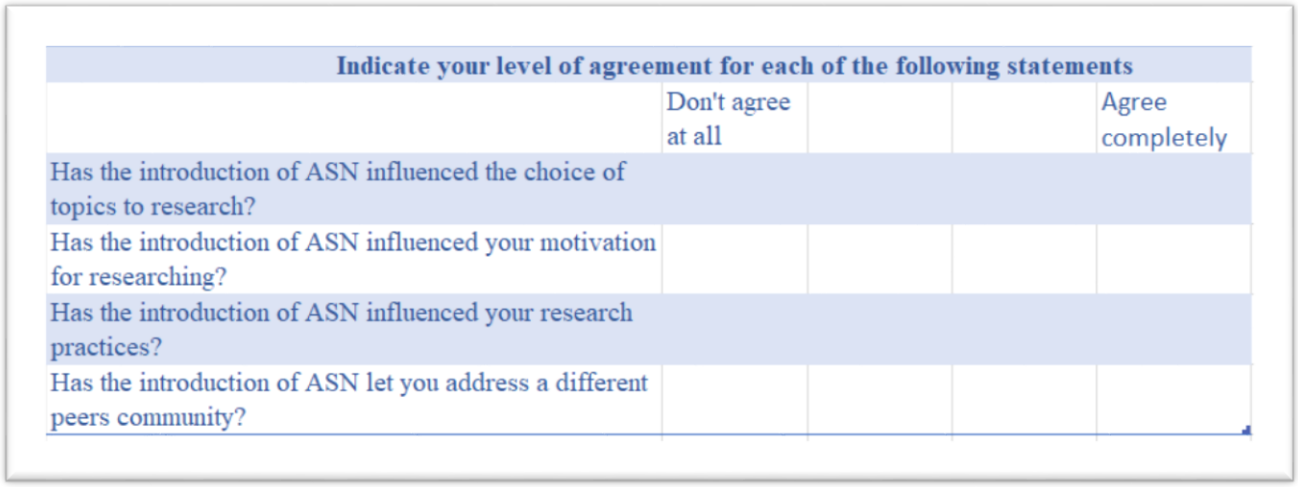
- *Participating to competitive calls for funding has subtracted me time for my own research agenda*;
- *Participating to competitive calls for funding induced me to change my research agenda plans.*

Even at the end of this section, we inserted a bar for typing a free comment, with which the respondent has the possibility to express his opinion on this topic (or on whatever he wants to).

4.1.6 Survey dimensions: the introduction of ASN

This section starts with a filter question, in which the respondent is asked whether his research has been influenced by the introduction of ASN. If the respondent replies positively, he will be addressed to the other questions on the influence of ASN on research agenda. Otherwise, he will be addressed to the last section of the survey.

The four following questions address the influence of the introduction of ASN on the different research agenda setting components (Fig. 9 below).



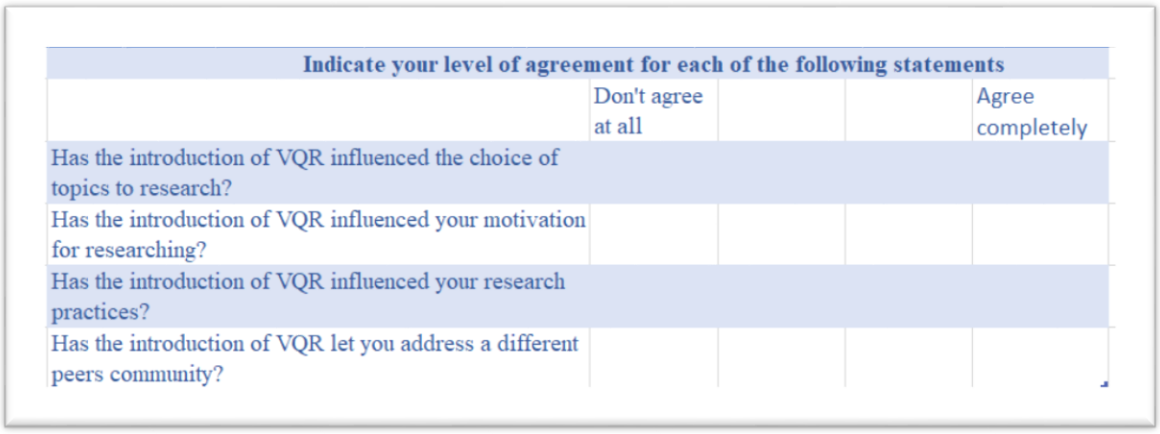
Indicate your level of agreement for each of the following statements				
	Don't agree at all			Agree completely
Has the introduction of ASN influenced the choice of topics to research?				
Has the introduction of ASN influenced your motivation for researching?				
Has the introduction of ASN influenced your research practices?				
Has the introduction of ASN let you address a different peers community?				

Figure 9: Survey extract on the influence of ASN on research agenda components on research agenda components

Even at the end of this section, we inserted a bar for typing a free comment.

4.1.7 Survey dimensions: the introduction of VQR

This section is very similar to the ASN related section, for the filter mechanism, and the questions on the influence of VQR on the research agenda setting components.



Indicate your level of agreement for each of the following statements				
	Don't agree at all			Agree completely
Has the introduction of VQR influenced the choice of topics to research?				
Has the introduction of VQR influenced your motivation for researching?				
Has the introduction of VQR influenced your research practices?				
Has the introduction of VQR let you address a different peers community?				

Figure 10: Survey extract on the influence of VQR on research agenda components on research agenda components

We positioned this set of questions at the end of the survey, to gradually address a topic that could have been perceived as sensitive by the respondents (Babbie, 2020).

This section ends with two free comments bar: in the first, respondents are encouraged to freely express themselves on the introduction of VQR, while, in the second bar, respondents are invited to write about the issue the prefer.

4.1.8 Survey design: further details and settings

The questionnaire was designed to be online, and we chose to adopt the open source LimeSurvey platform (https://www.limesurvey.org/), among the most commonly used online survey platform (such as Google Forms or SurveyMonkey), to upload and manage our webmail survey. LimeSurvey seemed to be the most suitable online platform for our purpose, since it has quite acceptable costs, for the need we had to collect a relative high number of responses, in comparison to SurveyMonkey (https://it.surveymonkey.com/). Still, LimeSurvey presents more functionalities than Google Forms (https://www.google.it/intl/it/forms/about/), and this represented a particularly important reason for choosing LimeSurvey, since it allowed us to easily structure the questions and answers options of the survey according to our needs.

The survey was set up in Italian, since we were expecting most of the respondents being Italians. Moreover, the use of Italian was deemed suitable to easily convey some concepts that may be tricky

to grasp in another language, and that could consequently push the respondents to drop the survey. Furthermore, to make the questionnaire understandable, we added some captions to explain the respondents how to proceed and to explain some concepts, along with an initial explanation of the concept of research agenda setting.

Respondents could access the questionnaire by clicking the hyperlink attached to the invitation email, and no password was required to access the form. The respondents were required to read and accept the privacy form attached to the survey webpage, before starting to fulfil the survey. The privacy form informed the respondents of all the privacy measures adopted (anonymization, measures for the storage of files and data). The completion of the survey required minimal digital skills (Carbonaro, Bainbridge, 2000). As previously seen, the survey is mainly composed of a series of closed-ended questions, except for the age-related question and the free comment bars, with no characters' limit. The questions, most of them categorical, were mainly of two types: nominal closed-ended questions, or Likert scale items. The respondents were required to reply to all the questions, except for the free comment-related questions (and, of course, the filtered questions that did not pop). Still, the respondent was not allowed to freely browse the survey, but just in the webpages in which the questionnaire was structured. At the end of the questionnaire, the respondents were invited to leave their email address if they would have appreciated to be contacted subsequently for the second phase of the research. By pressing the SUBMIT button, they could turn in the fulfilled form.

4.2 Survey management

4.2.1 Target population and collection of contact data

As already analysed in the previous chapter, the target population is composed of 6297 individuals[69], namely the social scientists' currently working in public and private universities. More specifically, we aimed to address full professors, associate professors, full-term researchers, and temporary

[69] As of the data collected in October 2021.

researchers from economics (13/A, 13/B, 13/C, 13/D), political analysis (14/A, 14/B) and sociology (14/C, 14/D). We decided to reach out the entire population with a census approach, thus not adopting a sampling technique, since the size of the target population at issue was not that big to make the submission of the survey, the survey management, and the collection of data unfeasible (Draugalis, Plaza, 2009, Kitchenham, Pfleeger, 2002), and to get the largest possible number of responses. It is worth citing, nonetheless, that when the entire population is surveyed, only with a 97% response rate we can speak of high level of confidence in the survey outcome, and that the self-selection bias could have occurred, since even though we personally addressed the target population, the responses to the survey questions may be biased because of the respondent's interest (both positive and negative) for the topic at issue.

The collection of the target population members' contacts was conducted manually, since the census of researchers working in Italian universities reports, among the personal information, only the university of affiliation, name, and surname. The collection of each researchers' email address was possible only by acceding the webpage of the university of afference. The collection of the institutional email addresses did keep three months, from August 2020 to the end of October 2020, comprising the time needed for the creation of the dataset, using Microsoft Excel. The target population members were grouped according to their afferent disciplinary field.

4.2.2 Pilot study

We conducted a pilot study in November 2020 with a survey uploaded on Google Forms platform. The primary objective was to ensure that the questions were easily understandable for the target population and to identify any typing error. The pilot survey was submitted to researchers from different disciplinary fields (sociologists, statisticians, economists and political analysts), afferent to La Sapienza University and to IRCRES-CNR[70]. It was of our concern to submit the survey to

[70] These two are the afferent institutions of the undersigned.

researchers from different disciplines, to attempt to have a quite thorough overview of the possible flaws of our survey. The pilot study had also the objective of enabling a discussion on the strengths and weaknesses of both the structure and the phrasing of the questionnaire (Jones et al, 2013), with a focus-group like approach. Indeed, in the email we directly asked our colleagues to feel free to advance suggestions and critiques for discussing them.

From the pilot study, some relevant suggestions emerged, specifically for the online platform previously adopted (namely, Google Forms). Indeed, the pilot study showed that managing a webmail survey through Google Forms to such a large number of people was rather difficult, especially for privacy-related issues, since respondents were required to insert their email address to submit the fulfilled questionnaire. Moreover, with Google Forms we had severe limitations in creating a survey with filter questions.

Furthermore, some suggestions were advanced on the phrasing of the many questions, and to reduce the number of the Likert scale items that tended to be repetitive. Still, some suggestions were particularly important for a correct translation from the original language of the MDRAI-R model (English) to the language used in the survey (Italian).

4.2.3 Timetable and organisation of the survey

The survey was drafted from July 2021 to October 2021, and finally revised in December 2021 after the pilot study. The collection of researchers' institutional email addresses and the creation of the dataset lasted from August 2021 to October 2021.

For sending the emails, we opted for an automated mailing system, attaching to the email the link to the LimeSurvey online platform, with a brief explanation of the goals of our research and of the meaning of the concept of researchers' research agenda setting. Otherwise, the most feasible alternative was to send the emails through the LimeSurvey platform, which is autonomously able to launch the survey to the email addresses uploaded on the LimeSurvey website.

Nevertheless, we chose to send the emails through the automated mailing system method, to better ensure respondents' privacy. Indeed, this method ensures survey respondents a total anonymity (while, on the other hand, the LimeSurvey platform can keep track of those who fulfilled the survey, and consequently sending the follow-up emails only to those who have not replied yet to the survey). The survey was sent to the academics' institutional email addresses on 5th January 2022, with one follow-up email for participating to the survey on 26th January 2022. In the second email we informed the recipients that the time window for filling out the survey was open till 11th February 2022.

While sending the surveys, in some cases (11) we noticed that the email addresses were not updated, and we were able to fit this issue and to correct the dataset. In other cases, some researchers could not be reached at all. Some recipients were also excluded from the target population since they were retired when the dataset was created.

4.2.4 Response rate

With the first sending of the emails, the survey has recorded a 29,51% response rate (Fig. 11): the pink section of the bars identifies the women percentage response rate, while the blue section stands for men; the bars are grouped into four main clusters, according to the academic role, and each bar represents the percentage of the survey respondents divided per academic role and disciplinary field, of which full and associate professors counted respectively for the 29,02% and the 31,01%. Still, full-term researchers' group recorded a 19,5% response rate, while temporary researchers 28,25% response rate. The recorded response rate in this first part is almost equally distributed between men and women, only with a slightly higher response rate for men (28,9%), in comparison to women response rate (26,2%).

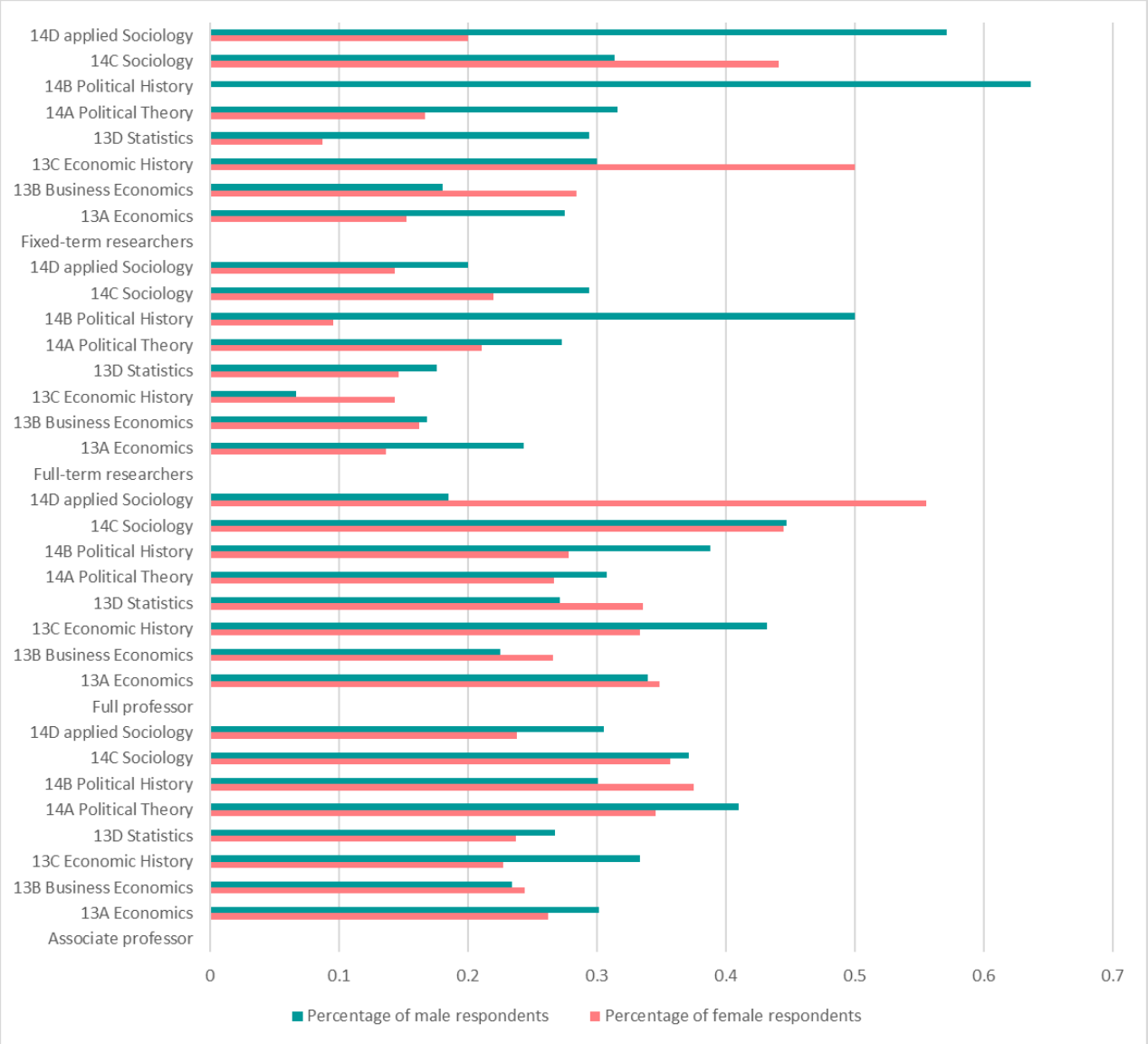


Figure 11: Response rate percentage by role and gender as of 20th January 2022

As it is apparent from Fig. 11, after the first emails sending a discipline subgroup resulted underrepresented, namely, male temporary researchers of *Storia Politica* disciplinary sector. The issue of achieving good representativeness in all the disciplinary subgroups was one of the main targets of the follow up emails.

The follow up emails were sent on January 26th, to all the researchers of the list, since, as already mentioned, it was not possible to know who had already replied to the survey (except for those who explicitly communicated to have participated to the survey). This has triggered some criticisms

expressed by some respondents, who complained about the pressure due to the recall. When this criticism was raised, it was of our concern to explain that the method used to send the emails was aimed to respect respondents' privacy as much as possible.

The total response rate recorded in this phase is at 36,64%, with all the subgroups represented (according to Fig. 12 below). The response rate trend has not showed significant changes, except for the overall increase recorded in each subgroup. Hence, associate and full professors remain the most represented subgroups (with, respectively, 38,64% and 38,8% response rate), temporary researchers response rate is at 35,13%, and full-term researchers at 26,6%. Moreover, associate and full professors remain the most represented subgroups (with, respectively, 38,64% and 38,8% response rate), temporary researchers response rate is at 35,13%, and full-term researchers at 26,6%.

The total number of respondents is 2299, which corresponds to a 36,55% response rate of the overall respondent population The response rate gap between men and women has slightly decreased, since men response rate goes from 28,9% to 36,52%, and women response rate from 26,2% to 35,6%.

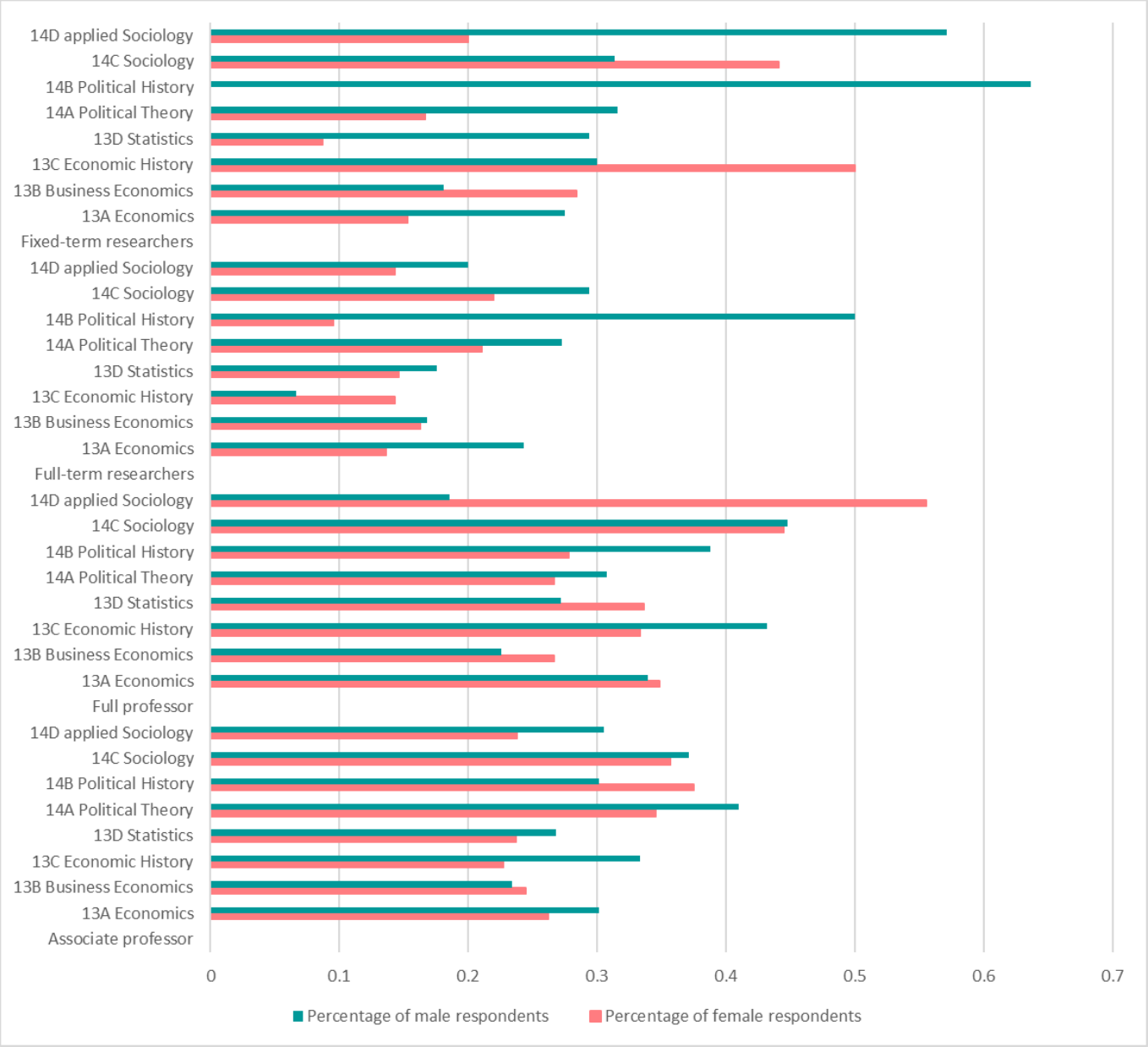


Figure 12: Response rate percentage by role and gender as of 11th February 2022

4.2.5 Data cleaning and survey dropouts

The data cleaning was conducted with Microsoft Excel and SPSS. From the LimeSurvey platform it was possible to directly download the survey outcome in different file formats. During the data cleaning process, that took about one month, we excluded from the final sheet the dropped out questionnaires recorded by LimeSurvey.

The questionnaire dropping out is as follows: 157 respondents dropped out at the initial part of the survey (the biographical information related part); 174 respondents dropped out before the beginning

of the first Likert scale; 1 respondent dropped out while answering the first Likert scale; 50 respondents dropped out during the "Academic collaboration" set of questions; 17 dropped out in the "Call for research funding" set of questions; 8 respondents dropped out at the ASN-related questions, and, finally, 2 respondents dropped out at the VQR-related questions. One possible explanation of this issue may be the cumbersome nature of some and complexity of the questions, as remarked by some respondents' replies to the survey. Seeing the data of the questionnaires dropped out, it seems apparent that the Likert scale section of the survey represented the most problematic part to fulfil. Some respondents, indeed, directly contacted us to ask for an explanation of some concepts of the Likert scale items.

4.2.6 Data-protection strategy and anonymization

It was possible to anonymize each fulfilled survey with the LimeSurvey privacy settings. Furthermore, the anonymization was also possible since we tried to make it impossible to trace back the respondent from the sociodemographic information.

Still, since when the respondents expressed their availability for the second phase of the research it was possible to trace back the respondents' fulfilled questionnaire, we split the list of email addresses from the file with the survey responses, to keep the anonymity.

The data were stored in our personal computer and protected by an entry password. The survey respondents were informed of the data protection measures by the privacy form attached to the questionnaire.

4.2.7 Respondents' reactions

Interestingly, it is noteworthy to report an unexpected relatively high level of personal participation of the people contacted for the survey, both in this first phase, and in the second phase.

In the first phase, 283 people directly replied to the email sender for a variety of reasons, mostly expressing their enthusiasm and interest for the issue proposed. Moreover, most of respondents

seemed to be overly concerned in sustaining the present PhD project, also for enabling the improvement of Italian research evaluation policies and the understanding of the current academics' working conditions. The respondents' emails were likewise to inform of the participation to the survey. Among these people, 10 of them provided some useful suggestions in order to improve both the conceptual framework of the questionnaire (with some tips on the variables to be considered in the analysis and in the subsequent phase of the research, or even the phrasing of the survey questions), and the survey management. Still, 26 people expressed interest in the results of the research and 13 made themselves available for the second qualitative research phase. Interestingly, some researchers directly called by phone to express their strong criticisms against the research evaluation process adopted by the Italian government.

Even in the second phase, more than 100 researchers have directly expressed their interest in the current research on the effects of research evaluation, to inform about their participation and to subsequently stay in touch for the advancement of the research itself (2 of whom have expressed interest in conducting further research projects based on the present research outcomes). Still, 4 people expressed their opinion on the conceptual framework of the questionnaire and expressed some useful suggestions on the following analytical phase.

4.2.8 Limits of the survey and of the use of telematic support

We can mainly identify two main limits of the survey. First, the study is mostly based on perception data, which can lead to biased responses, due to respondents' personal interpretation of concepts and phenomenon proposed in the survey, which depends on respondents' previous experiences, ideas and values (Kitchenham, Pfleeger, 2002). Some survey recipients, indeed, informed us that they had some difficulties in interpreting some concepts and items, even though the survey was endowed with some explanations captions, to make easier the comprehension.

Still, the choice of not sampling the target population may have determined a self-selection bias in the responses, and thus the survey responses might have recorded the opinion of researchers

particularly affected or interested in the topic, leaving aside other important opinions on the matter at issue. However, while running the analysis of the survey responses, as we will describe later, we have weighted the values recorded per each subgroup, in order to counterbalance the self-selection bias.

As for the limitations of the use of telematic tools in this specific phase of the research, as per what is specifically concerns this phase of the research, we did not encounter relevant difficulties in running the survey. Indeed, as previously specified, the main strengths of conducting a webmail survey are mostly linked to the possibility of easily reaching out a high number of people, breaking down any geographical barrier. Furthermore, the use of online tools was particularly helpful for the specific contingencies of the period in which the webmail survey was submitted, when, due to the Covid 19 pandemic, the obligation to keep the social distancing and the eventual difficulties and consequences in arranging face to face meetings. On the other side, among the weaknesses of the massive use of online tools we can mostly include the loss of some important information (due to the eventual bugs of the websites or programmes for collecting and analysing data).

4.3 Analysis of the survey outcome

4.3.1 Descriptive analysis of the survey

4.3.1.1 Descriptive analysis: gender, academic role, disciplinary field and university dimension

As previously said, the total number of respondents is 2299, which corresponds to a 36,55% response rate of the overall respondent population. Men (*Uomo*) count for the 60,8% of the respondent population, women (*Donna*) for the 38%, and the *I'd rather not answer* category for the 1,2% (Fig. 13 below).

Moreover, as shown in Fig. 14, the most represented academic role is associate professors (*Professore associato*), which represents 45,2% of the respondent population, followed by full professors (*Professore ordinario*), representing the 34,2%, then temporary researchers (*Ricercatore a tempo*

determinato di tipo B) follow, with the 10,4%, and full-term researchers (*Ricercatore ruolo ad esaurimento*), 10,1%.

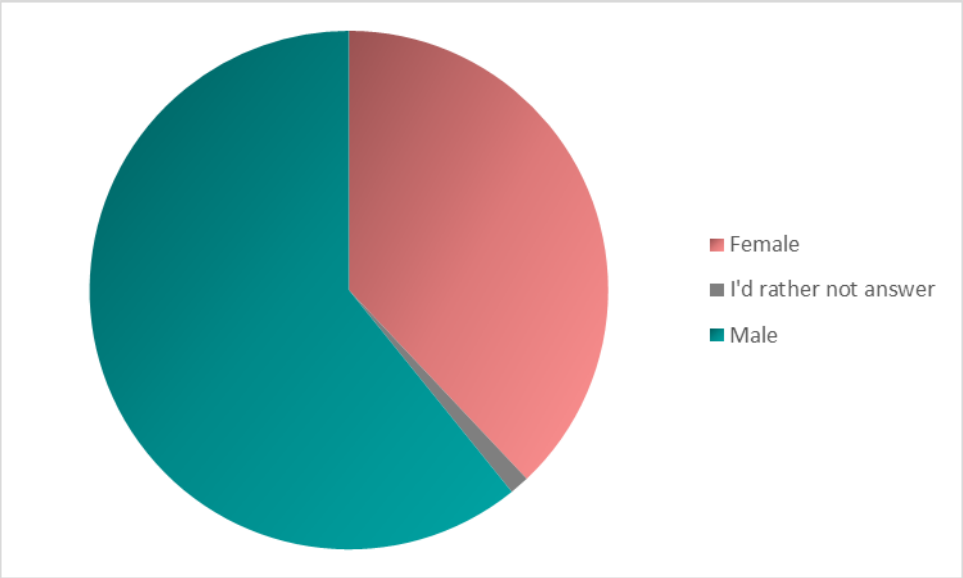


Figure 13: Respondent population distribution per gender

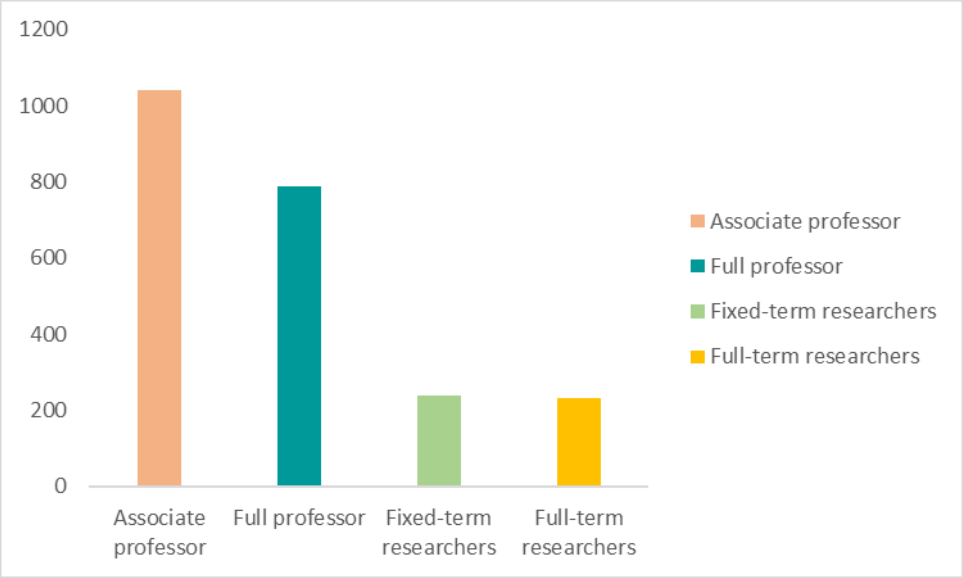


Figure 14: Respondents population distribution per academic position

From the questionnaire it emerges that 98,8% of the population claim to be Italian, 1,1% to come from a European country, and, at last, 0,1% from a non-European country. For the very limited number of non-Italian researchers, we have decided to recode this variable into two modalities,

instead of three, that is, Italian nationality, and the non-Italian nationality (which now counts for the 1,2% of the respondent population).

As said in the 4.2.4 section, all disciplinary sectors are represented in the respondent population (Fig. 15), with the 13/A (23%) and 13/B (31%) CUN areas representing more of half of the target population. Nonetheless, the groups of respondents referring to different disciplinary fields have recorded quite different response rate levels. Indeed, as it is apparent from Fig. 16, the 14 CUN areas (sociology and political analysis, with the highest percentage recorded by the 14/C area) are the most represented disciplinary fields, in comparisons to the response rate recorded in economics related subgroups (except for 13/C area, that recorded 41,6% response rate).

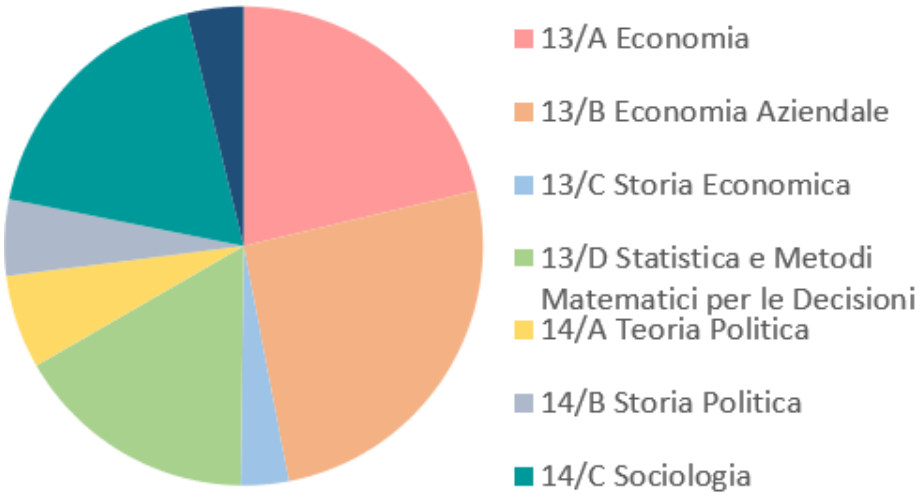


Figure 15: Percentage distribution of respondents' population per disciplinary sector

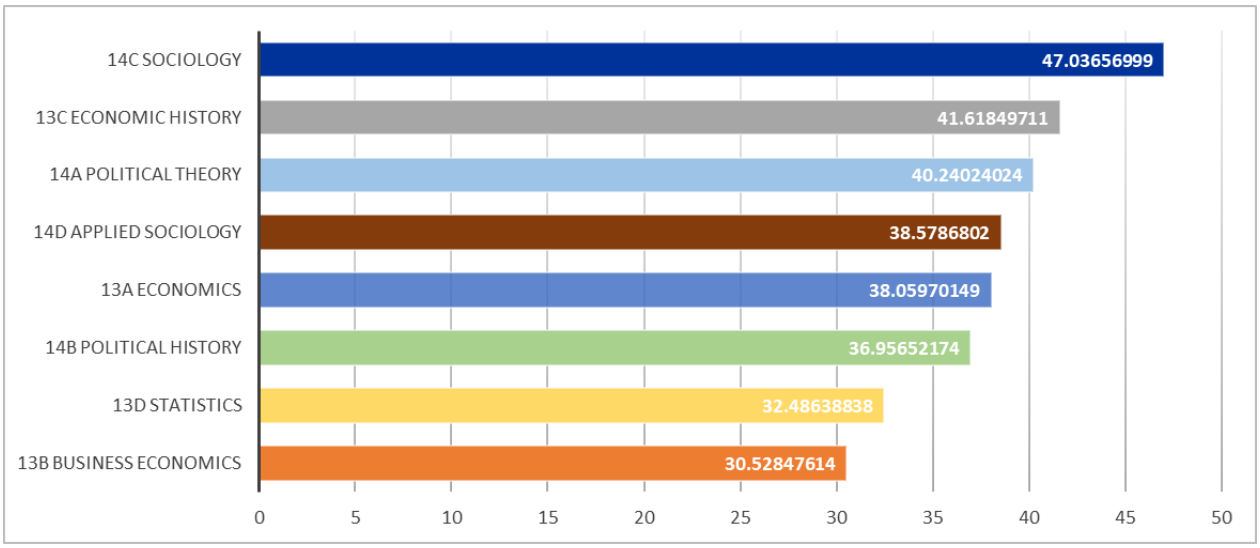


Figure 16: Percentage response rate per disciplinary field

Considering the variable *public/private university of affiliation*, it emerges that 88,8% of the respondents work in a public university, and the 11,2% of respondents work in a private facility. Indeed, of the 20 private universities of affiliation, only 5 are major universities, while the others are quite small facilities.

Moreover, the overall respondent population is distributed through all the modalities of the variables (the mode is the *Above 40.001 enrolled students* modality, with the 36% of the respondents' population). The data recorded are quite in line with the size of the target population groups obtained by clustering researchers' according to the dimension of their university of afference (Fig. 17).

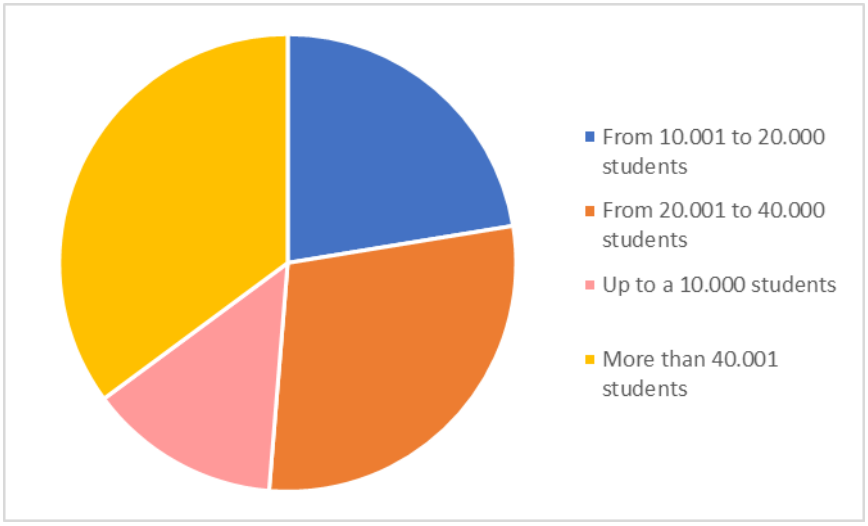


Figure 17: Respondents' population percentage distribution per university dimension

According to Fig. 18, it emerges that most survey respondents have a long-lasting experience in the academic working environment, since *Above 11 years of experience* is by far the modality with the highest number of cases. The relative low frequency of temporary researchers, of the *Till 5 years of experience* and *From 6 to 10 years of academic experience* categories may be easily explained considering the reduced economic resources invested for hiring academic personnel (Act of 2009, no. 9[71] and Act of 2010 no. 122[72]).

As a matter of fact, these results are in line with an analysis of respondents' age (as apparent from Fig. 19 below), since we have most respondents who claim to be 40 years old or older.

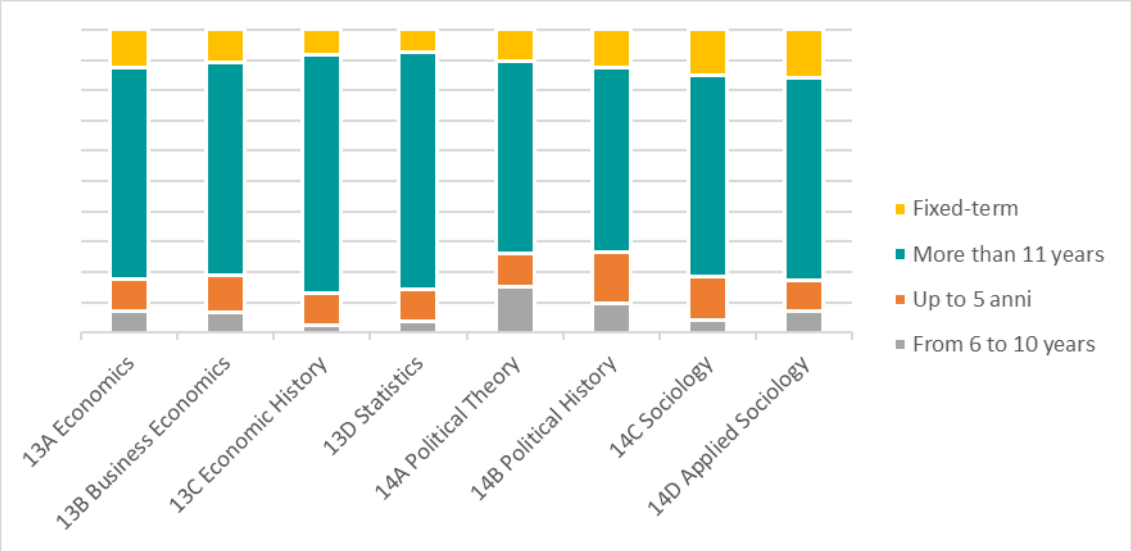


Figure 18: Respondents' population percentage distribution per academic seniority

[71] Legge 9 Gennaio 2009, n. 1: https://www.cun.it/uploads/storico/legge_1_09_01_2009.pdf.
[72] Legge 30 Lugliio 2010, n. 122: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2010-07-30;122.

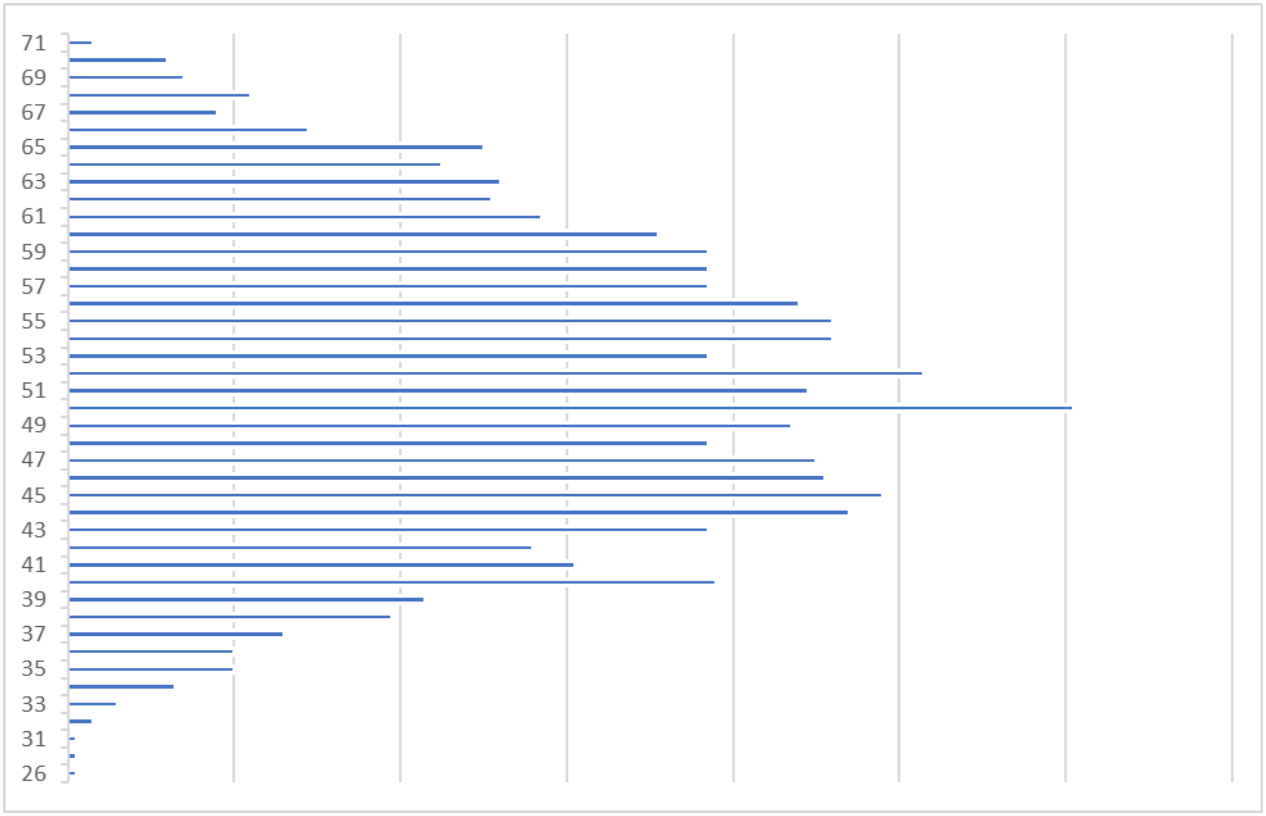


Figure 19: Respondents' population distribution per age

4.3.2 Analysis of the effects of the determinants of research agenda setting

4.3.2.1 VQR effects on research agenda setting

According to the responses collected from the question asking respondents whether they changed their research agenda setting because of the introduction of the VQR, the majority of respondents have affirmed that they did not change their research agenda setting process (Fig. 20 below). Indeed, 60,07% of respondents replied that they have not changed their research agenda (green area of Fig. 20 below), while 39,9% of respondents affirmed that they changed their research agenda (green area of Fig. 20 below).

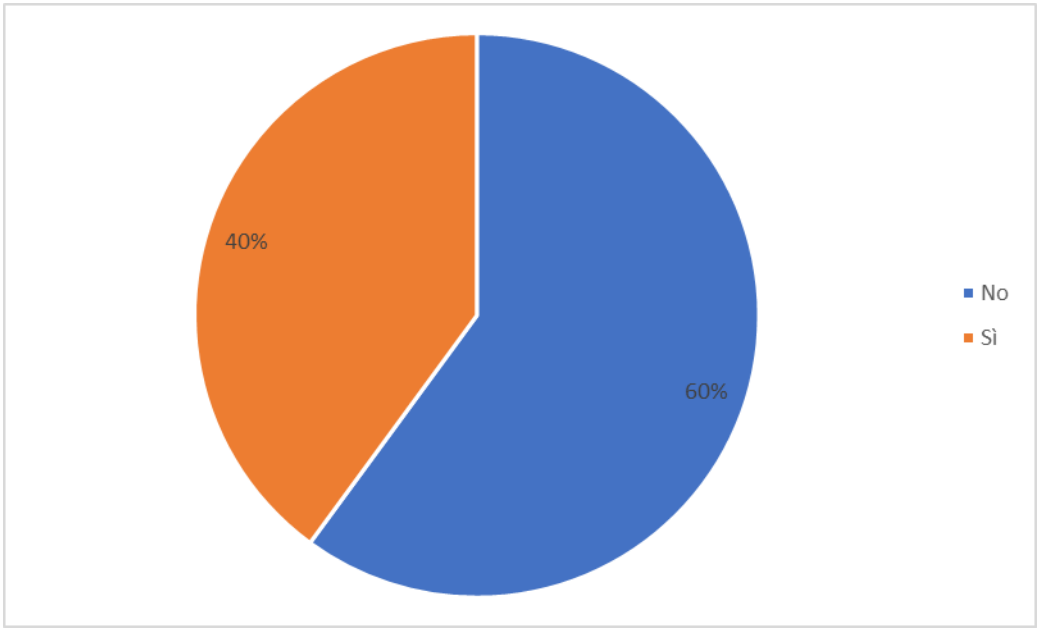


Figure 20: Percentage distribution of survey question asking respondents whether they changed their research agenda after VQR

This first result is very interesting, since it emerges that the indirect institutional pressure exerted with introduction of VQR in order to foster universities and researchers' research production, has not considerably affected the research agenda setting process of the majority of respondents.

A logistic regression model was adopted for the answering to our first research question (that is, identifying the effects of the introduction of VQR on researchers' research agenda setting-related behaviour) and to propose a more in-depth overview.

The logistic regression model (henceforth, logit) was deemed the most suitable model since in the survey the question[73] referring to the effects of the introduction of VQR on research agenda is a binary variable, where with 0 we indicate the certainty of the absence of effects of the introduction of VQR on research agenda, and 1 represents the certainty of the presence of effects of the introduction of VQR on research agenda. We used STATA for conducting the multivariate analysis. Before running the model, we had to insert an additional variable into the dataset to weight the different sizes of the sample, to counterbalance the eventual selection bias due to great variety among the different target

[73] Question number 21 of the survey: *Has your research agenda setting been influenced by the introduction of VQR?*

population subgroups. The weight was computed considering the size of the total number of people per academic profession, by gender, by disciplinary area and dimensions of afferent university.

In the computation of our model, the dependent variable was the presence/absence of the effects on research agenda after the introduction of VQR, while the independent variables where: academic position, gender, age, academic seniority, dimension of the afferent university, and disciplinary field, namely, the CUN area (Tab. 5 below). We had to run two models, because two of the independent variables, namely the academic position and the academic seniority (which is quite understandable, since, generally, the more is the researcher's experience, the higher is his academic position) are strongly correlated, making it difficult or impossible to estimate their individual regression coefficient reliably. Hence, in the first model the academic position was considered in the computation and the academic seniority excluded (left column of Tab. 5), while in the second one the academic seniority was included in the computation and the academic position excluded (right column of Tab. 5).

First of all, according to Tab. 5, the estimated coefficient associated with the independent variable *gen* (namely, the variable associated to the gender) is positive and statistically significant in both Model 1 and Model 2. We can thus affirm that women tend more likely to change their research agenda due to the introduction of VQR. Secondly, from Model 1 it emerges that the estimated coefficients of the different modalities of the independent variable *pos_acc* (identifying the respondents' academic position) are not statistically significant, whereas the estimated coefficients of the variable *eta_new*[74] are all positive and statistically significant. We can thus affirm that as the respondent's age increase, the probability of changing their research agenda related choices increases as well. In Model 2, the relevance of respondents' academic experience is even more apparent. Indeed, the *4.anz_acc* modality of the *anz_acc* variable[75], which identifies the group of academics with more than 11 years of academic experience, is the only modality with a positive and statistically

[74] This variable identifies the respondents' age. We recoded the variable into 4 modalities, namely 34 years old or younger (*1.eta_new*), 35-44 years old (*2.eta_new*), 45-55 years old (*3.eta_new*), and 56 years old and older (*4.eta_new*).
[75] Identifying the respondents' academic seniority.

significant estimated coefficient. We can consequently assume, according to both Model 1 and 2, that as the researchers' academic seniority increases, the probability of changing their research agenda setting-related choices considerably tend to increase.

VARIABLES (-anz_acc)	VQR_infl_ra		VARIABLES (-pos_acc)	VQR_infl_ra	
	Coefficient	Standard error	Coefficient		
gen	0.510***	(0.120)	gen	0.547***	(0.120)
2.pos_acc	0.191	(0.139)	2.anz_acc	0.114	(0.246)
3.pos_acc	-0.116	(0.242)	3.anz_acc	-0.314	(0.302)
4.pos_acc	0.290	(0.202)	4.anz_acc	0.657***	(0.240)
2.eta_new	0.446**	(0.174)	2.eta_new	0.143	(0.195)
3.eta_new	0.511***	(0.188)	3.eta_new	0.0857	(0.215)
4.eta_new	0.661***	(0.188)	4.eta_new	0.156	(0.214)
2.dim_at	-0.242	(0.198)	2.dim_at	-0.298	(0.200)
3.dim_at	-0.0813	(0.189)	3.dim_at	-0.120	(0.191)
4.dim_at	-0.192	(0.188)	4.dim_at	-0.251	(0.191)
2.cun	0.644***	(0.151)	2.cun	0.653***	(0.153)
3.cun	0.970***	(0.330)	3.cun	1.002***	(0.324)
4.cun	0.416**	(0.168)	4.cun	0.444***	(0.168)
5.cun	-0.486*	(0.259)	5.cun	-0.367	(0.255)
6.cun	0.113	(0.282)	6.cun	0.330	(0.281)
7.cun	-0.397**	(0.183)	7.cun	-0.304*	(0.179)
8.cun	-0.144	(0.320)	8.cun	-0.0543	(0.316)
Constant	-1.662***	(0.281)	Constant	-1.755***	(0.312)
Observations	2,264		Observations	2,264	
Robust standard errors in parentheses			Robust standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1			*** p<0.01, ** p<0.05, * p<0.1		

Table 5: Outcome of the computation of a logit model on the effects of the introduction of VQR on researchers' research agenda setting

Thirdly, from both Model 1 and 2 it emerges that dimension of the university of afference has not proven to be significant (since the estimated coefficients are negative).

Lastly, both from Model I and 2 it emerges that the disciplinary field (CUN area[76]) has proven to play an important role for the understanding of researchers' research agenda setting related choices, particularly for researchers afferent to the 13 CUN areas. Indeed, the estimated coefficients of 13/B, 13/C and 13/D CUN areas are both positive and statistically significant. We can thus affirm that researchers from 13 CUN areas, especially from Business Economics and Economics History, tend more likely to change their research agenda setting related choices as a result of the introduction of the VQR. On the other hand, researchers from 14 CUN areas tend considerably less likely to change their research agenda setting related choices because of VQR.

Therefore, from the analysis of Model 1 and 2, we can affirm that the gender, the academic seniority and the afferent disciplinary field explanatory variables represents all good predictors of researchers' tendency to change their research agenda setting process as a result of the introduction of VQR.

More precisely, the outcome from Model 1 and 2 suggests that:

- female researchers tend more likely to change their research agenda setting than male researchers,
- older and more expert researchers tend more likely to change their research agenda setting, in comparison to younger and less expert colleagues,
- and, at last, researchers afferent to 13 CUN area tend more likely to change their research agenda setting.

Hence, relying on the evidence emerged from this first phase of the analysis, we can affirm that the results of the research obtained so far are quite aligned to those emerged from other studies, especially for what it concerns the different behavioural pattern between male and female researchers. Indeed, the female researchers' more convergent behaviour and their major propensity to comply with their afferent institution's norms, as already emerged in other studies (D'Amico, Vermigli, Canetto, 2011,

[76] *2.cun* stands for the 13/B CUN area, namely Business economics, *3.cun* stands for 13/C CUN area, namely Economic History, *4.cun* stands for 13/D CUN area, Statistics, *5.cun* stands for 14/A CUN area, namely Political Theory, *6.cun* stands for 14/B CUN area, namely Political History, *7.cun* stands for 14/C CUN area, namely Sociology, and *8.cun* stands for 14/D CUN area, namely Applied Sociology.

Santos, Horta, Amâncio, 2021), can be quite easily explained considering the renowned gender disparities in higher education, as widely demonstrated by a large body of literature (Scott Metcalfe and Padilla González, 2013, Walker, 2018, Cullen, Luna, 1993). A more consistent behavioural pattern with the university institutions' norms and values, as a matter of fact, would increase female researchers' chances to get the tenure and move up in the academic hierarchy.

As concerns the second point emerged from the analysis, the higher propensity of older and more expert researchers to change their research agenda setting due to the VQR may be explained by the role senior researchers play in the academic organisation. Indeed, since they tend to be more involved in the organisational and financial process by the head of the departments (assignment of grants and funds according to the VQR results, creation of job positions, etc.), we could affirm that they tend to be more encouraged making their department have high score in the VQR evaluation process.

Lastly, researchers afferent to 13 CUN more propensity to change their research agenda setting process as the result of the VQR implementation is in line with the hypotheses that we advanced earlier, namely, on the relative more propensity of researchers from empirical data-based disciplines to comply with academic institutions' norms and standards (Lamont, 2009, Bonaccorsi, 2015). Moreover, this specific aspect in the Italian academic context may also be furthered by the issue that 13 CUN evaluation mechanism, even it is a so-called non-bibliometric sector (Baccini, 2014), can be considered hybrid, since it rests on the double use of pee review and bibliometric indicators, in contrast with the evaluation method adopted by GEV 14. Hence, the use of bibliometric indicators may explain the relatively more influence of VQR implementation on 13 area researchers' research agenda setting process.

As described in the previous chapter, the survey comprised a further section, which was filtered by the question on the effects of VQR on research agenda setting, which asked respondents to express themselves on the effects of the introduction of VQR on the four elements of which the agenda setting is composed of, namely, the choice of topics to research, the motivation for conducting research, research practices, and, at last, the choice of the peers community of afference.

We are hence going to show how the introduction of VQR has changed researchers' research agenda setting, considering, on the one hand, the four elements composing the research agenda, and, on the other hand, the two independent variables that are good predictors (according to Model 1 and 2) for understanding whether researchers have changed their research agenda, namely academic seniority, and disciplinary field.

The first research agenda setting component we are going to analyse is the researchers' choice of topics to study.

According to Fig. 21 below, indeed, it emerges that researchers with at least 11 years of academic experience represent the subgroup that has more considerably (20,5%) changed the topics to research, as the result of the introduction of VQR. On the other hand, by comparing the others subgroups (fixed-term researchers, researchers with maximum 5 years of academic experience, and researchers with an academic experience from 6 to 10 years), we can state that there are the percentage of people affirming that they have changed their research agenda topics because of the introduction of VQR is substantially lower (around 13%/14%), and homogenous among the subgroups.

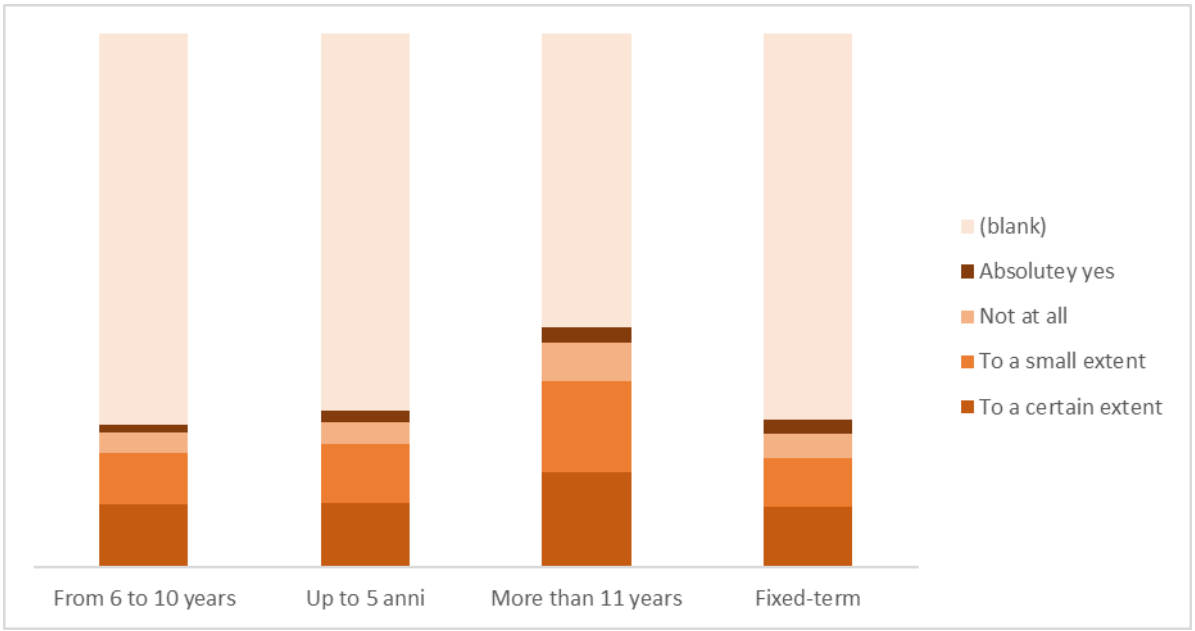


Figure 21: Percentage distribution of responses on changed choice of topics by researchers' academic seniority

According to Fig. 22 (below), if we consider researchers' disciplinary field, we can affirm that researchers afferent to the economic disciplinary sector have generally changed the topics on which conducting research, after the introduction of VQR. More specifically, it emerges that 13/B and 13/C CUN areas (namely, business economics, 27,76%, and economic history, 20%) have recorder the highest percentage of researchers confirming an influence on the selection of topics, followed by researchers afferent to 13/D CUN area (statistics, with the 18%), and 13/A CUN area (economics, with the 16,93%).

On the other hand, the disciplinary fields that recorded the lowest percentage of researchers affirming to have changed the selection of topics on which conducting research are researchers afferent to applied sociology (14/D, only 11,69% of respondents somewhat agreeing to the survey question on the influence of VQR on the selection of topics to research and, interestingly, none totally agreeing with the survey question), and researchers afferent to the 14/A CUN area (political theory, with 7% of respondents affirming to somewhat agree with the survey question, and 1,42% of respondents completely agree with the survey question).

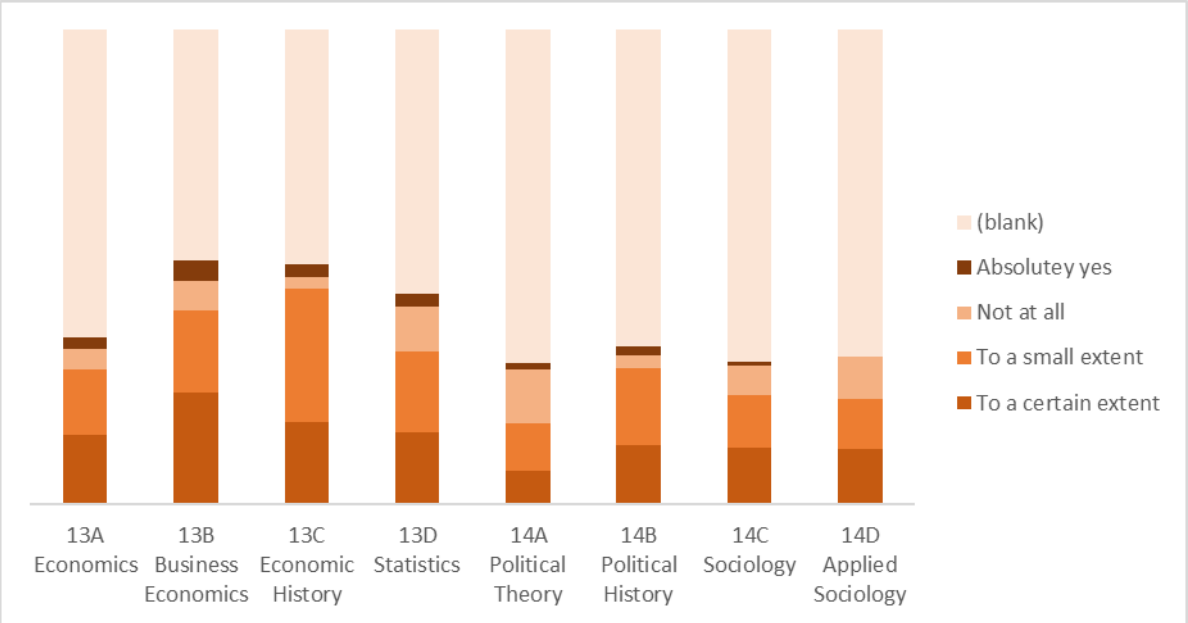


Figure 22: Percentage distribution of responses on influence on the choice of topics by researchers' disciplinary field

The second research agenda setting element we are going to analyse is the motivation for conducting research. According to Fig. 23, it emerges that researchers with more than 11 years of academic experience have changed more substantially their motivation for conducting research, with 16,58% of respondents somewhat agreeing to the survey question asking them whether their motivation has been influenced by the introduction of VQR, and 2,52% of respondents totally agreeing with the question. Researchers with more than 11 years of experience are then followed by research with 5 years of experience maximum, with 11,19% respondents somewhat agreeing with the survey question, and 1,44% respondents totally agreeing. On the other hand, the subgroup of researchers with an academic experience from 6 to 10 years is the one with the lowest percentage of respondents who affirmed to have experienced an influence of the introduction of VQR on their motivation for conducting research.

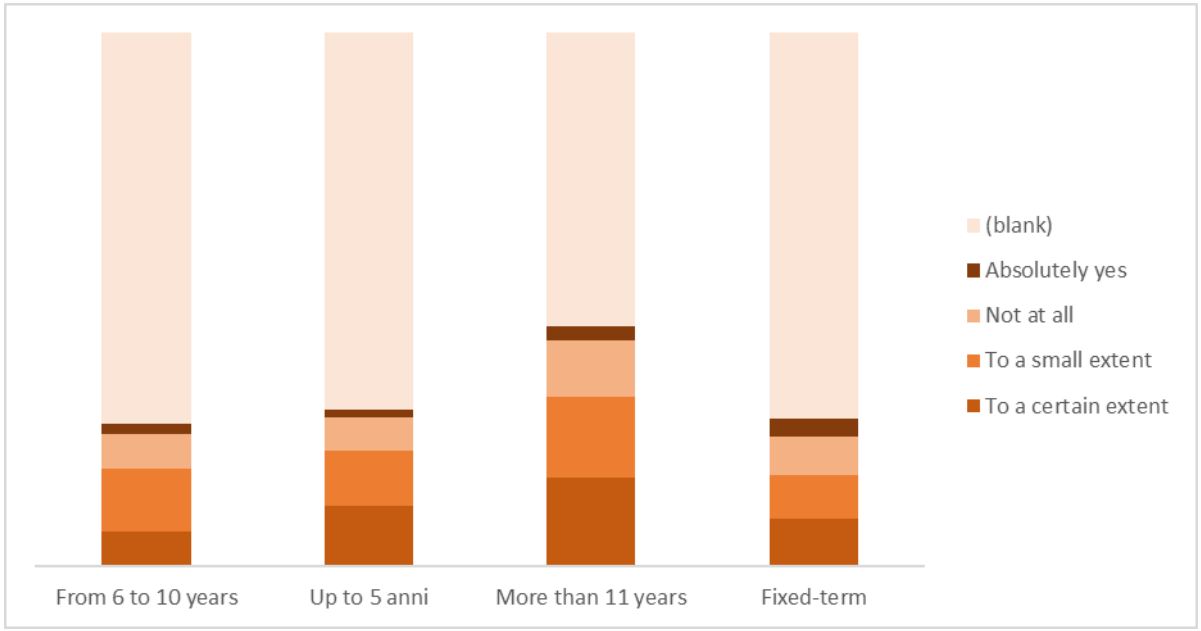


Figure 23: Percentage distribution of responses on influence on motivation for research per academic seniority

Moreover, from Fig. 24, which considers the influence of the introduction of VQR on researchers' motivation, per disciplinary field, it emerges that researchers afferent to 13/B are 13/C CUN area have more considerably experienced a change in their motivation for conducting research, followed

by 13/D, 13/A and 14/B CUN areas. On the other hand, 14/D and 14/A CUN areas are the subgroup with the lowest percentage of researchers affirming to have experienced a change in their motivation.

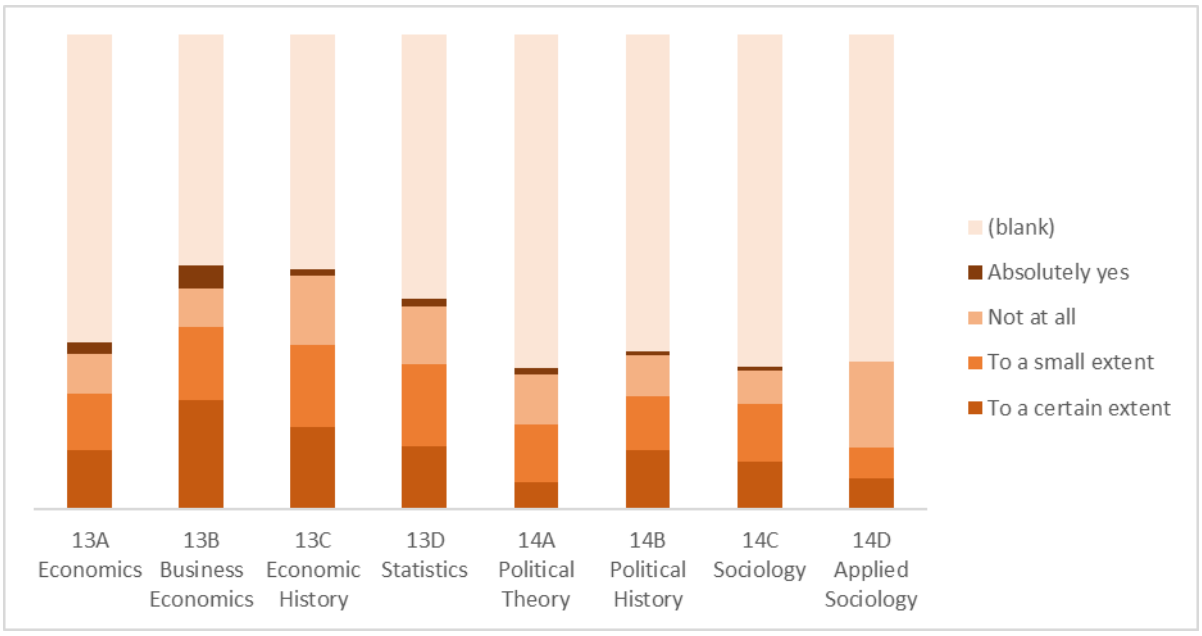


Figure 24: Percentage distribution of responses on influence on motivation for research per disciplinary field

The third research agenda setting component we are going to analyse are the research practices.

The most expert researchers (11 years of experience or more) are the subgroup with the highest percentage of researchers affirming to have changed research practices after the introduction of VQR (Fig. 25 below). More precisely, 18,18% of respondents somewhat agree with the survey question on the influence of VQR on their research practices, while 3,56% totally agree with the survey question. Interestingly, the most expert academics are followed by fixed-term research, with 10% of researchers somewhat agreeing with the survey question, and 3,33% totally agreeing with the survey question. On the other hand, we can affirm that the subgroup to which researchers with maximum 5 years of academic experience are afferent to, is the one with the lowest percentage of respondents claiming to have changed their research practices (7,58% somewhat agreeing to the survey question, 2,89% totally agreeing).

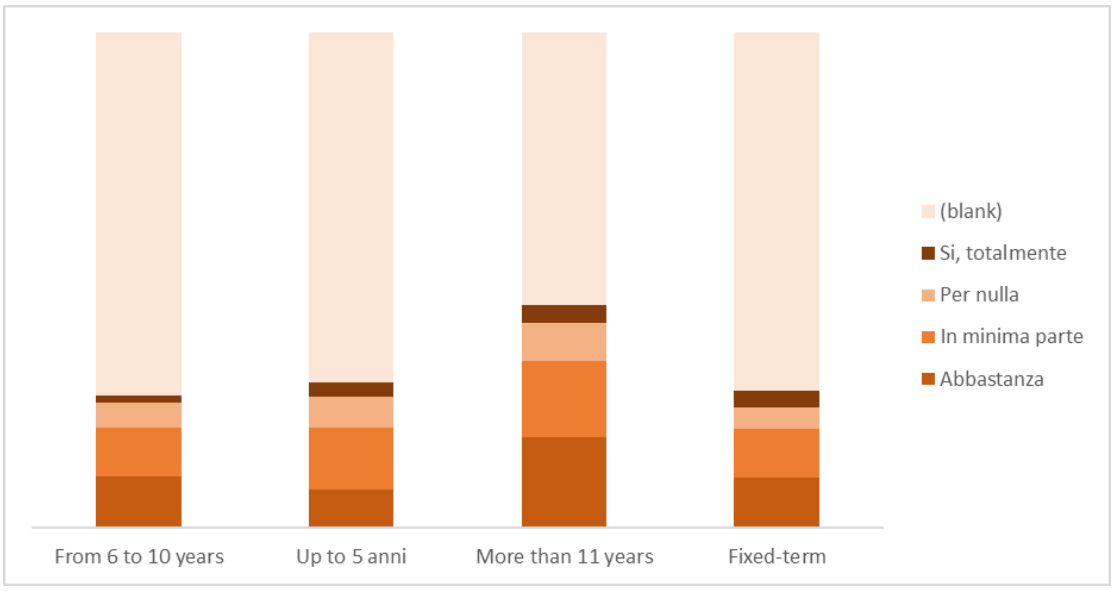


Figure 25: Percentage distribution of responses on influence on the choice of research practices for research per academic seniority

Fig. 26 (below) confirms the previously collected data on the influence of VQR on the research agenda setting dimensions, since researchers afferent to 13 CUN areas generally affirms to have changed more their research practices after the introduction of VQR, in comparison to researchers afferent to 13 CUN areas, with yet the lowest percentage of people affirming to have changed their research practices in correspondence of 14/D CUN area.

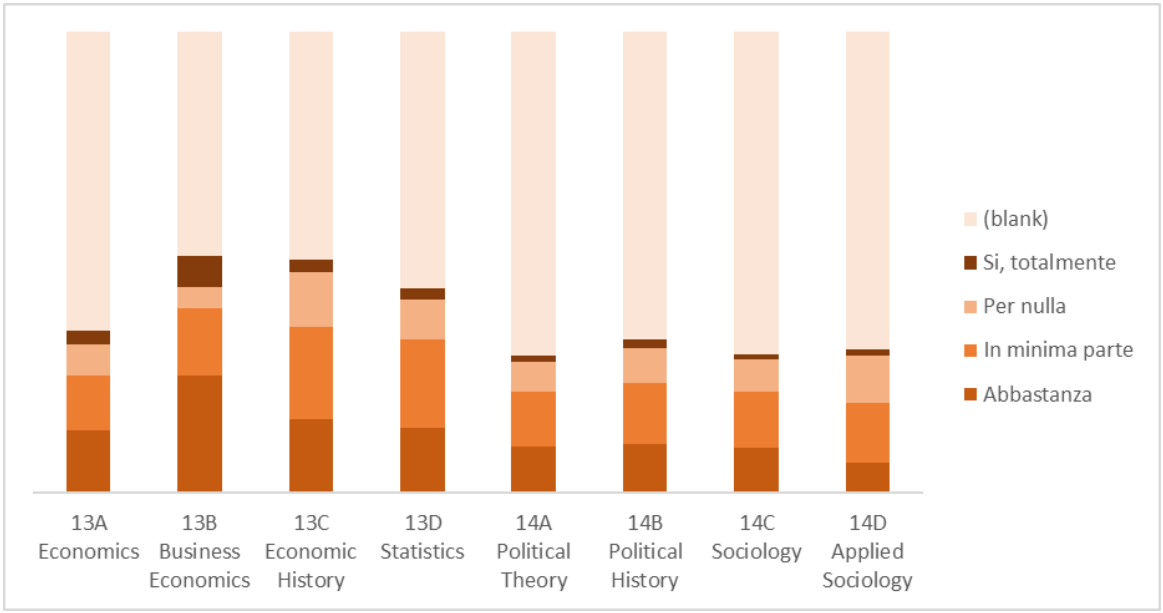


Figure 26: Percentage distribution of responses on influence on the choice of research practices for research per disciplinary field

The fourth research agenda setting component we are going to analyse is the researchers' choice of peers' community of afference.

Hence, Fig. 27 (below) confirm the trend previously seen, namely that the most expert researchers (with at least 11 years of academic experience and researchers with an academic experience from 6 to 10 years) are the ones who have more significantly changed their peers' community of afference, after the introduction of VQR.

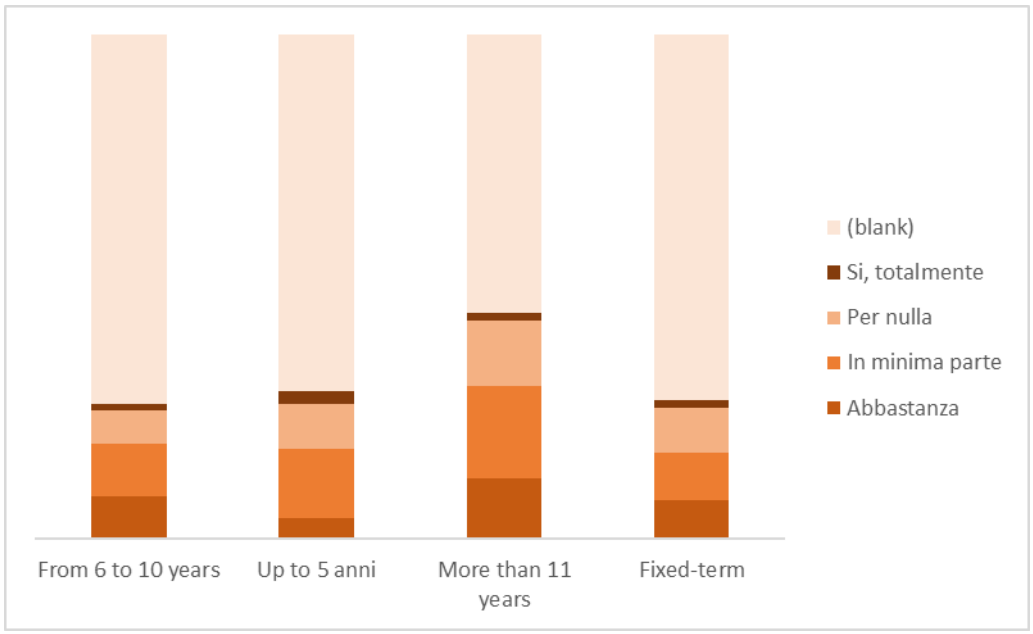


Figure 27: Percentage distribution of responses on influence on the choice of peer's community for research per academic seniority

Finally, from Fig. 28 (below) it emerges a soft change in the trend previously emerged. Indeed, researchers afferent to 13/B and 13/C CUN areas are followed by researchers from 14/B (political history), with 10,48% of respondents somewhat agreeing to the survey question asking them whether they changed their peers' community after the introduction of VQR, and 1,9% totally agreeing to the survey question.

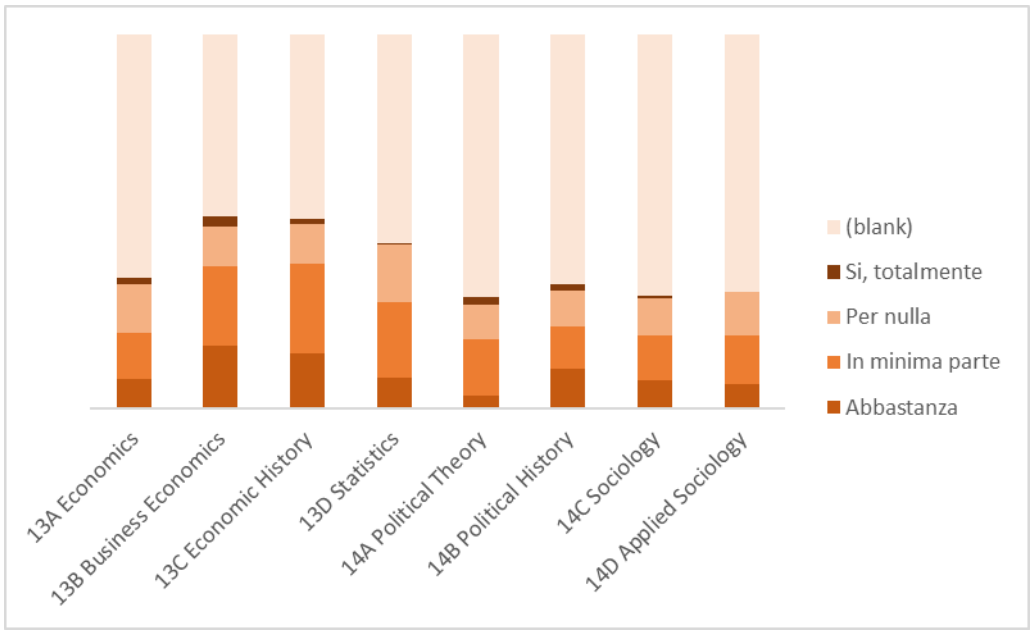


Figure 28: Percentage distribution of responses on influence on the choice of peers' community for research per disciplinary field

To sum up, according the responses collected with the survey and the analysis conducted, we can reply to our first research question and affirm that the introduction of the VQR has not produced relevant effects on the research agenda of the majority of the researchers surveyed. We can nevertheless affirm that the introduction of VQR has produced some effects on more than one third of the surveyed population, which is noteworthy, considering that the institutional pressure on researchers was indirect and mediated and intermediated by universities, which are actually the subjects under evaluation. This relatively remarkable effect on researchers' research agenda setting process may be explained by two factors: first of all, that the Italian research evaluation system can be considered *strong* (according to Whitley's research evaluation systems categorization, as seen in the first chapter), since it is characterized by high standardization and transparency, which generally tend to encourage the shaping of more homogeneous and standardized scientific communities (Whitley, 2007); on the other hand, the strict link between VQR outcomes and resources allocation may have strengthened its effects, due to a pressure exerted by universities elites on researchers, with the purpose of ensuring to the afferent university a good score in the national universities rankings.

Interestingly, senior researchers' research agenda setting process, especially from economic CUN areas, seem to have been more flexible and susceptible to the changed introduced by the VQR mechanism. Indeed, according to the evidence from the survey questions on the effects on the determinants of research agenda setting, senior researchers affirm to have changed their research practices, the selection of their research topics and their motivation for conducting research relatively more in comparison to younger fellows. We might accordingly presume that senior researchers have modified relatively more their research agenda setting related habits for the role of responsibility they have within the university hierarchical system, and for being tasked with comply with the benchmark required for the evaluation.

To conclude, we can reply to our third research question from the information collected so far, and affirm that economic sector has been affected more, in comparison to sociology and political analysis. This outcome is in line with the thesis advanced by Bonaccorsi (2015) and Lamonte (2009), who affirmed that disciplines with a strongly based on empirical evidence tend relatively more to adapt to the evaluation mechanisms, in comparison to disciplines focussed more on theoretical debate.

4.3.2.2 ASN effects on research agenda setting

According to the data collected from the question asking respondents whether they have changed their research agenda setting because of the introduction of the ASN, the majority of respondents have replied positively to the survey question. More precisely, 50,15% of respondents replies that they changed their research agenda setting after the introduction of ASN, while 49,85%, on the other hand, replied that they have not changed their research agenda setting (Fig. 29).

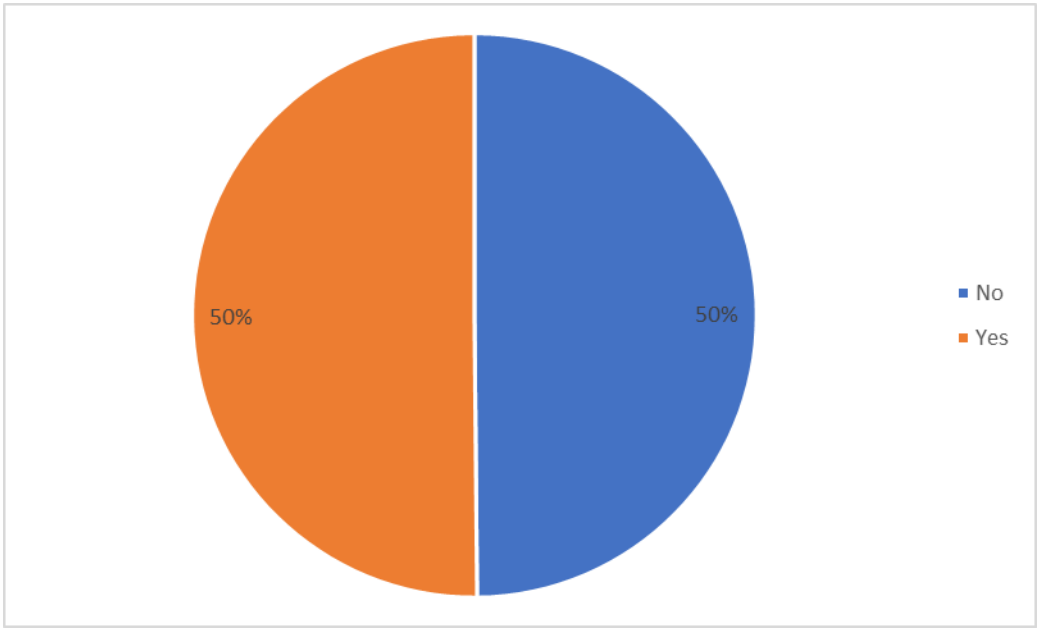


Figure 29: Percentage distribution of survey question asking respondents whether they changed their research agenda after ASN

From the data, it emerges that the direct institutional pressure exerted by the ASN has produced more relevant effects on researchers' research agenda setting, in comparison to the indirect institutional pressure exerted by the VQR.

As for the effects of VQR, even for the computation of the effects of the introduction of ASN on researchers' research agenda setting we used the logit model. As previously described, we had to insert an additional variable into the dataset to weight the different sizes of the sample, to counterbalance the eventual selection bias.

In the computation of the model, the dependent variable was the presence/absence of the effects on research agenda after the introduction of ASN, and the independent variables where the same as for what done in the VQR logit model, namely: academic position, gender, age, academic seniority, dimension of the afferent university, and disciplinary field, namely, the CUN area. Even in this case we run two models, because of the collinearity between the academic seniority and the academic position independent variables. Hence, in the first model (Model 1, on the left column, Tab. 6) we excluded the academic seniority from the computation, while in the second model (Model 2, on the right column) we excluded the academic position.

By considering both Model 1 and 2, it emerges that the coefficients referring to the variable gender (*gen*) are positive and statistically significant. For this reason, we can thus affirm that female researchers tend more likely than male researchers to change their research agenda setting, as a result of the introduction of ASN.

Moreover, the coefficients of the variables academic position (*pos_acc*) are positive and statistically significant in Model 1. Still, the "lower" the academic position, the higher the coefficient. Hence, we can thus affirm that younger researchers tend considerably more to change their research agenda setting behaviour, in comparison to older researchers. This outcome is consistent with the outcome that emerges from Model 2, where the coefficients of the academic seniority (*anz_acc*) referring to expert researchers are negative. For this reason, we can affirm that expert researchers tend definitely not to change their research agenda setting behaviour because of the introduction of ASN.

For what it concerns the disciplinary field (variable *n_cun*), the two models present slightly contrasting results. Indeed, on the one hand, from Model 1 it emerges that 13 CUN area and political history present positive and statistically significant coefficients, suggesting that researchers pertaining to these disciplinary field tend more likely to change their research agenda because of the introduction of ASN. On the other hand, from Model 2 it emerges that even 14/C CUN area (sociology) present a positive and statistically significant coefficient (even if not high, namely 0,348), in addition to the disciplinary fields that present positive and statistically significant coefficients in Model 1. Hence, we can say that excluding the academic position from the computation, even sociologists only slightly tend to change their research agenda setting because of the introduction of ASN.

To sum up, the data suggest that:

- female researchers tend more likely to change their research agenda setting than male researchers,
- lower position researchers tend more likely to change their research agenda setting, in comparison to more expert colleagues,

- and, at last, researchers afferent to 13 and 14/B CUN areas tend more likely to change their research agenda setting.

VARIABLES (-anz_acc)	ASN_infl_ra		VARIABLES (-pos_acc)	ASN_infl_ra	
	Coefficients	Standard error		Coefficients	Standard error
gen	0.400***	(0.121)	gen	0.565***	(0.119)
2.pos_acc	0.814***	(0.142)	2.anz_acc	-0.567**	(0.243)
3.pos_acc	1.323***	(0.243)	3.anz_acc	-1.086***	(0.277)
4.pos_acc	0.913***	(0.194)	4.anz_acc	-0.465*	(0.241)
2.eta_new	0.337*	(0.175)	2.eta_new	0.0787	(0.193)
3.eta_new	-0.0811	(0.186)	3.eta_new	-0.489**	(0.212)
4.eta_new	-0.555***	(0.188)	4.eta_new	-1.131***	(0.216)
2.dim_at	0.000769	(0.196)	2.dim_at	0.0126	(0.196)
3.dim_at	0.215	(0.191)	3.dim_at	0.242	(0.190)
4.dim_at	-0.166	(0.189)	4.dim_at	-0.190	(0.189)
2.cun	0.710***	(0.156)	2.cun	0.760***	(0.154)
3.cun	0.981***	(0.339)	3.cun	1.024***	(0.318)
4.cun	0.592***	(0.177)	4.cun	0.680***	(0.174)
5.cun	0.0298	(0.263)	5.cun	0.262	(0.258)
6.cun	0.553*	(0.295)	6.cun	0.909***	(0.297)
7.cun	0.101	(0.181)	7.cun	0.348*	(0.179)
8.cun	0.0656	(0.319)	8.cun	0.349	(0.320)
Constant	-1.382***	(0.276)	Constant	-0.319	(0.293)
Observations	2,264		Observations	2,264	
Robust standard errors in parentheses			Robust standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1			*** p<0.01, ** p<0.05, * p<0.1		

Table 6: Outcome of the computation of a logit model on the effects of the introduction of ASN on researchers' research agenda setting

For what it concerns the first point emerged from this second phase of the analysis, some considerations are in line with what expressed earlier on the relative more difficulties female researchers encounter in having a career in the academic sector. This issue is even stronger in this specific context, considering the direct effects that ASN results have on researchers' career, and, more importantly, considering the lower probability female researchers have to get scientific habilitation

(Marini, Meschitti, 2018). This issue is relevant in explaining also the reason why younger researchers' research agenda setting has been more influenced by ASN implementation, in comparison to older colleagues. Indeed, the precariousness that is inherent to academic profession in its early stage in Italy (Bozzon, Murgia, Poggio, Rapetti, 2017), may represent a very strong drive to have a more convergent research agenda setting approach.

As a matter of fact, the relative more propensity to adopt a convergent research agenda approach of researchers afferent to 13 and 14/B (applied sociology) CUN area can be explained for the relative importance of empirical data in these specific disciplines (Bonaccorsi, 2015, Lamont, 2009, Whitley, 2007).

In this section, we are going to present how the introduction of the ASN has determined a change on the components of research agenda setting. Indeed, we are going to consider respondents' replies to the questions referring to the ASN additional section, asking them whether the introduction of the ASN has induced them change their research topics, their motivation for conducting research, their research practices, and, lastly, their peers' community of reference. We are going to consider the independent variables that proven significant in the logit models, namely, academic position and disciplinary field.

The first research agenda setting component that we are going to analyse is the choice of topics.

According to Fig. 30, fixed-term researchers are the subgroup of researchers that affirm to adapt more the choice of topics to research, because of ASN (33,33% of respondents somewhat agree with the survey question on the influence of ASN on the choice of topics, and 8,33% totally agree).

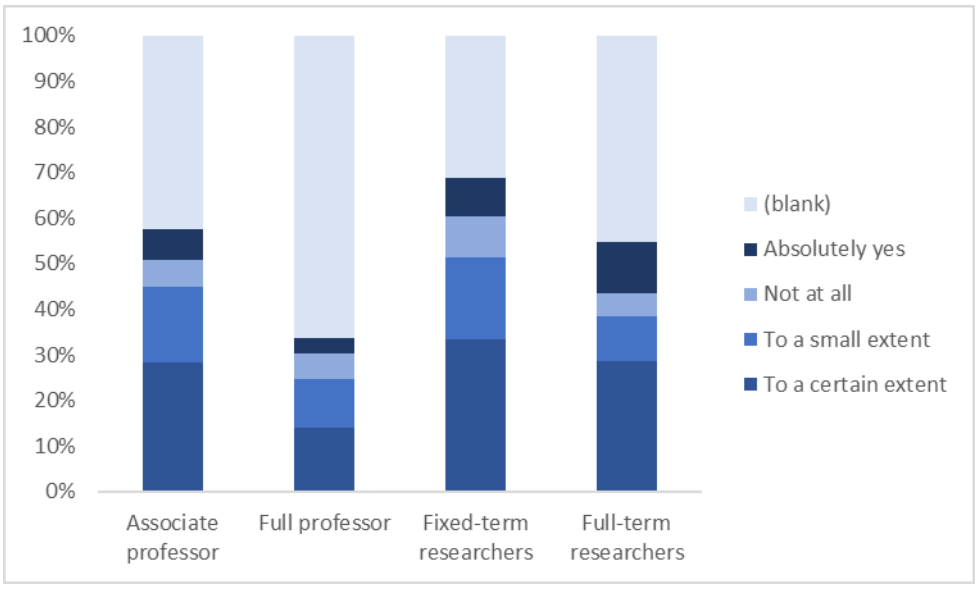


Fig. 30: Percentage distribution of responses on influence on the choice of topics for research per academic role

This outcome is understandable, since getting the habilitation is fundamental (but not enough, as previously said) for getting the tenure. The outcome recorded for associate professors and full-term researchers are quite similar, with full-term researchers higher percentage than associate (11,21% *versus* 6,73%) of respondents replying that they changed their topics to research because of ASN. This outcome could be easily explained by the fact that getting the habilitation for fixed-term researchers[77] represent the only way out to a dead-end street career, with no other possibility of career advancement. The outcome that emerges from the full professors category is quite understandable, since this researchers subgroup is even composed by researchers who got the tenure and became full professors afterwards before the implementation of Act of 2010, no. 240.

[77] Fixed-term researchers represent a professional academic category that no longer exists, after the implementation of Act of 2010, no. 240.

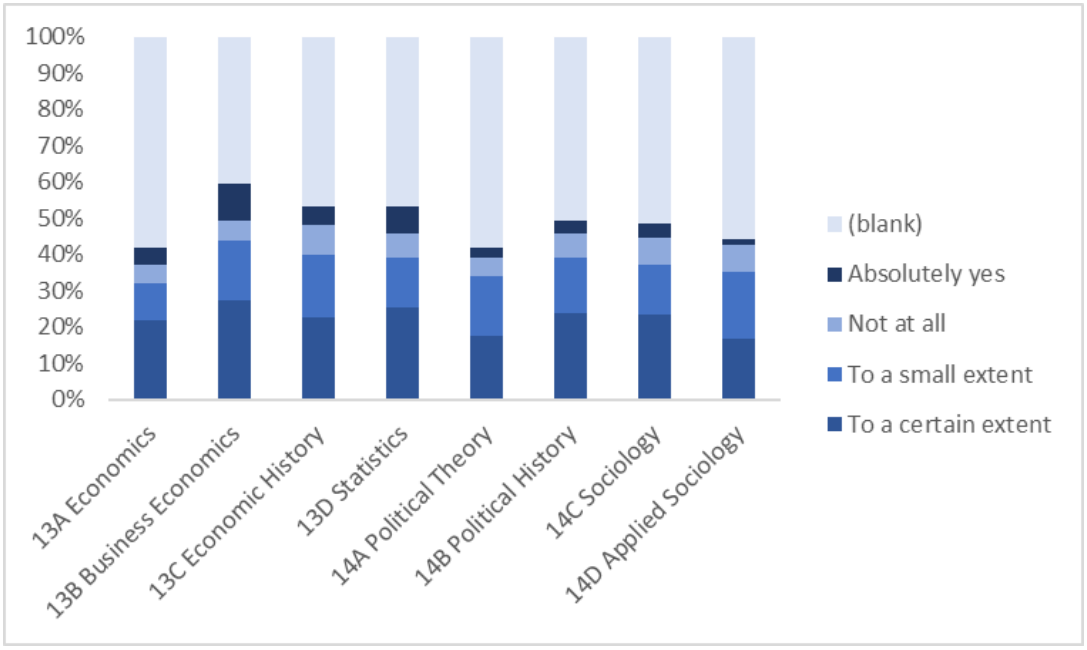


Fig. 31: Percentage distribution of responses on influence on the choice of topics for research per disciplinary field

From Fig. 31, it emerges that the percentage of respondents who somewhat agree to the survey question (on whether they changed the topics on which to conduct research) is generally high, varying from 16,88% (14/D CUN area) to 27,26% (13/B CUN area). This may explain, actually, the relative high logit parameters coefficients, in comparison to VQR coefficients. However, considering the percentage of respondents who totally agree with the survey question, we can see that 13/B, 13/C and 13/D present the higher values. We can thus affirm that researchers from economic sector have more considerably changed the topics to research, after the introduction of ASN.

The second research agenda setting we are going to consider is the motivation for conducting research. According to Fig. 32, which represents the percentage distribution of responses on the influence of ASN on the motivation for conducting research, it emerges, as before, that fixed-term researchers have changed their motivation for conducting research considerably more than other academic positions subgroups. Indeed, 26,67% of respondents somewhat agree with the survey question response on the question on the influence of ASN on the motivation of conducting research, while 8,75% totally agreed. Fixed-term researchers are followed by fixed-term researchers (33,18%), and associate professors (29,13%).

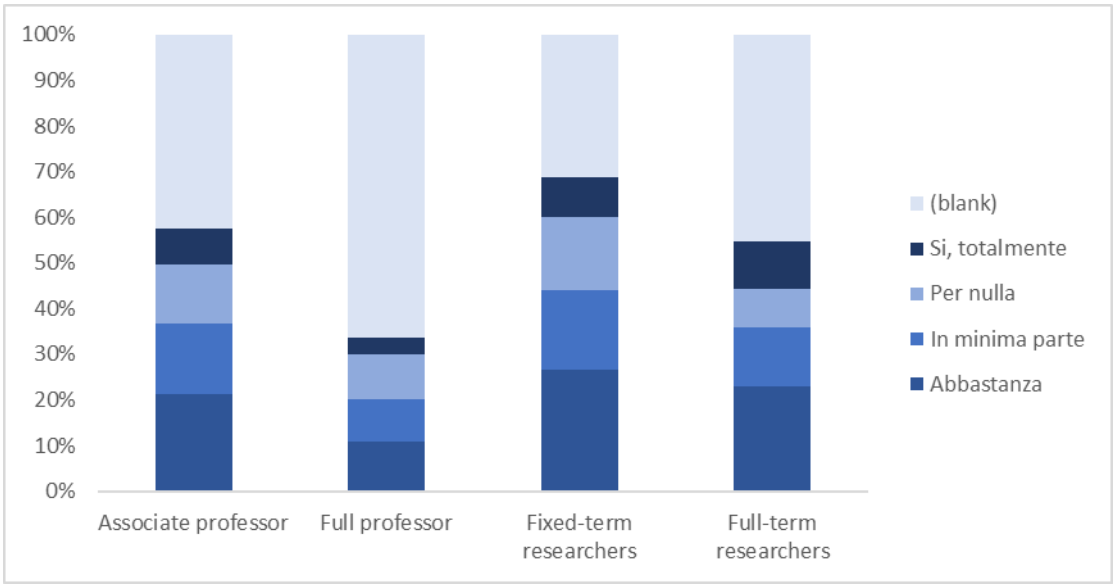


Fig. 32: Percentage distribution of responses on influence on the motivation for conducting research per academic role

Furthermore, Fig. 33 shows that change induced by the introduction of ASN on the motivation for conducting research per disciplinary areas. It emerges from the available data that the influence of ASN is not homogeneous among all the eight CUN areas at issue, in comparison to the effects of ASN on the choice of topics to research. However, from 13/A to 14/C CUN areas the data show that a relative high number of respondents somewhat agree with the survey question on the influence of the introduction of ASN on the motivation for conducting research. On the other hand, 13/D CUN area has recorded the lowest number of respondents who affirm to have changed their motivation for conducting research.

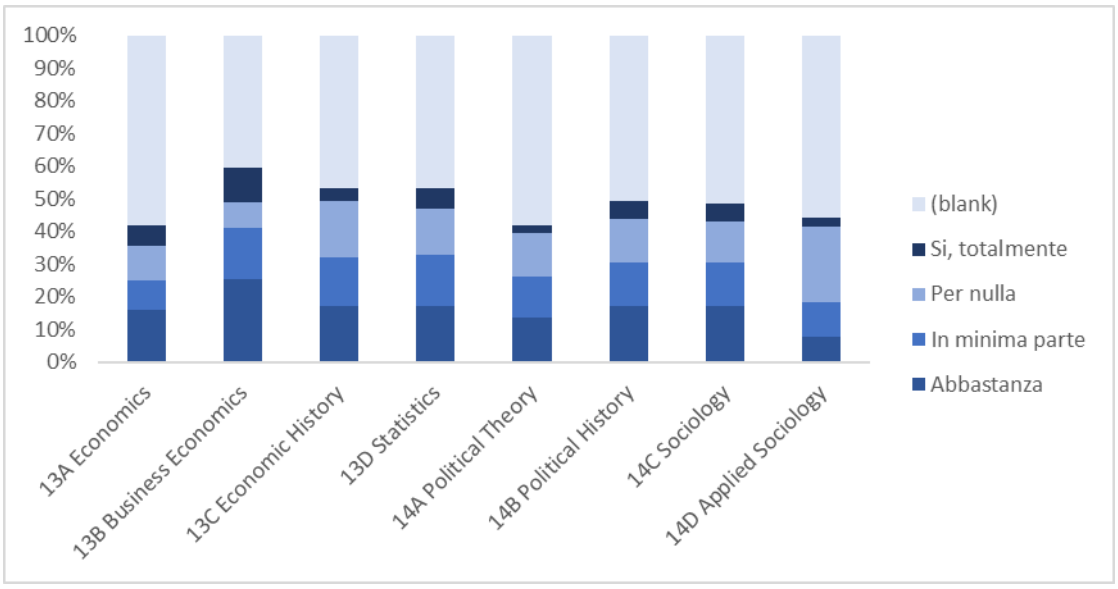


Fig. 33: Percentage distribution of responses on influence on the motivation for conducting research per disciplinary field

The third research agenda setting component we are going to analyse are the research practices. The data suggest, once again, that fixed-term researchers (Fig. 34) have changed more their research practices, after the introduction of ASN. Interestingly, from Fig. 34 it emerges that the outcome recorded for associate professors and fixed-term researchers are somewhat equivalent. Fig. 34 confirms the previous trend regarding full professors, notably that they have not changed their research agenda because of the introduction of ASN.

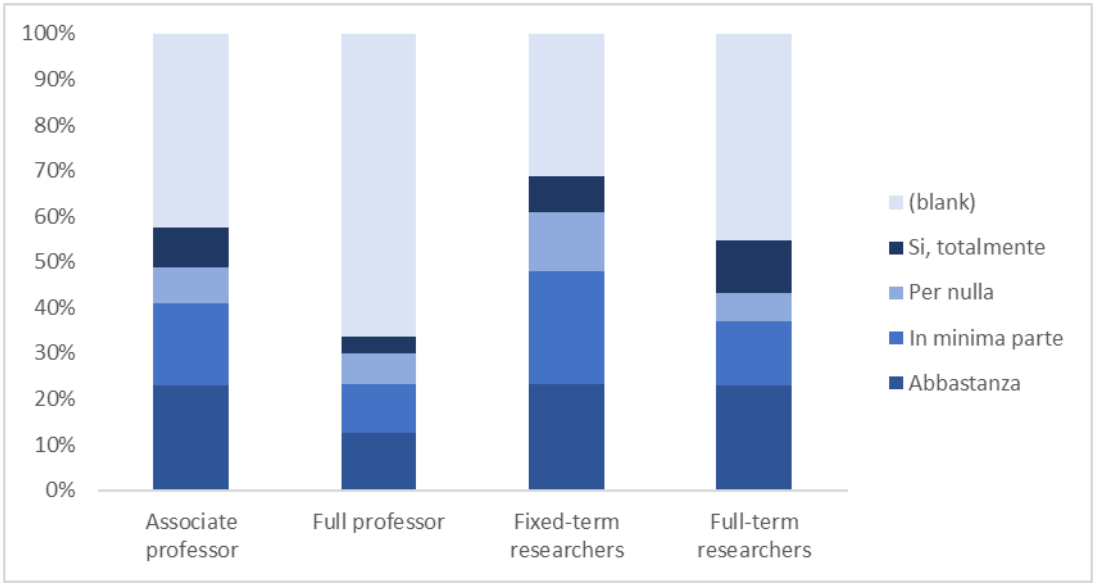


Fig. 34: Percentage distribution of responses on influence on research practices per academic position

Fig. 35, moreover, reports the percentage distribution of responses on influence of ASN on research practices per disciplinary field. According to Fig. 35, it emerges that there is considerable variety among the different disciplinary fields. The data show that 13/B (with 25,24% of respondents somewhat agreeing with the survey question and 13,55% totally agreeing) and 13/D (with 21,05% of respondents somewhat agreeing with the survey question and 7,76% totally agreeing) CUN areas are the subgroups which have recorded the higher amount of respondents affirming to have changed their research practices after the introduction of ASN. Still, 14/A and 14/D CUN areas have recorded the lowest percentage of respondents affirming to have changed their research practices.

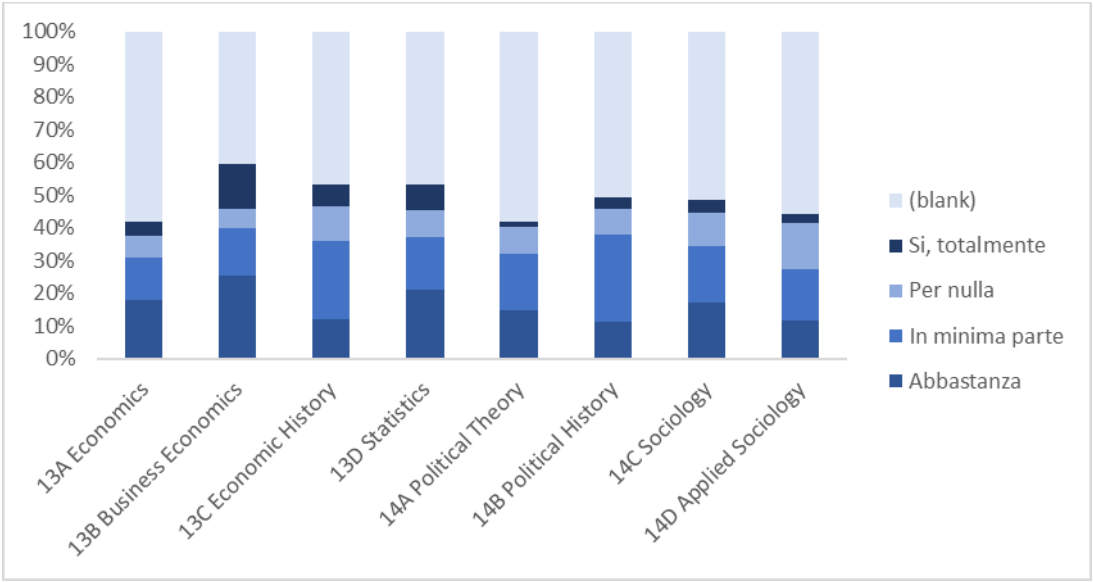


Fig. 35: Percentage distribution of responses on influence on research practices per disciplinary field

The fourth research agenda setting component we are going to analyse is the choice of the afferent peers' community.

Interestingly, according to Fig. 36, that represents the percentage distribution of responses on the influence of ASN on the choice of peers' community per academic role, the outcome recorded by associate professors, fixed-term research and full-term research are somewhat equivalent (summing up the respondents that somewhat agree to the survey question to the respondents that totally agree with the survey question). The previous trend regarding full professors' behaviour is confirmed.

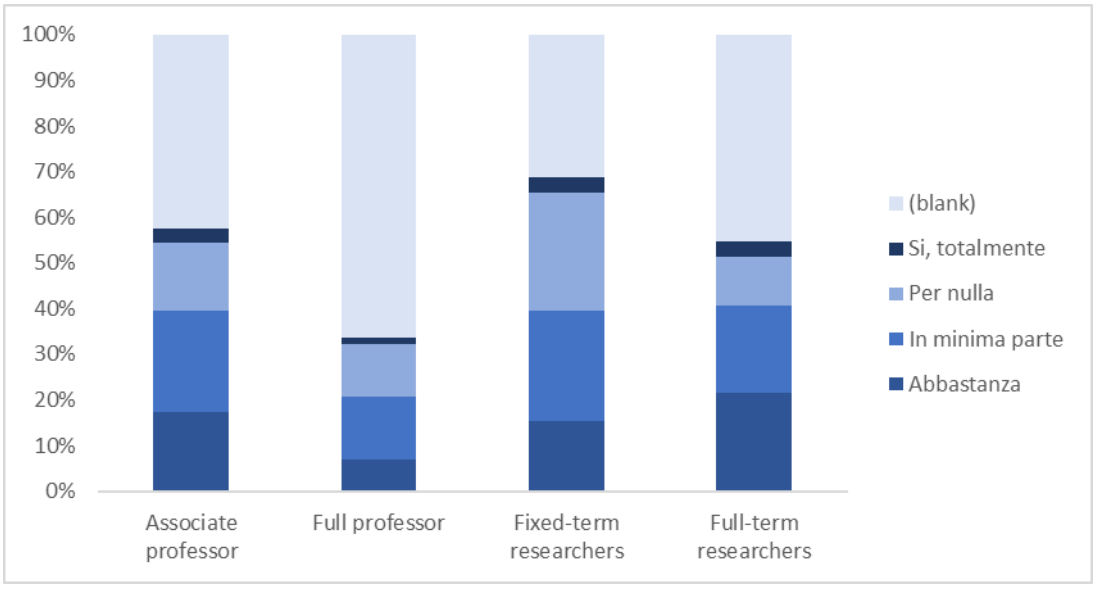


Fig. 36: Percentage distribution of responses on influence on the choice of peers' community per academic position

To conclude the analysis of the effects of the introduction of ASN on research agenda setting components, we analyse at last the percentage distribution of responses on the choice of peers' community per disciplinary field. Interestingly, from Fig. 37 (below), it emerges that only 13/A and 13/B CUN areas recorded a considerable percentage of researchers affirming to have changed their peers' community after the introduction of ASN. The other disciplinary fields, indeed, record lower percentage of respondents affirming to have changed their peers' community.

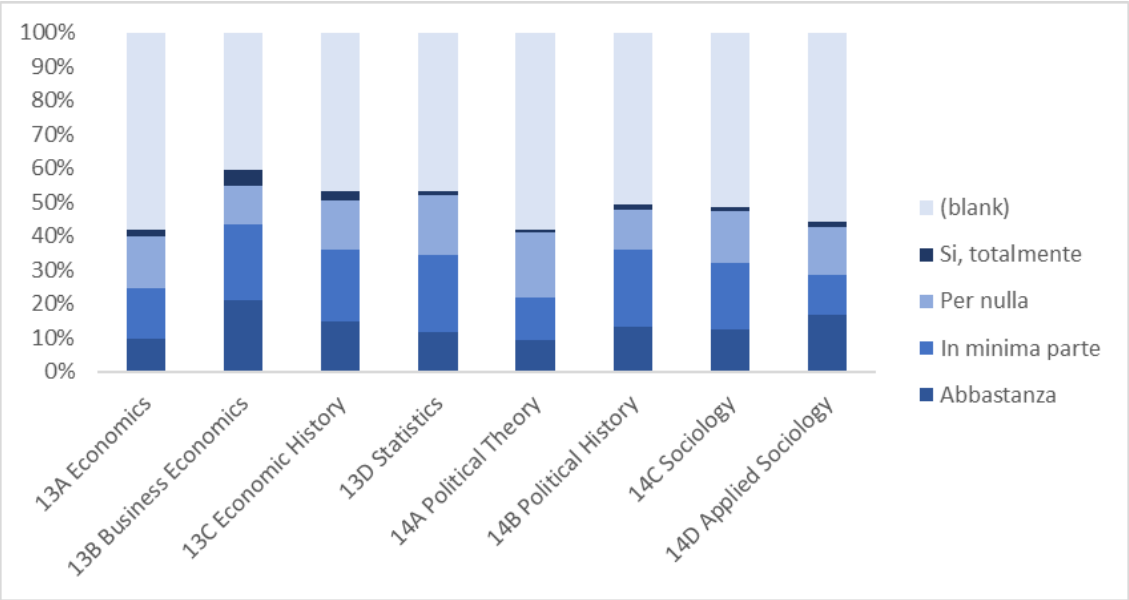


Fig. 37: Percentage distribution of responses on influence on the choice of peers' community per disciplinary field

To sum up, we can affirm that, accordingly, that we can reply to our second research question by saying that the introduction of the ASN has determined a relatively more remarkable effect on researchers' research agenda setting process, since more than half of respondents claimed to have changed their research agenda setting approach due to the introduction of ASN. This effect is most presumably due to the direct institutional pressure that ASN exerts on researchers, for the consequences it entails on researchers' career.

It is very important to remark, moreover, two differences of the effects of ASN on researchers' research agenda setting, in comparison to VQR: firstly, researchers without the habilitation, have remarkably changed their research practices, research topics and motivation, and peers' community, in comparison to senior researchers; secondly, the distinction of the effects of ASN on 13 and CUN areas is less clear-cut, and these changes also involve considerably researchers from sociology and political science.

The most important issue to remark, in our opinion, is that the introduction of ASN has introduced a clear-cut change in Italian academic profession, determining a cleavage between researchers with habilitation and researchers who have not gotten it yet. The ASN, indeed, seems to have introduced a mechanism that strongly affects researchers' behaviours till the moment they get the habilitation,

for which they are required to comply to a series of benchmarks, to "survive" in the academia. This highly competitive mechanism is perceived by researchers themselves as extremely negative for the quality of research production, and the element of strong competitiveness it introduces in the working environment. In conclusion, from the free comments to the ASN survey question, it emerged a *Hunger Games*-like mood, where individuals' efforts are not enough, most of times, for succeeding, since they directly say that the quality of research itself is not a guarantee for getting the tenure.

4.3.2.3 Descriptive analysis: academic collaboration

The survey, as previously mentioned, includes a section with a series of questions on scientific collaboration, starting with a filter question asking respondents whether they seek for scientific collaborations with colleagues. Except for respondents affirming not to seek at all for colleagues to collaborate with (namely, 70 respondents, corresponding to the 3% of the total amount of survey respondents), respondents were addressed to three subsequent questions on their collaboration attitude. Interestingly, the majority of respondents (77,8%, corresponding to the *Strongly agree* and *Agree* response options) affirm to frequently establish partnerships with colleagues (Fig. 38 below).

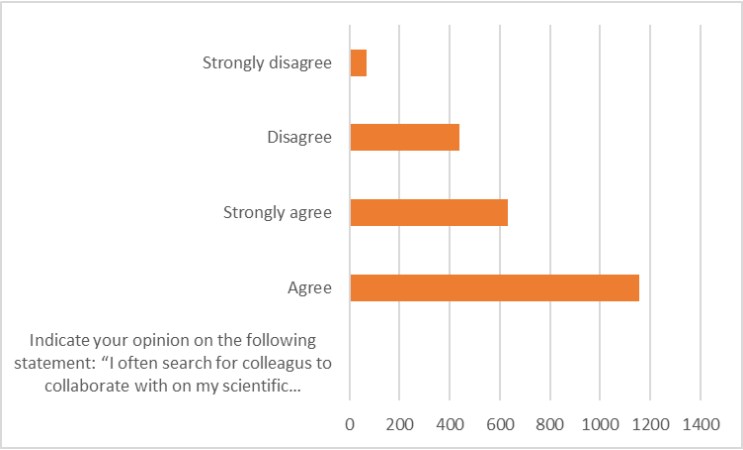


Fig. 38: Respondents' distribution on their academic collaboration attitude

According to the data collected, female researchers (once weighted the responses, since men are more represented than women in the respondent population) tend relatively more to collaborate than male

researchers (see Tab. 7 below). Indeed, for the *Strongly agree* and *Agree* response options women are respectively 1,4 and 1,25 odds willing to collaborate, while men odds of seeking for research collaboration are respectively 0,7 and 0,8.

		Men	Women
Indicate your opinion on the following statement: "I often search for colleagues to collaborate with on my scientific publications"	Strongly agree	0,7	1,41
	Agree	0,8	1,26
	Disagree	0,99	1,02
	Strongly disagree	0,85	1,17

Table 7: Odds of collaboration approach with gender as independent variable

According to Fig. 39, full professors, associate professors, and temporary researchers tend to search for collaborations with colleagues frequently, while, on the other hand, full-term researchers seem to be less interested in scientific collaborations.

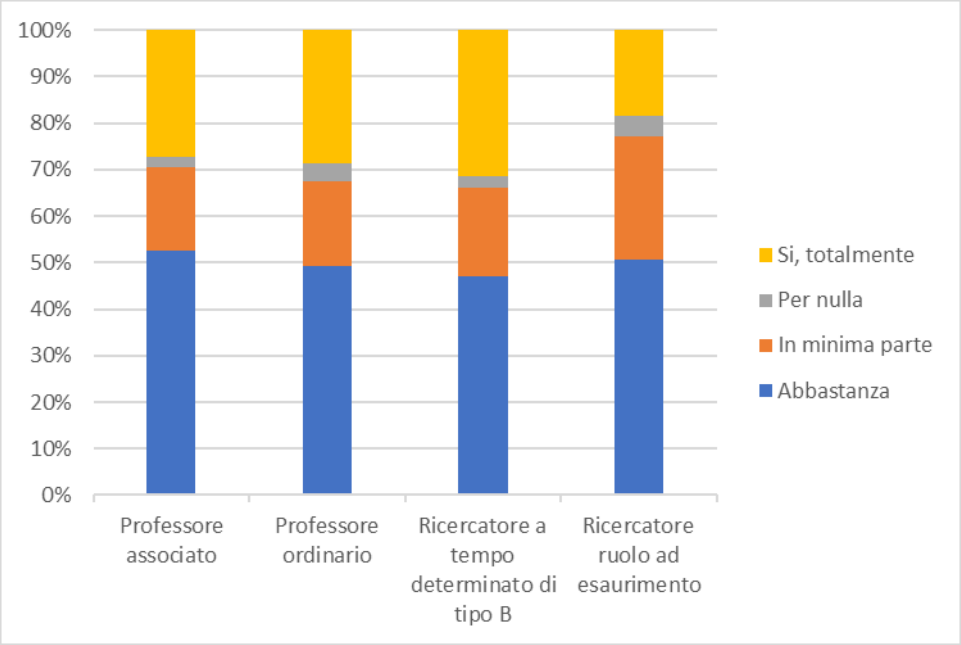


Fig. 39: Respondents' distribution on scientific collaboration per gender and academic position

Furthermore, for what it concerns the effects of collaboration on the research agenda setting with both colleagues from the same disciplinary field and from different disciplinary fields, Fig. 40 and 41 (below) show that researcher's research agenda setting is, in most cases, partially influenced by collaboration with colleagues. This especially with colleagues from the same disciplinary field, as it

is apparent from Fig. 40 below. Indeed, 70% of respondents affirm to have partially changed their research agenda due to the collaboration with colleagues from their same disciplinary field, and, on the other hand, only 24,1% disagree with the proposed statement (Fig. 40), while 57,2% of respondents affirm that collaboration with colleagues from different disciplinary fields has partially changed their research agenda (Fig. 41). This issue represents the most relevant outcome in terms of influence of collaboration on research agenda setting.

This outcome is also confirmed by respondents' responses to the survey questions on whether their research agenda setting has been considerably influenced by collaboration or not influenced at all by collaboration. Indeed, when respondents are asked whether either their research agenda has been influenced by collaboration, or their research agenda has not been influenced at all (both with colleagues from the same disciplinary field and a different disciplinary field), most of the 50% disagree or strongly disagree with the proposed statement (Fig. 40 and 41 below).

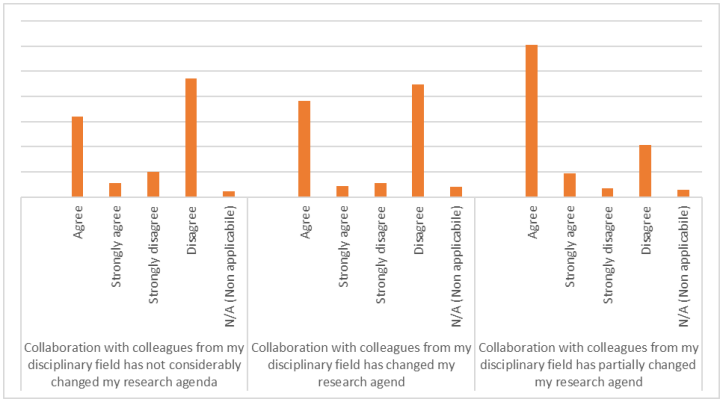


Fig. 40: Influence of collaboration with same-field colleagues on research agenda

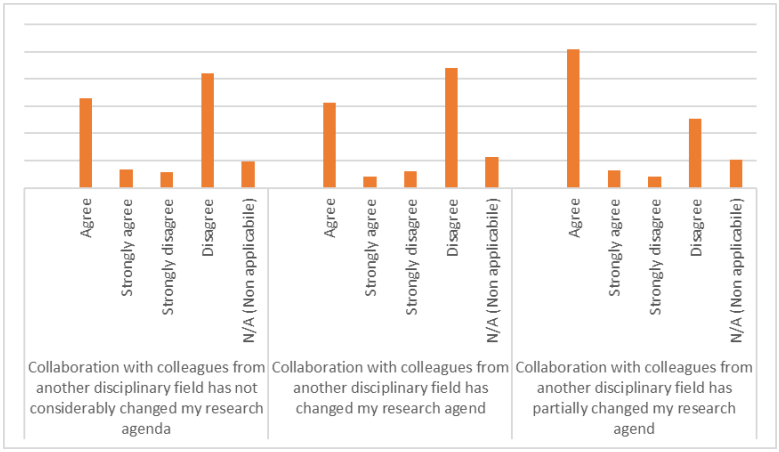


Fig. 41: Influence of collaboration with colleagues from another disciplinary field on research agenda

4.3.2.4 Descriptive analysis: influence of calls for research funding on research agenda

From the survey section referring to calls for research funding, it emerges that 89% of respondents affirm that they have participated to both national and international calls for research funding (Fig. 42). Interestingly, among the full-term academic positions, full-term researchers appear to participate less to calls for research funding, while full professors are the subgroup which affirms to have participated the most to calls for research funding (followed by associate professors and temporary researchers).

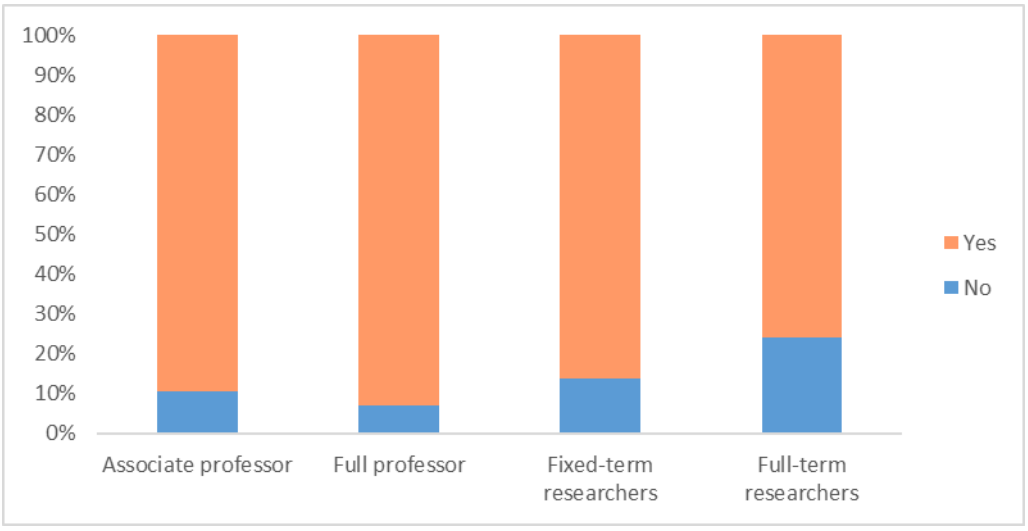


Fig. 42: Percentage distribution of respondents' participation to calls for research funding per academic role

This was a filter question, set to direct respondents who affirmed to have previously participated to a call for research funding to the following section, which was meant to understand whether and how

the participation to calls for research funding influenced researchers' research agenda setting. Those who replied to have never participate to calls for research funding were addressed to the subsequent section of the survey (that is, 253 people, corresponding to the 11% of respondents).

The subsequent question (which is a filter question too, as we will see afterwards) asked respondents whether participating to calls for research funding has influenced their research agenda setting. Respondents' responses are summarized in Fig. 43 below. Interestingly, 62% of respondents affirm they partially integrated the topics addressed in calls for research funding in their research agenda setting process, while 15% affirmed to have totally integrated the calls for research funding topics in their research agenda.

Respondents who affirmed not to have integrated at all calls for research funding topics in their research agenda (23%) were addressed to the following section, while the other respondents were asked to explain how the participation to calls for research funding has influenced their research agenda setting process.

Figure 43: Summary of respondents' answers on eventual influence of calls for funding on research agenda

The respondents who replied that participating to calls for research fund has influenced their research agenda, replied to the question referring to how the participation to calls for research fund influenced

their research agenda as follows: almost 60% (59,56%) replied that they had less time for conducting research on the topics pertaining to their research agenda because of the participation to calls for research fund, while almost 40% of respondents replied that participating to calls for research fund encouraged them to broaden their topics of interest.

4.3.2.5 Other determinants of research agenda setting

The survey comprises a series of questions asking for respondents' opinions on other possible determinants of research agenda setting. As it emerges from the data (Tab. 8 below), topics which may play a relevant role in the public arena the so-called mainstream topics) seem to be relevant to most of respondents in the research agenda setting process. Indeed, 84,8% of respondents affirm that conducting research on mainstream topics would grant high possibility of publishing.

Moreover, news items appear to have influenced the research agenda of 55,1% of respondents, while societal needs are perceived as determinants in shaping research agenda by 75,6% of respondents.

Interestingly, 53,3% of respondents affirm to have structured their research agenda in order to be recognized by their peer community (this concept was phrased differently, namely as being an active part of their peer community, to soften the concept, which may be a sensitive topic for the reference population).

The responses collected in this survey section seem to suggest that the social component of research production process (both the academic environment and society at large) play a relevant role in the research agenda setting process.

I have frequently oriented my researcg agenda for being an active member of my scientific community of reference	Somewhat agree	42,6%
	Strongly disagree	10,4%
	Somewhat disagree	36,3%
	Strongly agree	10,7%
Do you think that conducting research on a mainstream topic raises the chances of publishing?	Somewhat agree	42,6%
	Strongly disagree	3,0%
	Somewhat disagree	12,2%
	Strongly agree	42,2%
Do you think that new items play have played a role in the planning of your research agenda?	Somewhat agree	38,9%
	Somewhat disagree	32,2%
	Strongly disagree	12,7%
	Strongly agree	16,2%
Do you think that societal needs play or have played a role in the planning of your research agenda?	Somewhat agree	41,8%
	Strongly disagree	6,4%
	Somewhat disagree	18,0%
	Strongly agree	33,8%

Table 8: Respondents' replies on other determinants of research agenda

4.3.2.6 Analysis of the outcome of the Likert scale

As already said in chapter three, the survey comprised three Likert scale batteries, from the MDAI-R model. We inserted in our survey the MDAI-R model questions that let distinguish between the trailblazing and the cohesive research agenda approach, namely: collaboration, multidisciplinarity *versus* convergence, tolerance to low funding, and fema. As already analysed, a study conducted by Horta and Santos (2020) highlighted how, on the one hand, researchers' cohesive research agenda approach was linked to a commitment to the afferent organisation's values and habits and, on the other hand, researchers who have trailblazing research agenda approach tend not to adapt their research agendas to external pressure.

However, according to the data collected with the survey, the outcome that emerges from our context seem to partially confirm what emerged from Horta and Santos' study. Indeed, on the one hand, survey respondents seem to have a more trailblazing research agenda approach, and yet, as already analysed on the paragraph referring to the VQR influence, the majority of respondents affirm not to have changed their research agenda after the implementation of the VQR.

On the other hand, the majority of respondents affirm to have changed their research agenda after the introduction of ASN, as previously seen in the previous paragraphs, thus contradicting Horta and Santos' statement.

A possible explanation of this outcome may be identified in the different institutional pressure and function of VQR and ASN. Indeed, while, on the one hand, VQR exert an indirect institutional pressure and is mainly focussed on universities' performances, ASN is fundamental for researchers' career advancement, and consequently may have a direct effect in shaping researchers' practices.

4.4 Analysis of the survey free comments

As previously mentioned, the survey comprised five free questions, asking respondents to freely express themselves on the main topics covered, namely, scientific collaboration, calls for research funding, ASN and VQR. From the simple reading of all comments, it emerged a noteworthy interest in the topics covered by the survey, and high level of participation of the respondents. We split the content of the free comments per topic, following the thematic pattern of the survey, namely: collaboration, VQR, ASN, calls for research funding, other factors that may affect research agenda setting.

For the analysis of the first group of comments, we proceeded with an analysis of the free comments content by using some conceptual tags that, in our opinion, could best summarize the concept the respondent wanted to express. The tags were then grouped according to thematic clusters, to provide a thorough overview of respondents' opinions on a wide variety of topics.

According to Tab. 9 (below), many different opinions emerged on the effects of the determinants of research agenda setting, since the respondents were encouraged to express themselves on eventual topics not covered enough in the survey. Below some relevant notes from the analysis of the free comments.

From the collection of the comments on collaboration, calls for research funding, ASN and VQR, both negative and positive opinions on their impact on research agenda setting emerged. More

specifically, the introduction of ASN was widely discussed, for its impact perceived on researchers' practices and on the overall quality of research production (*ASN + quant – qual, ASN influences publication outlet/publication strategy*). Still, ASN is deemed responsible for not allowing young researchers to conduct research properly. For what it concerns the VQR, many negative comments were collected on the confused functioning and procedures, and, still, many respondents criticized the mechanisms of the selection of scientific journals for negatively conditioning their research agenda and their research choices (*RA influenced by publication outlets/scientific journals, RA conditioned by SSD and ANVUR system*). Calls for research funding are generally perceived as muddled and time-consuming, and mainly focussed on mainstream topics, with the risk of progressively set aside topics researchers would like to conduct research on. Scientific collaboration, by contrast, is generally perceived as a great opportunity for both career advancement and quality of research production. The effects of scientific collaboration on researchers' research agenda are perceived as enriching, rather than an hindrance. To conclude, many researchers expressed their concern over the transformation and struggles that academic career is facing: indeed, overwhelming administrative tasks, increasing pressure by their afferent university, scarcity of time and economic resources, seem to play a noteworthy role in negatively shaping the quality of researchers' research production, and the pursuing of their personal research interests.

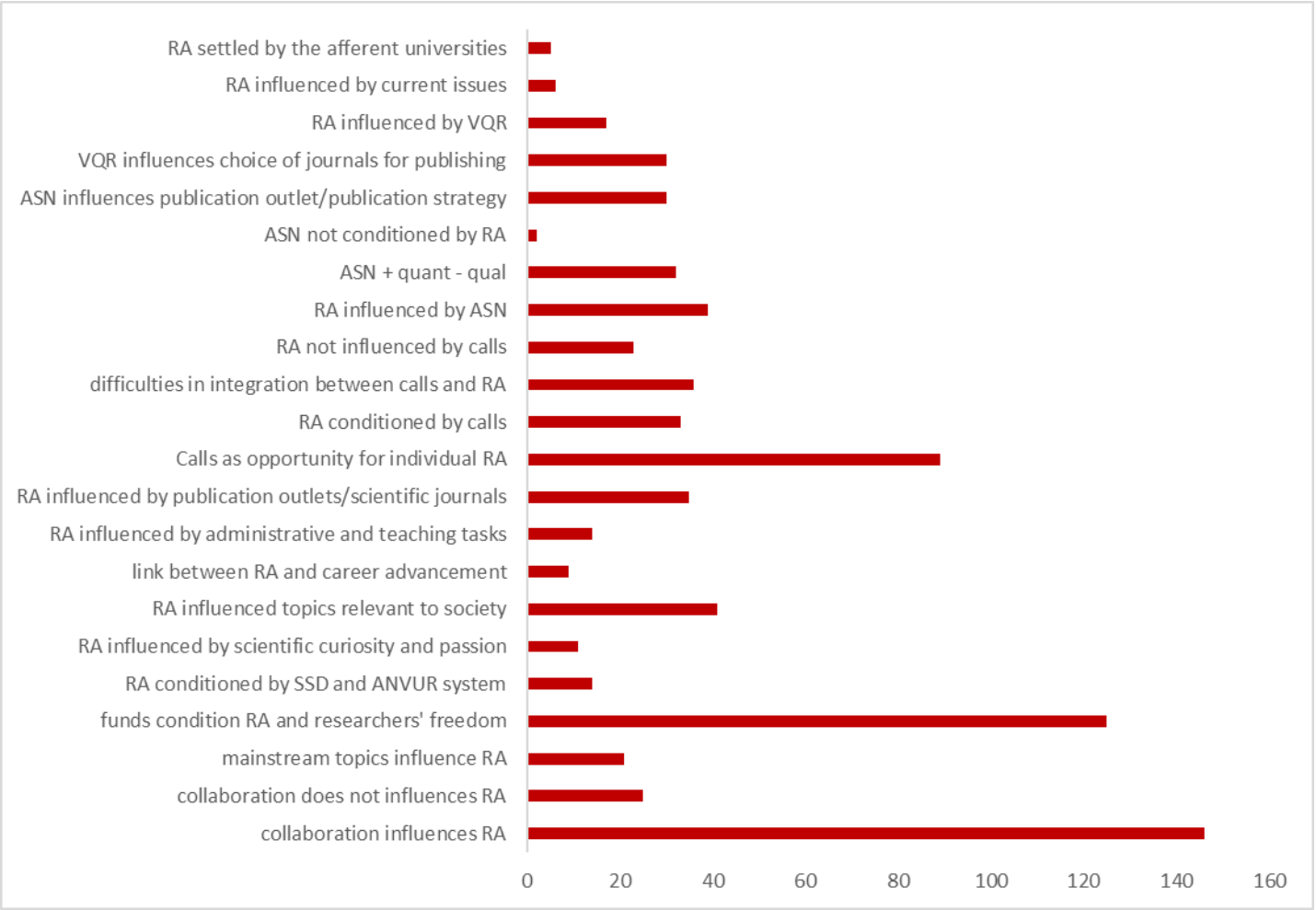


Table 9: Graphical summarize of the frequency of the analysis tags

4.5 Final remarks

The purpose of our research, that is disentangling the components of research agenda setting and understanding the effects of research evaluation, is very challenging for a series of reasons. Among these reasons, we would like to underline, first, that the various components of research agenda, as it emerges from our data, can easily coexist and not being mutually exclusive. Furthermore, the research agenda setting process is inherently due to personal motivations, that, perhaps, the researcher himself is not perfectly aware of. Indeed, a random researcher may have thought, before the Covid 19 pandemic outbreak, that news topics or being beneficial to society were not that important to his own research agenda setting, for, presumably, discovering the opposite after the pandemic outbreak.

However, the data collected let us better understanding whether, and eventually how, the research agenda setting of our target population have changed, with reference to the institutional pressure exerted by VQR and ASN. First, the introduction of VQR has not produced a change in the majority of researchers' research agenda setting, and, thus, our first hypotheses is confirmed. This outcome can be explained by the fact that VRQ exert an indirect l pressure to researchers and to research production, since universities are the actual units of VQR evaluation. However, the percentage of researchers affirming to have changed their research as a consequence of the introduction of VQR is in any case relatively high (nearly 40%). A possible explanation could be that the Italian research evaluation system can be classified as strong (Whitley, 2007), and thus being able to shape and standardize the contents and methods of research communities, since the outcomes of the VQR research evaluation exercises are at the basis of the allocation of a consistent part of the state funds for universities. This aspect could probably explain, as previously said, why from the survey it emerged that senior researchers tend to change their research agenda setting process relative more, in comparison to younger researchers, since they tend to be accountable for their department outcomes. Furthermore, we can also confirm our second hypotheses, namely the stronger effects of the introduction of ASN on research agenda, in comparison to VQR, due to its more direct institutional pressure on researchers' career and hiring process. This is even more understandable considering the increasing difficulties young researchers experience, due to the scarcity of economic resources for the personnel recruitment.

At last, as previously analysed, for what it concerns the influence of VQR on research agenda setting, a different effect between 13 and 14 CUN areas is to be highlighted. Indeed, economists and statisticians tend to affirm relatively more to have changed their research topics, their research motivation and research practices. This outcome, as a matter of fact, is linked to the inner features of the disciplines at issue: indeed, as already highlighted by Bonaccorsi (2015), Whitley (2007), and Lamont (2009), disciplines generally relying on a strong empirical basis tend to adapt more to standardized approach and practices, while, on the other hand, disciplines where the debate on

different theoretical approaches is pivotal, the standardization of practices and approaches is more unusual. We can hence confirm our third hypotheses, namely that disciplines that are empirical evidence oriented are more sensitive to research evaluation.

Chapter 5
Informed interviews of the experts of research evaluation

In this chapter we are going to describe the management and the outcomes of the informed interviews conducted with research evaluation experts. As previously said in the third chapter, we decided to conduct informed interviews (2007) focussed on the most relevant topics emerged from the analysis of the survey, with a twofold purpose: on the one hand, to better explore some relevant topics emerged from the free comments of the survey; on the other hand, to use what emerged from the interviews in order to propose a thorough and integrated comprehension of the effects of the implementation of research evaluation in the Italian academic context, combined with the main outcomes of the analysis of the survey[78], in accordance with the concurrent triangulation mixed method design. We thus opted for interviewing the researchers who participated to the survey, and, among the list of the volunteers, we chose the researchers who had proven experience in the field of research evaluation in the Italian academic context. We considered the proven experience in the field of research evaluation in the Italian academic context as, for instance, the participation of VQR or the ASN evaluation experts' panels, a proven expertise in the field of Italian academic and research evaluation system, the participation to the universities' evaluation centre units. Indeed, interviewing (and interacting with) researchers with proven experience in research evaluation would give us the possibility of having a double viewpoint (Morawska, 2018): on the one hand, the researcher's point of view, who, as an active member of the Italian academic system, has already experienced the evaluation as evaluated; on the other hand, the evaluators (at different levels of the research evaluation system). The choice of researchers with such a professional profile let us classify the interviewees of our research, if we consider the Gläser and Laudel's pattern (2009) to define the possible profiles of experts in a determined social context, as individuals with an expert role in the reference field.

[78] According to a concurrent triangulation mixed method approach, as already furthered in the third chapter.

Hence, for the purpose of our research, conducting informed interviews (where the interviewer should carefully study the interviewee's social field, before interviewing) seems to be the most suitable methodological approach because it enables an in-depth discussion between the interviewer and the interviewee, who would try to thoroughly understand and disentangling the issue of research agenda setting and of research evaluation through the communication process that takes place between them (Laudel, Gläser, 2007), starting from the main topics emerged from the previously conducted survey.

Hence, the present chapter is organized as follows: in the first part, we are going to describe out methodological choices and how we managed to organize the interviews; in the second part, we are going to illustrate the major outcomes of the analysis of the interviews; in the third part, we are going to sum up the interviews results and to draw the conclusions from this research phase.

5.1 Management of the interviews

5.1.1 Interviews protocols main dimensions

We structured the interviews protocols according to the main topics emerged from the survey free comments, which, in our opinion, deserved a better comprehension. The interviews were conceived as semi-structured informed interviews, in which we wanted to encourage a fluent conversation with the interviewees and keeping the interviews protocol only as conceptual reference point. Indeed, as we will see in the paragraph in which we will present the interviews analysis outcomes, the topics emerged from the interviews are more than those listed below.

The main topics we treated during the interviews were (also available in the Annex 2):

1) the VQR and ASN's influence on researchers' research agenda setting,
2) the afferent universities' pressure on researchers for conducting research on determined topics,
3) the role and influence of the classification scientific journals on researchers' research agenda;
4) the eventual VQR and ASN's pressure to conduct multidisciplinary research;
5) the role of VQR and ASN in the internationalization of Italian research production;

6) the eventual VQR and ASN' role in encouraging short term research, instead of long-term research projects;
7) VQR and ASN' role in changing the same concept of conducting research;
8) VQR and ASN' role in encouraging or discouraging academic collaboration;
9) the researcher's opinion on strategic collaboration, and,
10) VQR and ASN's effects on universities' availability of economic resources.

5.1.2 Selection and contact of the possible interviewees

As previously mentioned, we selected the interviewees from the list of survey respondents who expressed their interest in participating to the second phase of the research.

From the list of volunteers, which was separated at the first stage of the survey analysis in order to respect the privacy of respondents, we selected a total list of around 36 researchers.

The selection of the experts was conducted according to the following criteria:

1) the experts were supposed to have proven experience and knowledge in the Italian research evaluation system (former members of evaluation panels, technical support to universities or government councils on the topic of research, ecc..);
2) "representativeness" of the different disciplinary fields, gender, and geographical residence of the interviewees (Laudel, 2006).

We directly reached out the list of the selected researchers by email, to remember to the recipients the purpose of out research, and to invite them to an online interview (Seidman, 2006). In the email, we also provided details of the interviews (the duration, the topics at issue, the online platform to use) and encouraged them to propose a time slot to schedule the interview.

Of the 36 researchers contacted, 10 of them offered availability for the interview. It is noteworthy to remark that the interviewees' attitude during all the process (from the scheduling of the interview, to the running of the interview) was very proactive, and a sincere interest in the topics of our research clearly emerged.

5.1.3 Timetable and organization of the interviews

We conducted the interviews from the beginning of September 2022 to the end of October 2022.

The interviews were conducted through an online platform (Google Meet), for two main reasons: first, to avoid, as much as possible, any complications due to the Covid 19 pandemic, and, second, to easily reach out researchers from any geographical provenance. The link to open the call was sent 5 minutes before the scheduled time.

Once scheduled the interview, researchers were asked to fulfil an informed consent form, to be sent back before the beginning of the interview. Still, before starting we specified to the interviewees that we used a recorder to tape the interview. Interviewees were also informed that, while running the interviews, we would have used a notebook to take notes on interviewee's attitude and on our on perception of how the interview was going on.

The interviews were transcribed manually, since with any electronic device or application to transcribe the audio files was not suitable for adequately transcribing the audio files of the online interviews. The analysis of the interviews content lasted around three weeks (till the end of November, more or less).

5.1.4 Conducting online interviews: pros and cons

As previously said, the interviews were conducted remotely, using an online tool (Google Meet) to both run and record the interviews. As already mentioned, the decision of running the interviews online was mainly linked to need to address and solve eventual problems and unexpected events due to the Covid 19 pandemic. Indeed, this phase of the research was conducted in a moment in which the level of uncertainty made it very difficult to easily plan face to face interviews. Furthermore, the online tool let us easily reach out interviewees throughout the entire country, thus enabling a possible confrontation among researchers from different geographical and economic contexts.

However, the use of online tools presented some side effects, as well. First, in many cases the conduction of interviews was very difficult because of connection issues, that sometimes made it very hard or unpleasant to have a comprehensive conversation with the interviewees. It is worth underlining, however, that the interviewees had a very positive attitude, and this was extremely helpful in some situations. Second, it was not possible to use an online tool to automatically transcribe the recording of the interviews because the sound was not clear at all. This aspect, as a matter of fact, represented a rather time-consuming issue, since the manual transcription of the interviews too long time.

5.2 Interviews analysis

As previously mentioned, in this chapter we are going to outline the debate that emerged from the informed interviews. The analysis of the interviews has proceeded as follows:

1) we have in the first place isolated the interviews excerpts concerning the VQR and the ASN, attempting to identify some theoretical patterns;
2) from both the interview questions and the interviewees' speeches, it was possible to identify, for both the VQR and the ASN related excerpts the following key concepts: *VQR and ASN influence on research agenda*, *VQR and ASN effects on multidisciplinary research*, *VQR and ASN effects on internationalization of research*, *VQR and ASN long-term effects on research*, *VQR and ASN effects on academic collaboration*, *VQR and ASN effects on researchers' conception of research*, *VQR and ASN effects on young researchers' career*, *effects of VQR and ASN on the management and availability of universities' funds*, *final considerations on VQR and ASN*;
3) the remaining parts of the interviews were then organized according to some presumably relevant topics that emerged frequently across the various interviews; the main relevant topics identified are the following: *universities' pressure on researchers' research agenda setting and on other academic activities related aspects*, *role played by scientific journals and*

publication outlets on researchers' research activities, issues linked to the recruitment process of young researchers, broad considerations on the selection criteria of GEV members and referees;

4) we have created a table trying to summarize the interviewees' opinion and feeling on the topics discussed according to three main attitudes emerged: a positive attitude (named the *integrated ones*), a negative attitude (named the *apocalyptic ones*), and a rather neutral attitude (the *neutral ones*).

We are then going to discuss the main issues emerged as follows: the first paragraph is mainly focussed on VQR-related considerations; in the second part we are going to illustrate what emerged on ASN, while in third (and last) part we will describe interviewees' opinions on the various that may influence research agenda setting and report some interviewees' reflections and concerns on the academia status in a broad sense. The purpose was to provide a thorough overview of interviewees' attitudes towards the research evaluation and academic as a whole.

5.2.1 The effects of VQR

The first question was focussed to generally understand interviewees' opinion on the eventual effects on the introduction of the VQR on researchers' research agenda setting. It is noteworthy to remark that all the interviewees maintained that research agenda setting, and research practices have been somehow influenced by the introduction of the VQR and, subsequently, of the evaluation culture it entails. The interviewees, indeed, all generally agree that the introduction of the VQR has represented a remarkable element in the academic environment, and that, as in any evaluation system, the evaluator induces a change in the evaluated system's pattern of behaviour. The pressure to publish as much as possible, accordingly, would have induced researchers to adopt some strategies to comply with the VQR requirements. As a matter of fact, some interviewees maintain that, for instance, researchers tend more to conduct research on mainstream topics, which would ensure the researcher to publish more easily and get more visibility from the reference peers community with positive spill

over during the evaluation process, and to choose empirical or non-risky research, rather than focussing on theoretical or trailblazing research. This, according to many interviewees, would happen at the expense to the research quality and the researchers' creativity, since researchers do not have the possibility to properly devote time to some research products that would not ensure immediately consensus during the evaluation process:

"…c'è una fortissima tensione ad occuparsi di quello che viene chiamato mainstream. Sappiamo che si tratta di argomenti socialmente rilevanti e significativi, ma questo significa anche che ci sia una diffusione di più prodotti che affrontano la stessa tematica che viene affrontata in maniera ripetitiva, dovuto anche al bisogno di pubblicare. Quindi, spesso ci si trova a vedere pubblicazioni ravvicinate su pochi temi mainstream nel tempo con un incremento della conoscenza molto contenuto. L'idea che si ha è che un po' tutti dicano le stesse cose. Vengono così meno l'originalità della ricerca scientifica, visto che si fa ricerca su un tema su cui tantissimi hanno detto già tantissimo; inoltre, si deve riflettere che, come sociologi, si devono tenere in considerazione i temi che emergono, ma, con la rincorsa alla pubblicazione, molti di questi articoli mancano di una dimensione del pensiero tipico della sociologia, facendo prevalere una dimensione descrittiva, più che analitica. Come se si facesse solo della micro sociologia, senza però una dimensione di tipo esplicativo." (Interviewee 14_PO_F)[7980].

Some interviewees, furthermore, outlined how the introduction of VQR has pushed researchers to change their publishing patterns, thus preferring scientific papers rather than monographs and curatorship, which are far more time consuming. Still, the prestige of publishing a paper in international top journals, which might have an 80%-90% rejection rate, is mainly linked to the inherent difficulties and high competitiveness of being selected among a tremendous number of submissions. The issue of the role played by international scientific journals in shaping research

[79] In this paragraph, we have inserted some extracts from the interviews. We adopted a coding system in order to identify some features of the interviewees, while trying to ensure the anonymity. Hence, according to our coding system, we identify with the number 13 (13a stand for Economics and Economic History, 13b stands for Business Economics and Statistics) or 14 the interviewees' disciplinary field, PO stands for full professor and PA stands for associate professor, while M and F stand respectively for male and female.

[80] "There is a strong pressure to deal with what we call mainstream. We know that they are socially relevant and important topics, but this also means that there is a relevant dissemination of products that deal with the same topic that is treated in a repetitive way, due to the need to publish. Hence, we often see many products on few mainstream topics in a short time span, that determine only a little knowledge advance. The idea is that everybody just say the same things. The originality of scientific research decrease, since people conduct research on topics which many people have already studied. Moreover, we have to reflect, as sociologists, that we have to consider even the emerging topics, which lack, on the contrary, the typical sociological analytical pattern, letting prevail a more descriptive dimension, rather than analytic. As if we would do micro-sociology, without an explanatory dimension".

agenda setting has triggered a lively reaction among the interviewees, and it will be further addressed in the last paragraph.

Moreover, from the content of the analysis of the interviews it emerged that most of the interviewees agree on the stimulus that the implementation of the VQR determined on the internationalization of researchers' production (except for two interviewees, who affirmed not to have identified remarkable effects of VQR on the internationalization of researchers' productivity). Indeed, the VQR would have introduced a strong drive to publish in international scientific journals, with many positive effects on researchers' productivity, such as, a push toward scientific collaboration with foreigner colleagues, a standardization of the research practices and an increase of spreading scientific results because of the pervasive use of the English language. An interviewer even claimed that, in his opinion, the Italian scientific production, especially in some disciplinary field, has become less provincial:

> *"L'unico effetto assolutamente positivo* [della VQR] *è stata una maggiore spinta verso l'internazionalizzazione della nostra comunità accademica. Questo è un effetto congiunto di queste nuove forme di valutazione che sono state introdotte. Su questo siamo sicuri che è interessante l'impatto della vqr, che spinge anche le riviste italiane a internazionalizzarsi…. Comunque, la sprovincializzazione è un effetto positivo sotto tanti punti di vista."*[81] (Interviewee 14_PO_F)

Nevertheless, some interviewees maintain that the flipside of the current overemphasis on publishing in international scientific journals would be a loss of some specificities of the Italian academic tradition and culture, with a consequential underestimation of the Italian research production. Moreover, these interviewees affirm that the internationalization sometimes is more considered as an objective, rather than just a feature of a researcher's production, thus making coincide the quality of a scientific work with its international characteristics. Accordingly, publishing in an international scientific journal would be considered as a guarantee of an excellence certification, even though some journals' evaluation criteria are debatable.

[81] "The only definitely positive effect [of the VQR] has been a major drive toward the internationalization our academic community. This is a joint effect of these new forms of evaluation that have been introduced. On this we are sure that the impact of the VQR has been interesting, and that also pushes the Italian journals to get more international. By the way, the de-provincialism is a positive effect according to many points of view."

For what it concerns the eventual effects of VQR on multidisciplinary research, the interviewees' opinions were quite different. We can thus identify two main points of view. On the one hand, some interviewees affirm that the introduction of the VQR has encouraged multidisciplinary research, especially in determined fields, such as the niche topics, to get more visibility and citations. Still, the VQR call for notice specifies that researchers that work with interdisciplinary research groups may ask to submit their scientific works to the GEV panels they prefer.

On the other hand, other interviewees maintain that they have not remarked a significant effect of the introduction of VQR in encouraging multidisciplinary research. More specifically, they affirm that the positive or negative evaluation of multidisciplinary scientific works heavily rely on the GEV evaluators' opinion, and that it is not always a well-accepted practice. It also emerged that the daily tasks a professor is called to perform, and, among which, also the evaluation-relation duties, make very difficult to coordinate and work with multidisciplinary research teams.

From the question asking the interviewees their opinion on the effects of the introduction of VQR on scientific collaboration, interestingly, contrasting opinions emerged. It is important to remark, however, that most of the interviewees (6 out of 10) maintain that the introduction of VQR has strongly incentivized the scientific collaboration among researchers. Indeed, collaboration would be an important opportunity for researchers to publish more, get more visibility within the scientific community, and, accordingly, having an easier access to top international journals, in which a multidisciplinary and comparative approach may be appreciated. According to an interviewee, scientific collaboration, in social sciences particularly, would have much increased after the introduction of the VQR:

> *"… a otto mani ognuno mette qualcosa per una pubblicazione e ci si mette meno tempo, e questo si osserva secondo me andando a valutare l'incremento del numero degli autori nella produzione scientifica in ambiti che non sono quelli in cui tradizionalmente il numero degli autori non è elevato. Andiamo a guardare gli articoli nelle scienze sociali, e guardiamo quanti sono gli autori degli articoli! Negli ultimi anni il numero medio degli autori è cresciuto, passando da 1,2,3*

verso 3,4,5. Quello secondo me è il fatto, perché c'è l'incentivo a produrre di più; quindi, cooperando si riducono i tempi e tutti quanti hanno un prodotto della ricerca."[82] (Interviewee 13a_PO_M).

By contrast, some interviewees highlight how the introduction of VQR has genuinely encouraged the scientific collaboration among researchers afferent to the same university. Indeed, because of the high competitiveness that VQR introduced among different universities, collaboration with colleagues from other universities are inevitably discouraged, with higher negative effects on researchers afferent to medium and small universities. The competitiveness among universities and the overemphasis on productivity is, according to an interviewee, the symptom of the shifting of the conception of university from public place of culture, toward an enterprise.

Interestingly, the evolution of the conception of scientific research and universities have proven to be issues of much sensitivity for many of the interviewees. There is, indeed, a widespread feeling of decadence of the conception of research, because of research evaluation. The need to publish and to get good rating by the evaluators seems to be inducing to shift researchers' focus from the study and research activities as a creative and spontaneous "discovery" process, to a production-oriented process, where "more is more".

Accordingly, this would be encouraging a series of strategic behaviours among the researchers, such as:

- rejecting research projects and research topics that would not ensure a positive feedback from the peers' community and an easy publication process, with the risk of publishing less theoretically solid works;
- privileging only certain scientific journals, and subsequently the methodological thematical and approach they entail;

[82] "… a eight hands [publication] is faster, and this is also observable by evaluating the increased number of the authors in scientific production in disciplinary fields that are not those in which the number of authors is traditionally high. Let's have a look at the papers in social sciences, and check the number of authors per article! In the last years, the average number of authors is raised, from 1,2,3 to 3,4,5. This is, in my opinion, the issue, because there is an incentive to produce more; hence, cooperating the timing is reduced and everyone has his own research products".

- segmenting long-term research in more publications, with negative consequences on whole results and analysis of the research itself;
- opting for publishing results on papers rather than books or book chapters.

All these considerations are also to be considered in the frame of all the tasks professors have to manage, making more complex to conduct good quality research:

"*Se guardiamo le grandi tendenze, questa impostazione non mette l'accento sul processo di scoperta scientifica, ma sulla pubblicazione di per sé. Alcune pubblicazioni saranno utili, altre meno. Ci sono valutazioni anche sulla didattica, della terza missione e dei dottorati, e quindi le esigenze da soddisfare si moltiplicano. Tutto però è orientato a ottenere più fondi per fare anche più ricerca, a volte anche a scapito di altre attività. Chi entra nella logica e segue fino in fondo i vari esiti delle altre valutazioni, chi contribuisce viene premiato; invece, chi non contribuisce finirà per non fare niente. …. È un meccanismo che si alimenta da sé, e questo può influenzare anche il tipo di ricerca. …il problema è che siamo sempre meno manager di noi stessi, e il nostro manager è il manager di ateneo."*[83] (Interviewee 14_PO_M).

Nevertheless, few interviewees consider the pressure to publish as an opportunity to open up the traditional Italian research set-up to a more international and global approach, and that it is even possible to positively join the need to publish with long-term research projects.

While conducting the interviews, we warmly invited the interviewees to advance some considerations on some eventual aspects of the VQR they wanted to point out. Some important critics emerged.

First, the more critical aspect is that the VQR call for notice (explaining the evaluation criteria and the functioning of the entire process) is published after the period under evaluation is over. Still, even though an interviewee explained that the rationale of publishing the criteria *a posteriori* was to avoid researchers' strategic behaviours, and, still, that the evaluation criteria are after all easily predictable by researchers, many other interviewees claimed that this aspect was considered extremely negatively:

"La maggiore criticità è che le regole della vqr vengono esplicitate a posteriori. Il ricercatore è portato a fare una scommessa sulle future valutazioni rispetto al proprio lavoro. Questo è l'aspetto

[83] If you look at the great trends, this approach does not focus on the process of scientific discovery, but on the publishing *per-se*. There is also evaluation on teaching, on third mission, on doctorates, and then the needs to satisfy multiply. Everything though is oriented to get more funds for doing research, sometimes at the expense of other activities. Who joins the logic and follows the several evaluation results, is awarded; otherwise, who does not contribute will end up by not working any more. It is a self-nourishing mechanism, and this may influence the kind of research. The problem is that we are less and less managers of ourselves, and our university is our manager."

più critico perché i criteri sono assolutamente casuali. Nelle ultime VQR, ho pubblicato 50 articoli e ne ho dovuti scegliere 4, e li ho scelti del tutto casualmente"[84] (Interviewee 13a_PO_M).

Still, a certain distrust toward the objective and the VQR evaluation criteria, particularly considering that the goals of VQR and ASN seem to be in contradiction. The VQR evaluation criteria, more specifically, are perceived as the result of a top-down mechanism that has excluded the academic community from the debate. This element, accordingly, seems to be delaying the acceptance of the evaluation culture in the Italian academic community, where the evaluation rating on their scientific work is often badly accepted.

Moreover, according to many interviewees, conducting research, teaching, and managing the several administrative tasks is very hard, and this seems to be one of the main reasons why it is very difficult to find volunteers for the GEV panels and reviewers for carrying on the evaluation activities (and, for this reason, sometimes the reviewers are not that competent). This work overload seems to make the evaluation activities as GEV evaluator more difficult:

"Io ho fatto da valutatore per la vqr, per quel poco, ma siamo tutti oberati di incarichi amministrativi, che sono tanti. La burocrazia è tanta, la produzione di carta è tanta, e quindi questo riduce il tempo a disposizione per la ricerca. Poi sta alle persone decidere quanto dedicare ad un'attività di referaggio. Il tempo c'è ma c'è anche una questione di competenze, cioè di capire in tempi rapidi la qualità di un contenuto scientifico. Bisogna trovare un compromesso"[85] (Interviewee 13a_PO_M).

However, while some interviewees highlighted that the VQR mechanism should also consider the third mission and teaching activities, others also stressed that the introduction of VQR has helped standardizing the papers format, making it easier a confrontation with international researchers

5.2.2 The effects of ASN

[84] "The major criticality is that the rules of the VQR are published *a posteriori*. The researcher is called to make a bet on the future evaluations on his work. This is the most problematic aspect, because the criteria are totally random. In the last VQRs, I published 50 articles and I had to pick 4, and I chose them totally randomly".

[85] I have been an evaluator for the VQR, for a while, but we are overwhelmed by tasks, which are many. Bureaucracy is very much, the paper production is very much, and this reduces the available time for research. Then it is up to evaluators to decide how much to dedicate to review. There is time, but there is also a matter of competences, namely to quickly understand the quality of a scientific work. A compromise must be found".

The introduction of ASN in the selection process of professors has been producing remarkable effects on researchers' research agenda setting, according to all the interviewees. Indeed, the direct link between ASN result and the possibility of getting the tenure is a huge incentive for researchers to adapt their own research agenda and research practices to the ASN requirements:

"Io non so che dirle, nel senso che l'ASN certamente, ed è evidente a tutti, ha provocato questa corsa a pubblicare; quindi, è evidente che se cominci a dire che tante pubblicazioni servono come requisito di accesso per partecipare all'asn, è legittimo ipotizzare che ci sia o una maggiore superficialità nel trattare il medesimo tema o la tendenza a trattare tanti temi. Quindi sicuramente l'asn ha prodotto degli effetti strutturali e questi si vedranno tra 5/10 anni quando tutti i giovani assunti cominceranno ad avere dei problemi perché stiamo crescendo dei giovani che pensano che fare il professore sia scrivere dei paper, per cui finiscono in un'istituzione e non sanno gestirla...Sul modo di fare ricerca e di concepire il mestiere, vedo degli effetti strutturali della asn."[86] (Interviewee 14_PO_M)

As for the effects of VQR on the internationalization of research, from the interviews analysis it does not emerge a consensus on this specific matter. Indeed, while, on the one hand, some interviewees affirm that the introduction of ASN has remarkably encouraged the internationalization of research, on the other hand, others maintain that the internationalization is inherently linked to the creation of collaboration and publication channels with foreigner universities, and that researchers from universities which do not guarantee such international connections may be disadvantaged.

The functioning of the ASN, furthermore, seems to both encourage and discourage multidisciplinary, according to the interviewees. Multidisciplinary research, indeed, seems to be highly encouraged by the possibility of getting more citations, especially in the case of niche disciplinary fields or topics (as seen also in the previous paragraph), thus increasing the possibility of having a good rating by the ASN evaluation committees. By contrast, the ASN mechanism linked to the disciplinary sectors, for which researchers who wants to apply for a disciplinary sector have to submit scientific works which

[86] "I don't know what to say, I mean, it is evident to all, the ASN has determined this rush to publish; thus, it's evident that if you start saying that you need many publications as an access requirement to participate to the ASN, it is rightful to assume that there is a major superficiality in dealing with the same topic, or the trend to deal with many topics at the same time. The ASN has certainly determined structural effects and we will see them in 5/10 years, when the newly-hired young researchers will start having problems, because we are growing up young researchers who think that being a professor is writing papers, then they end up in an institution and don't know how to manage it. On the wat of conducting and conceiving research, I see some structural effects due to the ASN."

pertain to that specific disciplinary sector, discourage researchers from conducting research with multidisciplinary teams on topics that would be useless for the achievement of the habilitation.

By submitting the issue of scientific collaboration, with reference to the introduction of ASN to the interviewees, two different points of view emerged. On the one hand, some researchers affirmed that scientific collaboration is increasing due to the introduction of ASN, since young researchers have to publish as many research products as possible, to get the habilitation. Interestingly, most of the interviewees deem that the introduction of ASN has incentivized the scientific collaboration.

They are also more frequently involved in research projects by their supervisor, who can eventually sponsor them for future job positions at university. On the contrary, some interviewers maintain that the ASN would discourage scientific collaboration because its purpose is to evaluate the scientific capacity of the individual research (as opposed to the VQR, whose objective is to evaluate the universities and departments):

> *"Per l'asn è diverso, perché li si valuta l'apporto individuale del candidato. Quindi se uno presenta sempre e solo lavori di collaborazione, diventa difficile valutare il candidato dal punto di vista scientifico. Si tratta di capire qual è il tuo contributo individuale come studioso.*"[87] (Interviewee 14_PO_F)

The functioning of the ASN, moreover, seems to have remarkably changed young researchers' approach and conception of research. Interestingly, there is a consensus among all inteviewees on the possible negative consequences of the ASN on young researchers. Accordingly, the focus of young researchers seems to totally oriented toward the publication process, rather than a thorough comprehension and analysis of the topics at issue. The interviewees are actually concerned over an increasing superficiality of young researchers' theoretical background, since they tend to publish many scientific works with a repetitive methodological approach, with very little consideration of the

[87] For ASN it's different, because it evaluates the individuale candidate's contribution. Thus, if a researcher only submits scientific works in collaboration, it becomes difficult to evaluate the candidate according to a scientific point of view. It is all about understanding your contribution as a scholar."

theoretical basis of the study. A major risk is that the rush to publish and the evaluation culture may jeopardise on the long-term the quality and originality of the whole academic community:

"Sicuramente l'asn fa correre troppo i giovani, non li fa meditare, e noi non ci rendiamo conto, soprattutto il sistema non si rende conto, che se tu devi diventare associate professor a Princeton, ti chiedono un articolo pubblicato e due sottomessi a importanti riviste. Da noi per diventare associati devi avere 10 titoli. Il vero punto è l'asn che crea questa stortura mostruosa per cui tu vai a cercare pubblicazioni. Ovviamente ce n'è uno su dieci che riesce a tenere velocità, quantità e qualità, ma quelli c'erano anche prima. Mentre molti che potrebbero produrre cose sensate e ponderate e solide, tendono a disperdersi in questa pressione per pubblicare e pubblicare. Il vero problema per me è dato dall'asn e dall'impatto sui giovani. C'è da preoccuparsi perché se tu parti così pensi che il mestiere sia quello lì. …. Che non venga valutato l'insegnamento nell'asn, è una scelta di legge. Io sono sempre stato contrario all'abilitazione, …. L'asn valuta la tua capacità di fare ricerca e questo è interessantissimo, ma non prevede la prova didattica. Quindi è veramente una stortura, secondo me l'abilitazione è uno strumento perverso che sta creando l'idea che poiché sono abilitato, ho diritto ad essere professore. invece la legge è chiara sul fatto che è solo un requisito minimale."[88] (Interviewee 14_PO_M)

Moreover, the fact that with the introduction of ASN, as already seen in the second chapter, many researchers with habilitation have not then been hired by any university has become a rather political issue that has been long debated. Many rectors, for instance, during the electoral campaign, often promise to hire more associate professors, since it is a much sensitive question for the academic community.

To conclude, it is also noteworthy that researchers that are attempting to get the habilitation do sometimes have to shape their research activities according to the objectives they are ultimately willing to pursue. Indeed, the scientific journals deemed of quality for the ASN are not the same for

[88] "The ASN certainly makes the young researchers run too much, not meditating, and we are not realizing, especially the system is not realizing, that if you must become an associate professor at Princeton, they ask you for one published article and two submitted to important journals. [In Italy] you have to have 10 articles to become associate professors. The real point is that the ASN creates this monstrous distortion for which you are looking for publications. Obviously, only one researcher is able to keep the speed, quantity and quality, but these were there even before. Otherwise, many others that may produce solid, thoughtful and solid things, tend to get lost because of this pressure to publish. The real problem in my opinion is the ASN and its impact on the young researchers. We need to worry because the young researchers think that the academic profession is that one. That the ASN does not evaluate the teaching, it is the law. I have always been against the habilitation, … The ASN evaluates your researching skills and that is very interesting, but it does not entail teaching tests. Hence, it is a real distortion, I think that the habilitation is a perverse mechanism that is suggesting the idea that, since I am habilitated, I have the right to be a professor. Instead, the law is clear in the fact that it is only a minimal prerequisite."

the VQR, and thus the researcher may have to fragment and to re-shape his own research agenda because of this specific feature of VQR and ASN.

5.2.3 Further considerations on the evaluation system and the academic environment

Some more interesting issues emerged from the interviews analysis, specifically about the role of the classification of scientific journals, the eventual role played by universities in shaping researchers' research agenda setting, and whether the introduction of VQR has changed universities' management of research funds.

First, all the interviewees agree on the remarkable role played by the classification of scientific journals in shaping their research agenda and research practices, for a series of reasons. First, because the classification has not always been permanent over time, basing the entire procedure on principles that are perceived by the academic community as arbitrary and not scientifically funded, considering also that the evaluation criteria of the journals are often debatable. The issue is indeed much debated among scholars, to the point that the academic organizations have been able to make ANVUR including a vast number of scientific journals in the list. However, the difficulties in this system are still perceived as problematic, especially in some research contexts the researchers are embedded in:

"*È evidente [l'influenza delle riviste sull'agenda di ricerca], perché adesso la pressione è di pubblicare su riviste di fascia a o grandi riviste internazionali. Questo in qualche modo anche se non è stato detto crea veramente una discriminazione tra chi ha accesso a certe riviste e chi no, o a certe case editrici e chi no. Comincia con il contare molto il gruppo di ricerca in cui sei inserito. Gli accessi, infatti non sono completamente liberi e meritocratici, ma legati a cordate e a rapporti con i gruppi editoriali; quindi, in qualche modo questo crea molte differenze anche all'interno dell'ambiente accademico (a seconda delle università in cui lavori, o dei gruppi di ricerca in cui sei collegato).*"[89] (Interviewee 14_PO_F)

[89] It is apparent [on the influence of journals on research agenda], because the pressure now is to publish on *riviste di fascia a* or great international journals. This somehow determines a discrimination among who has access to certain journals and who has not. It starts to be very important the research group you are in. The accesses indeed are non-completely free and meritocratic, but connected to editorial groups, and this is somehow creating many differences within the academic environment (depending on the university you are working in, or of the research groups in which you are in)".

The chaotic way in which ANVUR handled the classification of the scientific journals and its effects on researchers' planning of their research activities was perceived as really problematic, since the evaluation criteria of scientific works were, as previously said, published *a posteriori*.

It is wort citing, however, that an interviewer positively considers the introduction of the lists of scientific journals as a way to counterbalance the pre-eminence and power that certain *baroni* previously had in addressing researchers' publications only on determined outlets that were under their supervision.

Secondly, interviewees all generally agree that their afferent universities do not directly attempt to shape their research agenda. They are generally only invited to take charge of certain new research projects in some occasions. Nevertheless, the European calls for research funding, the PNRR, and the third mission, can sometimes address researchers' research agenda in order to get more research funds or to devote their own research agenda and research activities to tackle societal issues:

"*Al limite può succedere che arriva al dipartimento una richiesta di fare una determinata ricerca, e il direttore del dipartimento può chiedere chi è disponibile a fare ricerca in quel determinato tema. Quindi non si parla di pressione. La terza missione però sta diventando una funzione molto importante nell'università, che si sostanzia nel rapporto con il territorio e i suoi problemi. Questo quindi stimola e spinge l'agenda dei ricercatori verso alcune tematiche ed esigenze della società."*[90] (Interviewee 14_PO_F)

Interestingly, it is worth citing a peculiar case on this specific matter. Indeed, an interviewee maintains that his afferent university exerts a much proactive pressure to address researchers' research agendas, since in the territory in which that specific university is there are many private enterprises and local institutions directly requiring the university a local-oriented and problem-solving research production.

Lastly, some interviewees maintained that the managing of research funding as considerably changed after the introduction of VQR. As a matter of fact, the allocation of resources to universities according

[90] It may happen, at the latest, that a request to carry on certain research is made to the department, and the department director can ask who is available for conducting research on that specific topic. Thus, we cannot speak about pressure. The third mission is though becoming a very important university function, that substantiates itself into the relation with the territory and its issues. This hence stimulates and pushes researchers' agenda towards some societal issues."

to an awarding mechanism has determined a strong competitiveness among universities, as previously said, and universities and departments have managed more efficiently the available resources. Interestingly, some interviewees affirmed that the freedom in managing the economic resources at department level should be extended even to the recruitment process, to make universities free and able to hire and pay professors' salaries, in order to make public universities as competitive as private universities.

5.3 A summary of the key elements from the interviews and some final remarks

In this chapter, we have proposed an analytical summary of the interviews conducted with the experts of the Italian research evaluation system, collecting and clustering the interviewees excerpts on a thematical basis. With Table 10 (below) we have tried to graphically sum up the major outcomes of the interviews, in order to propose a few considerations, also in light of the survey major results illustrated in the previous chapter. This table has been structured considering the content tables used to illustrate and summarize the interviews analysis in Reale and Seeber's work (2011). Still, in Tab. 10 we also added the absolute frequency of the comments collected on the main specific topics reported.

	Apocalyptics	Neutrals	Integrateds
VQR: influence on RA	negative effects on research quality; less originality; debatable ev.criteria (6)	no effects on RA (3)	Italian research improves; less provincial (1)
VQR: influence on internazionalization	flipside of pressure to internationalization (4)	no effects on internationalization (1)	positivedrive to international collaboration (5)
VQR: influence on multidisciplinarity	SSD and GEVs discouraging multidisciplinary (4)	no effects on multidisciplinary (4)	increased multidisciplinary (3)
VQR: influence on collaboration	collaboration discouraged, high competitiveness (3)	no VQR effects on collaboration (1)	collaboration positively incentivized (6)
VQR: influence on conception of research	*publish-or-perish* , fragmented and evaluation oriented research (8)	no VQR effects on conception of research (1)	positive effets againts *baroni* (1)
VQR: criticalities	*a posteriori* rules, debatable criteria, difficulties as evaluators (8)		standardized Italian papers, VQR improving (2)
ASN: influence on RA	strong influence on RA (9)		
ASN: influence on internazionalization	classification of journals makes internationalization difficult (4)		internazionalization encouraged (4)
ASN: influence on multidisciplinarity	ASN features hinder multidisciplinarity (3)		multidisciplinarity encouraged (2)
ASN: influence on collaboration	ASN features hinder collaboration (4)		seniors help juniors for the ASN (6)
ASN: influence on conception of research	segmented, career-oriented conception of research (7)		
Effects of classification of scientific journals	mainstream topics, less originality, power of editors (7)		fake problem, counterblance to *baroni* (2)
Universities' role in shaping RA	no direct influence, only slightly (6)		
Funds management	improved funds management (5)		

Table 107: Graphical summary of interviews content analysis

The first consideration that is worth saying is that we perceived that the academic community is considerably concerned of the consequences and effects of the relative recent changes implemented in the academic system.

This was apparent from the active participation to the survey (just to remind, a 36% response rate), and to the interviews as well. During the interviews, indeed, the interviewees were all very helpful and talkative, even when the connections or electronic devices issues made the interviews more complicated. Their lively reactions to some VQR and ASN-related questions proved that these issues are quite sensitive. The importance covered by these topics is also apparent by how the interviews

excerpts are distributed in Tab. 10: for some topics (VQR and ASN's influence on RA, ASN and VQR influence on conception of research) interviewees' opinions were extremely polarized, and often the majority of the comments are more concentrated around the negative pole of the (rough) continuum between apocalyptic and integrated. More specifically, we adopted the apocalyptic/integrated dialectics, borrowed from Umberto Eco's essay (1994), to identify two major researchers' attitudes toward the implementation of VQR and ASN: in our sense, the apocalyptic ones are the researchers that are remarkably sceptical on VQR and ASN, both rationale and consequences, while the integrated ones are those who have a more positive attitude. However, to Tab. 1 we have added the "neutrals" modality to identify a more nuanced approach to the issues at stake.

As it emerges from Tab. 1, we find a very polarized attitude with reference to ASN-related topics, on VQR criticalities, on the effects of the classification of scientific journals, on universities' role in shaping research agenda, and, lastly, on the universities' management of funds. The VQR-related comments, by contrast, are more diversified.

The outcomes of the interviewees are in line with the survey outcomes, namely, that the ASN is generally perceived by academics as more affecting, in comparison to VQR. Indeed, as it emerged from the interviews, the ASN is perceived as a clear-cut cleavage in the academic environment, especially for what it concerns young researchers' training and the long-term research production.

On the other hand, the introduction of the VQR is perceived as less affecting the researchers' daily activities, and it is generally perceived as really relevant only in the period of submission of scientific works and the publications of evaluation results. However, a rather common feeling that emerges, even from the survey free comments, is that the VQR evaluation criteria and the functioning should have been more shared with the entire academic community, and would need a general regular framework, in order to better plan their research activities. Still, what emerges from interviewees is that the academic community is still adjusting itself, after the introduction of VQR and ASN, and that many researchers are still wondering what are the long-term effects of them on the series of aspects we dealt with.

A second consideration, which is linked to the previous sentence, is that, broadly referring to our research questions, the introduction of ASN and VQR has determined a change in the "epistemological curve". According to both the survey and the interviews, we can affirm that ASN and VQR are contributing to a re-shaping of the concept of research and its scope, and, consequently, on researchers' research agenda, since we considered it as the planning of the actions to concretely do, in order to conduct meaningful research. This is actually the core comment of the apocalyptic researchers: they strongly criticize how, with the introduction of the shareable and acceptable principle of research evaluation, its medium and long-term effects would determine a shift from the passion for conducting research and the process of discovery toward a more enterprise-like approach, with all its consequences on the quality of research.

Discussions and conclusions

For the scope of our research, that was to understand whether the introduction of VQR and ASN has somehow influenced the way in which researchers' research agenda setting and to provide a thorough overview of Italian researchers' general opinions the research evaluation system. For this reason, we opted for a concurrent triangulation mixed method approach (a webmail survey and the management of informed interviews with survey respondents), in order to match the strengths of both the qualitative and quantitative approach, and counterbalancing where possible, their weaknesses.

We conducted a self-administered webmail survey, with whose results we are able to reply to our three research questions. The first research question was focussed on understanding whether the introduction of VQR has been influencing Italian researchers research agenda setting process. As a matter of fact, most of Italian researchers affirmed that they have not changed their research agenda due to the introduction of VQR, thus, confirmed our hypothesis. Moreover, with regards to our second research question, on whether the effects of ASN might be stronger on researchers' research agenda setting, we can maintain that ASN has had stronger effects, since most of the survey respondents affirmed that they changed their research agenda because of research agenda setting. The different effects are inherently linked to the VQR and ASN's different functioning and objectives. Indeed, on the one hand, the VQR exerts an indirect institutional pressure to researchers' productivity, since its main purpose is to evaluate universities and departments' performance, and the consequences of universities' rating do not directly affect them (but may affect the allocation of universities' fundings). On the other hand, the ASN is the *condicio sine qua non* for a researcher to get the tenure, and thus has directly consequences on researchers' productivity and career (especially for what it concerns the junior researchers, as the survey outcomes highlight). Indeed, as it emerged from both the survey free comments and the interviews, the introduction of the ASN has been perceived as a breaking point in the academic environment, which, as they admit, has entailed a series of adaptation strategies, by both junior and senior researchers.

However, the share of people affirming to have changed their research agenda as the result of the introduction of the VQR is not that low (around 39%), and this fact may be explained by the same features of the VQR: indeed, according to Whitley (2007), highly standardized and conducted on a regular basis national research evaluation systems tend to influence and harmonize more researchers' research production and research practices, especially when the research evaluation outcomes may bring some consequences or rewards for universities and research institutes.

Our third question was also confirmed, since we identified different research agenda setting behaviours between 13 and 14 CUN sectors. Indeed, economists and statisticians have affirmed to have changed their research topics, their research motivation and research practices relatively more, in comparison to sociologists and political analysts. As already illustrated in the third chapter, some authors (Bonaccorsi, 2015, Lamonte, 2009, Whitley, 2007) had already identified a different behavioural pattern among the scholars from disciplines which have a strong empirical basis (such as, economists and statisticians), affirming that they tend to have a more adapting approach towards the guidelines and objectives proposed by the academic elite of the afferent scientific community. Moreover, even the hybrid evaluation methodology adopted by the GEV of the 13 CUN area may have induced a more cohesive research agenda approach (Santos, Horta, 2018).

Considering the survey outcomes and the interviews analysis, we can affirm that the introduction of VQR (more slightly) and ASN (more strongly) are somehow influencing the aggregate "epistemological curve" of researchers working in Italian academia, partially as a result of the series of requirements that researchers have to comply with to be an active part of the entire academic system. In our opinion, it is too early to define or even guess the features of this phenomenon in the medium-term. However, according to the interviews analysis, this trend seems to remarkably concern researchers, who complain about the levelling of the research topics to few mainstream topics, and the neglecting of original, but challenging, topics. Hence, many concerns emerged on the effects of the introduction of ASN and the VQR on the overall quality of Italian academic production, since

adopting a more trailblazing approach is not "convenient" at all for researchers, especially the youngest.

To conclude, a final remark on what emerged from the recurring interactions we had with researchers during this period. An element that emerges loud and clear is that the academic community is still adapting to the introduction of the large-scale evaluation, and with many difficulties as well. A general distrust emerged toward the purpose and management of VQR and ASN (as confirmed also by the relatively high number of *apocalyptics* in our interviews), and they are often seen as a further complication in their already busy working lives. Research and teaching are sometimes perceived as harmed by VQR and ASN, and the high competitiveness (among colleagues and universities as well) introduced by both VQR and ASN exacerbates this feeling. Interestingly, many interviewees' reflections were mostly focussed on the introduction of ASN, which is having strong effects on young researcher's research practices and their conception of research and encourage them to focus more on the quantity rather than the quality of research production.

As mentioned in the third chapter, a major limitation to our study is that it is focussed on perception data, that may be biased the individual's perception and personal experience. Furthermore, the newness and complexity of the concepts proposed to the survey participants might have make even harder the comprehension of some concepts and questions.

A further limitation is due to the choice of targeting the entire social scientists community, without adopting a sampling approach, that may have determined a self-selection bias in the responses.

Since the topic and the approach adopted are rather new and unexplored, there are many new possibilities in the future to further this specific topic, both on the point of view of the method and on the topic itself. Indeed, it would be interesting to adopt both a comparative and a more in-depth approach: on the one hand, it would be interesting indeed to expand the target population to the entire Italian academic community, and, furthermore, to reflect on an international comparison with rather similar national research evaluation systems; on the other hand, it may be interesting to further the

issue of how the different dimensions of research agenda in the different academic contexts (such as, in a collaboration or multidisciplinary context).

References

Abramo, G., D'Angelo, C. A., & Rosati, F. (2014). Career advancement and scientific performance in universities. *Scientometrics*, *98*(2), 891-907.

Abramo, G., & D'Angelo, C. A. (2015). An assessment of the first "scientific habilitation" for university appointments in Italy. *Economia Politica*, *32*(3), 329-357.

Abramo, G., & D'Angelo, C. A. (2015). The VQR, Italy's second national research assessment: Methodological failures and ranking distortions. *Journal of the Association for Information Science and Technology*, *66*(11), 2202-2214.

Abramo, G., & D'Angelo, C. A. (2016). Refrain from adopting the combination of citation and journal metrics to grade publications, as used in the Italian national research assessment exercise (VQR 2011–2014). *Scientometrics*, *109*(3), 2053-2065.

Åkerlind, Gerlese S. (2008). An academic perspective on research and being a researcher: An integration of the literature. *Studies in higher education*, *33*(1), 17-31.

Anfossi, A., Ciolfi, A., Costa, F., Parisi, G., & Benedetto, S. (2016). Large-scale assessment of research outputs through a weighted combination of bibliometric indicators. *Scientometrics*, *107*(2), 671-683

Ancaiani, A., Anfossi, A. F., Barbara, A., Benedetto, S., Blasi, B., Carletti, V., ... & Sileoni, S. (2015). Evaluating scientific research in Italy: The 2004–10 research evaluation exercise. *Research Evaluation*, *24*(3), 242-255.

ANVUR. (2011). *Valutazione della Qualità della Ricerca 2004-2010 (VQR 2004-2010). Bando di partecipazione*. https://www.anvur.it/wp-content/uploads/2011/11/bando_vqr_def_07_11.pdf

ANVUR (2015). *Linee Guida VQR 2011-2014.*
https://www.anvur.it/wp-content/uploads/2015/07/dm_27_06_2015_vqr_11-14.pdf

ANVUR. (2019). *Report of the Group of Experts charged by ANVUR to advice on the process 'Valutazione della Qualità della Ricerca (VQR)*': http://www.anvur.it/wp-content/uploads/2019/03/High-Experts-Report-on-VQR.pdf

ANVUR. (2020). *Integrazione delle Linee guida per le valutazione della qualità della ricerca (VQR) 2015-2019.* https://www.anvur.it/wp-content/uploads/2020/09/Decreto-Ministeriale-n.444-dell11-08-2020.pdf

Article 33, Italian Constitution, https://www.senato.it/istituzione/la-costituzione/parte-i/titolo-ii/articolo-33

Atkinson, M. (2001). 'Peer review'culture. *Science and engineering ethics*, *7*(2), 193-204.

Babbie, E. R. (2020). The practice of social research. Cengage learning.

Baccini, A. (2014). La VQR di Area 13: una riflessione di sintesi. *Statistica & Società*, *3*(3), 32-37.

Baccini, A., & De Nicolao, G. (2016). Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise. *Scientometrics*, *108*(3), 1651-1671.

Barclay, S., Todd, C., Finlay, I., Grande, G., & Wyatt, P. (2002). Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaire studies of GPs. *Family practice*, *19*(1), 105-111.

Bellotti, E., Kronegger, L., & Guadalupi, L. (2016). The evolution of research collaboration within and across disciplines in Italian Academia. *Scientometrics*, *109*(2), 783-811.

Bethlehem, J. (2010). Selection bias in web surveys. *International statistical review*, *78*(2), 161-188.

Binswanger, M. (2014). Excellence by nonsense: The competition for publications in modern science. In *Opening Science* (pp. 49-72). Springer, Cham.

Bonaccorsi, A. (2015). La valutazione possibile, il Mulino.

Borrelli, D., & Stazio, M. (2018). La «grande trasformazione» dell'università italiana. *Rivista Trimestrale di SCIENZA dell'AMMINISTRAZIONE*, *1*, 1-19.

Bourdieu, P. (1988). *Homo academicus*. Stanford University Press.

Bozzon, R., Murgia, A., Poggio, B., & Rapetti, E. (2017). Work–life interferences in the early stages of academic careers: The case of precarious researchers in Italy. European Educational Research Journal, 16(2-3), 332-351.

Bruce, A. C., Tait, L. J., & Williams, R. (2004). Interdisciplinary integration in Europe: The case of the Fifth Framework programme. Futures, 36, 457–470

Bryman, A., Becker, S., & Sempik, J. (2008). Quality criteria for quantitative, qualitative and mixed methods research: A view from social policy. *International journal of social research methodology*, *11*(4), 261-276.

Burrows, R. (2012), 'Living with the h - index? Metric Assemblages in the Contemporary Academy', The Sociological Review, 60/2:355–72.

Butler, L. (2003). Explaining Australia's increased share of ISI publications—the effects of a funding formula based on publication counts. *Research policy*, *32*(1), 143-155.

Campbell, David F. (2013). New university governance: how the academic profession perceives the evaluation of research and teaching. In *The Work Situation of the Academic Profession in Europe: Findings of a Survey in Twelve Countries* (pp. 205-228). Springer, Dordrecht.

D'Amico, R., Vermigli, P., & Canetto, S. S. (2011). Publication productivity and career advancement by female and male psychology faculty: The case of Italy. *Journal of Diversity in Higher Education*, *4*(3), 175.

Cantwell, B. (2011). Transnational mobility and international academic employment: Gatekeeping in an academic competition arena. *Minerva*, *49*(4), 425-445.
Capano, G. (2014). The re-regulation of the Italian university system through quality assurance. A mechanistic perspective. *Policy and Society*, *33*(3), 199-213.
Carbonaro, M., & Bainbridge, J. (2000). Design and development of a process for web-based survey research. *Alberta Journal of Educational Research*, *46*(4).
Cerroni, A. (2009). Valutare la scienza sociale nell'epoca della società della conoscenza. *Quaderni di sociologia*, (49), 169-181.
Colarusso, S., & Giancola, O. (2020). *Università e nuove forme di valutazione: Strategie individuali, produzione scientifica, effetti istituzionali* (Vol. 59). Sapienza Università Editrice.
Colwell, R., Blouw, M., Butler, L., Cozzens, S. E., Feller, I., Gingras, Y., ... & Woodward, R. (2012). Informing research choices: Indicators and judgment. *The Expert Panel on Science Performance and Research Funding*, *142*.
Corbetta, P. (2014). *Metodologia e tecniche della ricerca sociale*. Bologna: il Mulino.
Coryn, C. L., Hattie, J. A., Scriven, M., & Hartmann, D. J. (2007). Models and mechanisms for evaluating government-funded research: An international comparison. *American Journal of Evaluation*, *28*(4), 437-457.
Creswell, J. W., Clark, V. L. P., Gutmann, M. L., & Hanson, W. E. (2003). ADVANCED MIXED. *Handbook of mixed methods in social & behavioral research*, 209.
Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications, Incorporated
Creswell, J. W. (2014). *Research design: qualitative, quantitative, and mixed methods approaches*. 4th ed. Thousand Oaks, California, SAGE Publications
Cullen, D. L., & Luna, G. (1993). Women mentoring in academe: Addressing the gender gap in higher education. Gender and Education, 5(2), 125-137.
Dahler-Larsen, P. (2012). Evaluation as a situational or a universal good? Why evaluability assessment for evaluation systems is a good idea, what it might look like in practice, and why it is not fashionable. *Scandinavian Journal of Public Administration*, *16*(3), 29-46.
De Paola, M., Ponzo, M., & Scoppa, V. (2015). Gender differences in attitudes towards competition: Evidence from the Italian scientific qualification. *Available at SSRN 2575039*.
De Philippis, M. D. (2015). Multitask agents and incentives: the case of teachingand research for university professors.
Di Benedetto, A. Un'analisi metodologica della valutazione dei prodotti nella VQR 2004-2010.
Draugalis, J. R., & Plaza, C. M. (2009). Best practices for survey research reports revisited: implications of target population, probability sampling, and response rate. *American journal of pharmaceutical education*, *73*(8).
Eco, U. (1994). *Apocalittici e integrati* (Vol. 26). Bompiani.
Elton, Lewis (2000). The UK research assessment exercise: unintended consequences. *Higher Education Quarterly*, *54*(3), 274-283.
Enders, J., De Boer, H., File, J., Jongbloed, B., & Westerheijden, D. (2011). Reform of higher education in Europe. In *Reform of higher education in Europe* (pp. 1-10). Brill Sense.
Entman, R. M. (1989). How the media affect what people think: An information processing approach. *The journal of Politics*, *51*(2), 347-370.
Ertmer, P. A., & Glazewski, K. D. (2014). Developing a research agenda: Contributing new knowledge via intent and focus. *Journal of Computing in Higher Education*, *26*(1), 54-68.
Eurostat. (2021). *R&D expenditure in the EU at 2.3% of GDP in 2020*. https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20211129-2
Eurostat. (2022). *Share of female researchers by sector of performance*. https://ec.europa.eu/eurostat/databrowser/view/rd_p_femres/default/bar?lang=en
Fasanella, A., & Martire, F. (2017). Considerazioni metodologiche sulla Vqr 2011-2014 e possibili sviluppi della valutazione. *Sociologia e ricerca sociale*.
Favale, T., Soro, F., Trevisan, M., Drago, I., & Mellia, M. (2020). Campus traffic and e-Learning during COVID-19 pandemic. *Computer networks*, *176*, 107290.
Fondazione CRUI. (2015). *L'utilizzo dei dati VQR per la valutazione dipartimentale negli atenei italiani Metodologie ed esperienze*
Franceschet, M., & Costantini, A. (2011). The first Italian research assessment exercise: A bibliometric perspective. *Journal of informetrics*, *5*(2), 275-291.
Franceschini, F., & Maisano, D. (2017). Critical remarks on the Italian research assessment exercise VQR 2011–2014. *Journal of informetrics*, *11*(2), 337-357.
Frey, B. S. (2003). Publishing as prostitution?–Choosing between one's own ideas and academic success. *Public choice*, *116*(1), 205-223.
Frey, B. S., Eichenberger, R., & Frey, R. L. (2009). Editorial ruminations: publishing Kyklos. *Kyklos*, *62*(2), 151-160.
Gadd, E. (2015). When are journal metrics useful? A balanced call for the contextualized and transparent use of all publication metrics. *Impact of Social Sciences Blog*.

Gambardella, D., Grimaldi, E., & Lumino, R. (2019). L'Università italiana e i paradossi della valutazione: processi di depoliticizzazione e spazi di riflessività. *D'Albergo E., Moini E., a cura di, Politica e azione pubblica nell'epoca della depoliticizzazione. Roma: Sapienza Università Editrice*.
Geuna, A., & Martin, B. R. (2003). University research evaluation and funding: An international comparison. *Minerva*, *41*(4), 277-304.
Laudel, G., & Gläser, J. (2007). Interviewing scientists. *Science, Technology & Innovation Studies*, *3*(2).
Gläser, Jochen, and Grit Laudel. "On interviewing "good" and "bad" experts." *Interviewing experts* (2009): 117-137.
Gläser, J., & Laudel, G. (2016). Governing science: How science policy shapes research content. *European Journal of sociology/Archives Européennes de sociologie*, *57*(1), 117-168.
Greco, G., & Scarcello, F. (2013). Fair division rules for funds distribution: The case of the Italian Research Assessment Program (VQR 2004–2010). *Intelligenza Artificiale*, *7*(1), 45-56.
Grisorio, M. J., & Prota, F. (2020). Italy's national research assessment: Some unpleasant effects. *Studies in Higher Education*, *45*(4), 736-754.
Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology* (Vol. 561). John Wiley & Sons.
Hammarfelt, Björn, and De Rijcke, Sarah (2015). Accountability in context: Effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of Arts at Uppsala University. *Research Evaluation*, *24*(1), 63-77.
Henkel, Mary (2005). Academic identity and autonomy in a changing policy environment. *Higher education*, *49*(1), 155-176.
Hessels, Laurens K., and Harro Van Lente. "Re-thinking new knowledge production: A literature review and a research agenda." *Research policy* 37.4 (2008): 740-760.
Hicks, D. (2012). Performance-based university research funding systems. *Research policy*, *41*(2), 251-261.
Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429-431.
Hills, P. V., & Dale, A. J. (1995). Research and technology evaluation in the United Kingdom. *Research Evaluation*, *5*(1), 35-44.
Hood, C., & Jackson, M. (1992). The new public management: a recipe for disaster. *Hazard management and emergency planning: Perspectives on Britain*, 109-125.
Horta, H., & Santos, J. M. (2016). An instrument to measure individuals' research agenda setting: the multi-dimensional research agendas inventory. *Scientometrics*, *108*(3), 1243-1265.
Santos, J. M., & Horta, H. (2018). The research agenda setting of higher education researchers. *Higher Education*, *76*(4), 649-668.
Horta, H., & Santos, J. M. (2020). Organisational factors and academic research agendas: an analysis of academics in the social sciences. *Studies in Higher Education*, *45*(12), 2382-2397.
Santos, J. M., & Horta, H. (2020). The association between researchers' conceptions of research and their strategic research agendas. *The association between researchers' conceptions of research and their strategic research agendas*, (4), 56-74.
Istat. (2022) *Livelli di istruzione e ritorni occupazionali. Anno 2021*. https://www.istat.it/it/archivio/276497
Italy, Commissione Affari Costituzionali, Commissione Cultura, Scienza e Istruzione. (1989) *SEDUTA DI MERCOLEDI' 29 MARZO 1989.*
http://legislature.camera.it/_dati/leg10/lavori/stencomm/0107/Leg/Serie010/1989/0329/stenografico.pdf
Italy, Commissione Affari Costituzionali, Commissione Cultura, Scienza e Istruzione. (1989) *SEDUTA DI MERCOLEDI' 5 APRILE 1989.*
Italy, Commissione VII Cultura, (2010), Bollettino delle giunte e delle Commissioni Parlamentari, Seduta del 15 Settembre 2010,
http://documenti.camera.it/leg16/resoconti/commissioni/bollettini/html/2010/09/15/07/comunic.htm#data.20100915.com07.bollettino.sede00020.tit00010
Italy, Commissione VII Cultura, (2010), Bollettino delle giunte e delle Commissioni Parlamentari, Seduta del 21 Settembre 2010
http://documenti.camera.it/leg16/resoconti/commissioni/bollettini/html/2010/09/21/07/comunic.htm#data.20100921.com07.bollettino.sede00020.tit00020
Italy, Compiti e Istituzioni, Ministero dell'Istruzione, dell'Università e della Ricerca (http://www.cnvsu.it/ilcomitato/compiti.asp).
http://legislature.camera.it/_dati/leg10/lavori/stencomm/0107/Leg/Serie010/1989/0405/stenografico.pdf
Italy, *Linee Guida per la Valutazione della Ricerca*, CIVR, http://web.mclink.it/MJ4450/criteri-valutazione-ricerca.pdf
Italy, Osservatorio per la valutazione del sistema universitario (1997), *Incontro Nazionale sulla Valutazione del Sistema Universitario*. http://www.cnvsu.it/_library/downloadfile.asp?id=10648
Italy, *Resoconto della I Commissione Permanente* (1996) Resoconto del 18 Dicembre 1996, http://leg13.camera.it/_dati/leg13/lavori/bollet/frsmcdin.asp?percboll=/_dati/leg13/lavori/bollet/199612/1217/html/01/&pagpro=38n2&all=off&commis=01

Italy, *Resoconto della I Commissione Permanente* (1997) Resoconto del 14 Gennaio 1996, http://leg13.camera.it/_dati/leg13/lavori/bollet/frsmcdin.asp?percboll=/_dati/leg13/lavori/bollet/199701/0114/html/01/&pagpro=23n1&all=on&commis=01
Italy, *Resoconto della I Commissione Permanente* (1997) Resoconto del 16 Gennaio 1996, http://leg13.camera.it/_dati/leg13/lavori/bollet/frsmcdin.asp?percboll=/_dati/leg13/lavori/bollet/199701/0116/html/01/&pagpro=34n2&all=off&commis=01
Italy, *Resoconto della I Commissione Permanente* (1997) Resoconto del 21 Gennaio 1996, http://leg13.camera.it/_dati/leg13/lavori/bollet/frsmcdin.asp?percboll=/_dati/leg13/lavori/bollet/199701/0121/html/01/&pagpro=4n2&all=off&commis=01
Italy, *Resoconto della VII Commissione permanente* (2006), Resoconto del 11 Ottobre 2006, http://leg15.camera.it/_dati/leg15/lavori/bollet/frsmcdin.asp?percboll=/_dati/leg15/lavori/bollet/200610/1011/html/07/&pagpro=86n2&all=off&commis=07
Italy, *Resoconto della VII Commissione permanente* (2006), Resoconto del 12 Ottobre 2006, http://leg15.camera.it/_dati/leg15/lavori/bollet/frsmcdin.asp?percboll=/_dati/leg15/lavori/bollet/200610/1012/html/07/&pagpro=84n3&all=off&commis=07
Italy, *7^ Commissione permanente (Istruzione pubblica, beni culturali) in sede referente* (2010), Resoconto N. 267 del 16 Dicembre 2010
https://www.senato.it/japp/bgt/showdoc/frame.jsp?tipodoc=SommComm&leg=16&id=515617&part=doc_dc-sedetit_isr
Jones, T. L., Baxter, M. A. J., & Khanduja, V. (2013). A quick guide to survey research. *The Annals of The Royal College of Surgeons of England*, *95*(1), 5-7.
Karlsson, S. (2017). Evaluation as a travelling idea: Assessing the consequences of Research Assessment Exercises. Research Evaluation, 26(2), 55-65.
Karvalics, L. Z. (2013, October). From scientific literacy to lifelong research: a social innovation approach. In *European Conference on Information Literacy* (pp. 126-133). Springer, Cham.
Karlsson, S. (2017). Evaluation as a travelling idea: Assessing the consequences of Research Assessment Exercises. Research Evaluation, 26(2), 55-65.
Kitchenham, B. A., & Pfleeger, S. L. (2002). Principles of survey research part 2: designing a survey. *ACM SIGSOFT Software Engineering Notes*, *27*(1), 18-20.
Kitchenham, B., & Pfleeger, S. L. (2002). Principles of survey research part 4: questionnaire evaluation. *ACM SIGSOFT Software Engineering Notes*, *27*(3), 20-23.
Kitchenham, B., & Pfleeger, S. L. (2002). Principles of survey research: part 5: populations and samples. *ACM SIGSOFT Software Engineering Notes*, *27*(5), 17-20.
Knuuttila, Tarja. "Science in a new mode: good old (theoretical) science versus brave new (commodified) knowledge production?." *Science & Education* 22.10 (2013): 2443-2461.
Kuhn, T. S. (1970). *The structure of scientific revolutions* (Vol. 111). University of Chicago Press: Chicago.
Laudel, G. (2006). The art of getting funded: how scientists adapt to their funding conditions. *Science and Public Policy*, *33*(7), 489-504.
Lamont, M. (2009). *How professors think: Inside the curious world of academic judgment*. Harvard University Press.
Latour, B., Woolgar, S., (1979), Laboratory life. The social construction of scientific facts, Cambridge University Press
Laudel, G. (2006). The art of getting funded: how scientists adapt to their funding conditions. *Science and Public Policy*, *33*(7), 489-504.
Leahey, Erin (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. *American sociological review*, *72*(4), 533-561.
Leisyte, Liudvika, Enders, Jürgen, and De Boer, Harry. (2008). The freedom to set research agendas—illusion and reality of the research units in the Dutch universities. *Higher education policy*, *21*(3), 377-391.
Lepori, B., Van den Besselaar, P., Dinges, M., Potì, B., Reale, E., Slipersæter, S., ... & Van der Meulen, B. (2007). Comparing the evolution of national research policies: what patterns of change?. *Science and Public Policy*, *34*(6), 372-388.
Lepori, B., Reale, E., & Spinello, A. O. (2018). Conceptualizing and measuring performance orientation of research funding systems. *Research Evaluation*, *27*(3), 171-183.
M. Lewis, J. (2015). The politics and consequences of performance measurement. Policy and Society, 34(1), 1-12.
Limoges, C., Scott, P., Schwartzman, S., Nowotny, H., & Gibbons, M. (1994). The new production of knowledge: The dynamics of science and research in contemporary societies. *The New Production of Knowledge*, 1-192.
Lombi, L. (2015). *Le web survey* (pp. 1-128). Franco Angeli.
Luukkonen, T., & Thomas, D. A. (2016). The 'negotiated space'of university researchers' pursuit of a research agenda. *Minerva*, *54*(1), 99-127.
Marini, G. (2017). New promotion patterns in Italian universities: Less seniority and more productivity? Data from ASN. *Higher Education*, *73*(2), 189-205.

Marini, G., & Meschitti, V. (2018). The trench warfare of gender discrimination: evidence from academic promotions to full professor in Italy. *Scientometrics*, *115*(2), 989-1006.
Martin, B. R. (2011). The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster?. *Research evaluation*, *20*(3), 247-254.
McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public opinion quarterly*, *36*(2), 176-187.
McGill, Monica M., and Settle, Amber (2012). Identifying effects of institutional resources and support on computing faculty research productivity, tenure, and promotion. *International journal of doctoral studies*, *7*, 167-198.
McGinn, Michelle K. (2012). Being Academic Researchers: Navigating Pleasures and Pains in the Current Canadian Context. *Workplace: A Journal for Academic Labor*, (21).
Merton, R. K. (1942). A note on science and democracy. *J. Legal & Pol. Soc.*, *1*, 115.
Minelli, E., Rebora, G., Turri, M., & Huisman, J. (2006). The impact of research and teaching evaluation in universities: comparing an Italian and a Dutch case. *Quality in Higher Education*, *12*(2), 109-124.
Minelli, E., Rebora, G., & Turri, M. (2008). 12. The structure and significance of the Italian research assessment exercise (VTR). *European Universities in Transition: Issues, models and cases*, 221.
Middlehurst, Robin. "Higher education research agendas for the coming decade: a UK perspective on the policy–research nexus." *Studies in Higher Education* 39.8 (2014): 1475-1487.
Morawska, E. (2018). Qualitative migration research: Viable goals, open-ended questions, and multidimensional answers. In *Qualitative research in European migration studies* (pp. 113-131). Springer, Cham.
Moscati, R. (2001). Italian university professors in transition. *Higher Education*, *41*(1), 103-129.
Moss, Gloria, and Kubacki, Krzysztof (2007). Researchers in higher education: a neglected focus of study?. *Journal of further and higher education*, *31*(3), 297-310.
Musselin, C. (2004). Towards a European academic labour market? Some lessons drawn from empirical studies on academic mobility. *Higher Education*, *48*(1), 55-78.
Musselin, C. (2013). How peer review empowers the academic profession and university managers: Changes in relationships between the state, universities and the professoriate. *Research Policy*, *42*(5), 1165-1173.
Neave, G. (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*, 7-23.
Ochsner, M., Kulczycki, E., & Gedutis, A. (2018, September). The diversity of european research evaluation systems. In *STI 2018 Conference Proceedings* (pp. 1235-1241). Centre for Science and Technology Studies (CWTS).
Ochsner, M., Kulczycki, E., Gedutis, A., & Peruginelli, G. (2020). 2.3 National research evaluation systems. *Handbook of Bibliometrics*, 99-106.
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.
Osuna, C., Cruz-Castro, L. and Sanz-Menendez, L. (2011), 'Overturning Some Assumptions about the Effects of Evaluation Systems on Publication Performance', Scientometrics, 86: 575–92.
Paradeise, C., Reale, E., Bleiklie, I., & Ferlie, E. (Eds.). (2009). *University governance*. Dordrecht: Springer Netherlands.
Van Der Meulen, B., Whitley, R., & Glaser, J. (2007). The Changing Governance of the Sciences.
Paradeise, Chaterine, Thoenig, Jean-Claude (2013). Academic institutions in search of quality: Local orders and global standards. *Organization studies*, *34*(2), 189-218.
Patti, F. P. (2019). Career Paths into Legal Academia in Italy. *Bocconi Legal Studies Research Paper*, (3895852).
Pinar, M., & Horne, T. J. (2022). Assessing research excellence: evaluating the Research Excellence Framework. *Research Evaluation*, *31*(2), 173-187.
Pinheiro, D., Melkers, J., & Youtie, J. (2014). Learning to play the game: Student publishing as an indicator of future scholarly success. *Technological Forecasting and Social Change*, *81*, 56-66.
Polanyi, M. (2012). *Personal knowledge*. Routledge.
Potì, B., & Reale, E. (2009). *ERAWATCH Country Reports 2009: Italy. Analysis of Policy Mixes to Foster R&D Investment and to Contribute to the ERA* (No. JRC53706). Joint Research Centre (Seville site).
Rauhut, H., Johann, D., Jerke, J., Rathmann, J., & Velicu, A. (2020). The Zurich survey of academics: Methods, design, and data.
Reale, E., and Marini, G. (2017). The Transformative Power of Evaluation on University Governance. In *Managing Universities* (pp. 107-137). Palgrave Macmillan, Cham.
Reale, E., & Pennisi, C. (2013). La valutazione dell'Università e della ricerca in Italia: stato dell'arte e prospettive. *Prove di valutazione. Libro Bianco sulla valutazione in Italia*, 23-62.
Reese, S. D., & Danielian, L. H. (1989). A closer look at intermedia influences on agenda setting: The cocaine issue of 1986. *Communication campaigns about drugs: Government, media, and the public*, 47-66.
Rijcke, S. D., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. *Research evaluation*, *25*(2), 161-169.
Salter, A. J., & Martin, B. R. (2001). The economic benefits of publicly funded basic research: a critical review. *Research policy*, *30*(3), 509-532.

Reale, E., Barbara, A., & Costantini, A. (2007). Peer review for the evaluation of academic research: lessons from the Italian experience. *Research evaluation*, *16*(3), 216-228.
Reale, Emanuela, and Seeber, Marco (2011). Organisation response to institutional pressures in Higher Education: the important role of the disciplines. *Higher Education*, *61*(1), 1-22.
Rossi, P. (2016). Luci e ombre della VQR.
Rostan, M. (Ed.). (2011). *La professione accademica in Italia: Aspetti, problemi e confronti nel contesto europeo*. LED Edizioni Universitarie.
Santos, J. M. D. C. S. M. D. (2020). The science of scientific agendas: insights into the agenda setting process of researchers.
Santos, J. M., Horta, H., & Amâncio, L. (2021). Research agendas of female and male academics: a new perspective on gender disparities in academia. *Gender and Education*, *33*(5), 625-643.
Schlipphak, B., & Isani, M. (2020). Designing survey questions and choosing survey formats. In *Handbuch Methoden der Politikwissenschaft* (pp. 351-371). Springer VS, Wiesbaden.
Scott Metcalfe, A., & Padilla González, L. (2013). Underrepresentation of women in the academic profession: A comparative analysis of the North American region. NASPA Journal About Women in Higher Education, 6(1), 1-21.
Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. Teachers college press.
Shattock, Michael (2014). Can we still speak of there being an academic profession?. *History of Education*, *43*(6), 727-739.
Shattock, M. (2014). The context of 'modernising'reforms in university governance. In *International trends in university governance* (pp. 15-28). Routledge.
Shore, Cris, and Wright, Susan (2004). Whose accountability? Governmentality and the auditing of universities. *parallax*, *10*(2), 100-116.
Small, H. (2004). On the shoulders of Robert Merton: Towards a normative theory of citation. *Scientometrics*, *60*(1), 71-79.
Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International journal of public administration*, *18*(2-3), 277-310.
Statistics.com, *Collinearity*, *https://www.statistics.com/glossary/collinearity/*
Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, *133*(5), 859.
Statnano. *Research in R&D (per million people)*. https://statnano.com/report/s90
Van Den Besselaar, P., Hemlin, S., & Van Der Weijden, I. (2012). Collaboration and competition in research. *Higher Education Policy*, *25*(3), 263-266
Vaughan Dickson, V., Lee, C. S., & Riegel, B. (2011). How do cognitive function and knowledge affect heart failure self-care?. *Journal of Mixed Methods Research*, *5*(2), 167-189.
Vessuri, Hebe (2008). The role of research in higher education: Implications and challenges for an active future contribution to human and social development
Walker, K. A. (2018). Gender gap in professional entomology: Women are underrepresented in academia and the US government. Annals of the Entomological Society of America, 111(6), 355-362.
Wang, X. (2022, May). Italian Research Assessment VQR: Framework, Achievements and Controversies. In *Proceedings of the 7th International Conference on Distance Education and Learning* (pp. 261-267).
Watermeyer, Richard, and Olssen, Mark (2016). 'Excellence' and exclusion: the individual costs of institutional competitiveness. *Minerva*, *54*(2), 201-218.
Weinberg, A. M. (1963). Criteria for scientific choice. *Minerva*, 159-171.
Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences?. *Scientometrics*, *62*(1), 117-131.
Whitley, R. (2007). Changing governance of the public sciences. In *The changing governance of the sciences* (pp. 3-27). Springer, Dordrecht.
Whitley, R., Gläser, J., & Engwall, L. (Eds.). (2010). *Reconfiguring knowledge production: Changing authority relationships in the sciences and their consequences for intellectual innovation*. OUP Oxford.
Whitley, R., Gläser, J., & Laudel, G. (2018). The impact of changing funding and authority relationships on scientific innovations. *Minerva*, *56*(1), 109-134.
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.
Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. *Minerva*, 66-100.

Annex 1: Survey on the effects of research evaluation on researchers' research agenda setting

1. Academic position
   - Full professor
   - Associate professor
   - Temporary researcher
   - Full time researcher

2. Gender
   - Male
   - Female
   - I'd rather not answer

3. Nationality
   - Italian
   - European country (UE27 + UK, CH, NO)
   - Extra-European country

4. Indicate your age

5. Academic seniority
   - Temporary researcher
   - Till 5 years
   - From 6 to 10 years
   - Over 11 years

6. Indicate the dimension of your afferent university per number of enrolled students
   - Till 10.000 students
   - From 10.001 to 20.000 students
   - From 20.001 to 40.000 students
   - Over 40.001 students

7. Indicate if your afferent university is public or private
   - Public university
   - Private university

8. Afferent CUN sector o
   - 13/A Economics
   - 13/B Business economics
   - 13/C Economic History
   - 13/D Statistics
   - 14/A Political Theory
   - 14/B Political History
   - 14/C Sociology
   - 14/D Applied Sociology

9. Indicate or level of agreement per each question

	Totally agree	Somewhat Agree	Somewhat Disagree	Totally disagree
My specialization is focused on a single discipline				
Conducting research in new fields is not part of my plans				
Multidisciplinary research is more interesting that one-discipline-oriented research				
I prefer working with multidisciplinary teams than with colleagues from one singe discipline				

10. Indicate or level of agreement per each question

	Totally agree	Somewhat Agree	Somewhat Disagree	Totally disagree
The availability of funds for research on a specific field does not influence my decision on conducting research on that topic				
My publications are improved by collaboration with other authors				

11. Indicate or level of agreement per each questions

	Totally agree	Somewhat Agree	Somewhat Disagree	Totally disagree
I structure my research agenda to let my works being selected by my university for the evaluation				
Having to collocate my products into disciplinary fields determined by ANVUR binds my research agenda				
I rather prefer "innovative" research to "safe research", even when the				

probability of success are low				

12. Indicate your opinion on this sentence: “I often search for colleagues with who I can collaborate on my publications”.
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

13. How do you think that collaboration with peers from your disciplinary field has influences your research agenda setting?
    - Collaboration with colleagues have not influenced my research agenda.
    - Collaboration with colleagues pushed me to change the planning of my research agenda.
    - Collaboration with colleagues pushed my to partially change the planning of my research agenda.

14. How do you think that collaboration with peers from other disciplinary fields has influences your research agenda setting?
    - Collaboration with colleagues have not influenced my research agenda.
    - Collaboration with colleagues pushed me to change the planning of my research agenda.
    - Collaboration with colleagues pushed me to partially change the planning of my research agenda.

15. More comments on the effects of collaboration with colleagues on the structuring on your research agenda.

16. Indicate your opinion on the following statement: “I have frequently oriented my research agenda to gain social recognition from my peers”:
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

17. Do you think that conducting research on a mainstream topic increases your chances that your work will be published? (for mainstream topic we mean a prevalent and very common topic in the scientific community debate)
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

18. Do you think that current events are relevant when programming your research agenda? (for instance, immigration, Covid-19 pandemic)
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

19. Do you think that societal needs are relevant when programming your research agenda? (for example, unemployment, poverty, social inequalities)
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

20. Have you ever participate to calls for research funding?
    - Yes
    - No

21. Do you think that attending competitive calls has influenced your research agenda?
    - I have totally integrated in my research agenda the topics addressed in the calls for research funding I participated to.
    - I have partially integrated in my research agenda the topics addressed in the calls for research funding I participated to.
    - Participating to calls for research funding have not influenced my research agenda setting.

22. How do you think that attending competitive calls has influenced you research agenda?
    - Participating to calls for research funding left me less time to conduct research on my research topics as usual.
    - Participating to calls for research funding induced me to change the planning of my research agenda.

23. More comments on how participating to calls for research funding influenced my research agenda setting.

24. "My research agenda has been influenced by the introduction of ASN":
    - Yes
    - No

25. Has the introduction of ASN influenced the choice of topics to research?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

26. Has the introduction of ASN influenced your motivation for researching?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

27. Has the introduction of ASN influenced your research practices?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

28. Has the introduction of ASN let you address a different peers community?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

29. More comments on the influence of ASN on your research agenda setting.

30. Has your research agenda been influenced by the VQR?
    - Yes
    - No

31. Has the introduction of VQR influenced the choice of topics to research?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

32. Has the introduction of VQR influenced your motivation for researching?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

33. Has the introduction of VQR influenced your research practices?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

34. Has the introduction of VQR let you address a different peers community?
    - Yes, totally
    - Somewhat
    - To a small extent
    - Not at all

35. More comments on the influence of VQR on your research agenda setting.

36. More comments on the factors that have influenced your research agenda setting.

37. We sincerely thank you for your collaboration. Please insert your email address below if you are available to participate in the future to the second phase of our research to deepen the topic at issue.

Annex 2: Protocols of the interviews with the experts

1. Free comment on the influence of ASN and VQR on researchers' research agenda setting.
2. Eventual pressure of the afferent university to conduct research in a determined field.
3. Free comment on the choice and selection by ANVUR of scientific journals on researchers' research agenda setting.
4. VQR and ASN influence on research attitude toward multidisciplinary research.
5. VQR and ASN influence on research attitude toward internationalization of research.
6. VQR and ASN influence on researchers' preference on short or long-term research projects.
7. VQR and ASN influence on researchers' conception of research.
8. VQR and ASN influence on scientific collaboration.
9. VQR and ASN influence on scientific collaboration with young researchers.
10. VQR and ASN influence on universities' resources funds management.

Acknowledgments

I would like to express my deepest appreciation to my tutor and the CNR-IRCRES team for the support for the finalization of the work, from the very beginning to the final phase.
I would also like to express my sincere thanks to Professor Maria Grazia Galantino and Professor Felice Addeo, who kindly helped me during the process, by providing concrete support, especially during the first phase of the research.
Special thanks, moreover, to all the researchers and professors who actively participated to the research and expressed their enthusiasm in the topic of the research.
Lastly, I'd like to thank my helpful families, my supportive partner, and friends, with a special mention to those people who made this work possible, by proposing projects and ideas that I would have never thought without them.