

# Monetary incentives versus public funding in healthcare research: what matters the most?

Gabriele Letta<sup>a</sup> Luca Salmasi<sup>a</sup> Gilberto Turati<sup>a</sup>

<sup>a</sup> Department of Economics and Finance, Università Cattolica del Sacro Cuore, Milan

May 2025

## Abstract

This paper studies the impact of two policy interventions on scientific research productivity in a major private Italian hospital. The first is a performance-based monetary incentive program (a Management-By-Objectives, or MBO, bonus) introduced by the hospital management to reward non-academic physicians for publishing research. The second is the hospital's recognition as an *Istituto di Ricovero e Cura a Carattere Scientifico* (IRCCS), which allowed its academic medical researchers to access dedicated public research funding. Using detailed panel data on physicians' publications from 2012 to 2022, we employ several difference-in-differences strategies to evaluate each policy's effect. We find that the introduction of monetary incentives did not lead to any significant impact on research output of the previously less research-active (non-academic) physicians, unless they were both treated by the MBO policy and had also access to IRCCS funding ("double-treated"). The IRCCS recognition caused instead a major boost in the publication rates of academic doctors and Medical Directors affected by both policies. We also assess the impact of such policies on the research quality of the hospital, by accounting for citations. We document increased cross-collaboration between the monetary-incentivized groups, indicating the emergence of knowledge spillovers; however, such increase was quite substantial for the non-academic doctors, while being modest for structured researchers in relative terms. Our findings might be able to inform the design of policies to incentivize research in healthcare organizations, highlighting the lower significance of performance-based incentives in absence of adequate research funding means.

**Keywords:** Healthcare management, Medical research, Monetary incentives, Management-By-Objective, Public funding to research.

**JEL Classification:** I10, I23.

# 1 Introduction

The increasing pressure borne by European healthcare systems is a relevant issue nowadays. As a matter of fact, given the demographic transition, the aging population is driving up the utilization of healthcare and its costs (Tang et al., 2022), bringing about major consequences in terms of macroeconomic growth (Acemoglu and Restrepo, 2017; Aksoy et al., 2019) and labor markets (Hao et al., 2024). As health impairs with age, chronic diseases become more prevalent, which boosts further long-term care expenses (D. E. Bloom et al., 2020; S. Chen et al., 2024; Hacker, 2024; Ye et al., 2023). In addition to such structural issues, the COVID-19 outbreak highlighted the need for a better resilience of health systems to epidemics (D. E. Bloom et al., 2022; S. Chen et al., 2021; Wu and Wang, 2024) which, joint with the other mentioned factors, contribute to a constant increase in health expenditures, in doing so challenging the sustainability of the systems.

In such regards, scientific research and innovation in medical practices turn out to be crucial for the enhancement of the efficiency and effectiveness of healthcare. Whereas the short-term impact of new treatments on efficiency is not to be taken for granted (Grant and Buxton, 2018), as innovative health technologies may initially raise costs (Chandra and Skinner, 2012), a steady and consistent research activity is expected to enhance the quality of care, by improving the long-run health outcomes of the affected population, fostering prevention, and triggering cost savings in a lengthened perspective. This is the reason why a high-standard medical research activity is usually favored by institutions, and healthcare facilities are often ranked and assessed by third-parties in accordance to the quality of the research they are able to carry over.

At the hospital level indeed, research may yield multiple benefits. Research-active institutions are better able to attract and retain high-quality medical professionals (AMS, 2020; Maynou et al., 2024); in addition, if they are able to “translate” research outcomes into medical practices, they tend to improve clinical standards of care (Barrenho et al., 2021, 2025). Such hospitals can also develop comparative advantages that ultimately benefit patient outcomes (N. Bloom, Propper, et al., 2015; N. Bloom et al., 2020; Ghandour et al., 2022). These potential benefits notwithstanding, conducting research in health-care settings requires strong incentives and huge resources, as clinical duties often take priority over the rest. Which raises a key policy question: how should individual institutions incentivize physicians to engage in research?

In academic environments, physicians and researchers are inherently motivated to publish by career-

oriented motives (tenure, promotions, prestige), as well as altruistic reasons or due to scientific curiosity (Rousseau et al., 2021). These *academic incentives* mean that even without a direct monetary rewarding of scientific contributions, academic clinicians conduct research “by definition” to progress in their job and, by doing that, they tend to improve care to a broad extent, although not necessarily for the nearer environment they are involved with, liked patients with whom they have direct interaction. This may raise even further concerns regarding the unintended consequences of placing too much weight on research activity. Achieving a higher research productivity in academic institutions is also simplified by the availability of funding and resources, usually granted through public funds. Outside academia, such incentives may be weaker, which is why additional mechanisms might be necessary to trigger research activities among practitioners.

We provide with empirical evidence on two different approaches to foster research in healthcare: performance-based monetary incentives versus increased public funding. We leverage the case study of a large, private, Italian teaching hospital, where both types of policies were implemented in quick succession. First, in 2017 the hospital introduced a performance-based Management-By-Objectives (MBO) bonus program, aimed at financially rewarding the non-academic hospital physicians for publishing papers within an yearly time-span. One year later, in early 2018, after succeeding in a recognition procedure ushered concomitantly with the MBO introduction, the hospital obtained the IRCCS status (Scientific Research Hospital designation), which led to additional government funding for research, which mainly benefited the hospital’s academic doctors. Our research design provides with a unique comparison between a direct monetary incentive for individuals and a broader institutional research funding intervention, within the same organization and time frame.

The present work also contributes to several strands of literature. First, being the MBO a form of corporate policy devised by the hospital management, we relate to the literature about management quality and incentives in imperfect markets and public services. Prior studies underlined the importance of management practices in hospitals (N. Bloom, Propper, et al., 2015; N. Bloom et al., 2020; Goodall, 2011) and schools (Muralidharan and Sundararaman, 2011, N. Bloom, Lemos, et al., 2015), as well as the impact of higher competition and incentive reforms in the healthcare sector (N. Bloom, Propper, et al., 2015; Gaynor et al., 2012, 2013; Kessler and McClellan, 2000; Longo et al., 2017; Propper et al., 2004). As a matter of fact, performance-based incentives in the public sector have been shown to improve relevant outcomes in some contexts (Burgess et al., 2017; Dal Bó et al., 2013). In the medical field, numerous studies examine how physicians respond to financial incentives (Bertoli and Grembi, 2019;

Brosig-Koch et al., 2024; Gruber and Owings, 1996; Molitor, 2018; Shurtz, 2013, 2014). We add to this literature by evaluating a performance-pay scheme for hospital physicians aimed at research outputs, rather than clinical outcomes and practices, which is the commonest metrics for the above literature.

Second, our study contributes to the literature on research funding and scientific productivity. Several studies have investigated the returns to public R&D grants and funding on academic research output and innovation. For instance, prior scholars have inquired on how grant funding affects scientists' subsequent publications, citations, entrepreneurship, and mobility (Azoulay et al., 2011; Babina et al., 2023; Banal-Estañol et al., 2023; Baruffaldi et al., 2020; Benavente et al., 2012; Ganguli, 2017; Ghirelli et al., 2023; Jacob and Lefgren, 2011). While additional funding often correlates with higher research productivity, the causal evidence is mixed, as some authors find weak or non-existent effects of grants on publications (Jacob and Lefgren, 2011). We fit into this research setting by examining the effect of a new modality of granting public funding to medical research (i.e., the *IRCCS* recognition) on the research output of academic physicians, and by directly comparing it to the effect of the incentive-based policy rewards.

Third, we refer to the literature on knowledge spillovers and collaboration networks in science. Academic outcomes can indeed be shaped by peer effects, like exposure to productive researchers or co-authorship networks (Azoulay et al., 2010; Bosquet et al., 2022; Brogaard et al., 2014; Colussi, 2018; Waldinger, 2010). By focusing on the interactions between the two different groups of physicians (those eligible for the MBO bonus and those receiving the IRCCS public funding), we study whether incentives for one group spill over to the other. The intuition behind it is that the MBO scheme might lead doctors not affiliated to universities to seek for collaborations with academic physicians, in doing this affecting their output indirectly, and vice versa.

We provide evidence on three main points. First, we compare the effectiveness of private performance-based monetary incentives versus public, career-oriented funding incentives in stimulating research output. Second, we examine individual determinants of research productivity in a private hospital operating in an "imperfect" market (non-profit healthcare), where academic motivations may be less dominant. Third, we assess potential research spillovers resulting from the combination of these two policy interventions, including increased cross-group co-authorship. To our knowledge, this is the first study to jointly evaluate a direct monetary reward program and a public research funding boost within the same setting. To preview our findings, we find that the MBO-monetary incentive did not have any significant impact on the publication productivity of non-academic physicians, without being able to narrow the gap between them and their university counterparts. On the other side, the IRCCS recognition and

associated funding increased research output among the hospital’s academic physicians by more than 100%. A significant, positive effect of the MBO is observed only when both incentives applied (i.e. non-academic physicians included in the group of IRCCS-funds’ recipients), which would suggest the effects were cumulative. However, the impact is relevant only with respect to individuals subject to the incentive scheme only or subject to no scheme at all, turning negative and statistically significant when comparing such group to academic physicians only subject to the public-funding improved access. We also retrieve that the policies led to greater collaboration between the two groups of researchers, with the non-academic physicians seeking for collaboration with IRCCS-recipients, while no effect is observed on citations overall. A battery of robustness checks validate the fact that results are not driven by pre-existing trends, by the COVID-19 shock, or by compositional changes and spillovers.

The remainder of the paper is organized as follows. Section 2 focuses on the institutional framework, describing the hospital and the two implemented policies. Section 3 dwells on the data sources and sample construction. Section 4 presents the analysis of the MBO incentive, including the empirical strategy, the main results, and some additional findings achieved through a triple-difference estimation. Section 5 focuses on the IRCCS funding and its effects. Section 6 reports several robustness and validation checks. Section 7 examines other outcomes such as collaborative publications and citation counts. Section 8 concludes. Additional figures and results are provided in the Appendix.

## 2 Institutional Framework

The study is set in a leading private hospital in Rome, Italy, affiliated with a prestigious private Italian university. The hospital is a major healthcare hub which offers both clinical services and medical education. It is also renowned for research: in recent years it has ranked among the top hospitals in Italy for research output. As a matter of fact, it has been consistently listed among the top 50 hospitals worldwide over the recent years, and more than 100 of its affiliated researchers are included in the top 2% ranking of scientists globally (according to standardized citation metrics). These characteristics make it an ideal environment for the studying of policies aimed at boosting research activity. In 2017–2018, the hospital underwent two significant research-related policy changes. First, in 2017 (in anticipation of a forthcoming evaluation for obtaining research status), the hospital management implemented a performance-based MBO policy to encourage publication by hospital physicians who did not have university affiliations (“non-academic” physicians). This MBO program (which has been active every year

since 2017) has offered monetary bonuses for publications: each physician would receive a payout for each peer-reviewed journal article published in a given year, with the amount of the payment being proportional to the journal's Impact Factor. Specifically, for every Impact Factor point of a published article (as indexed by the Web of Science website), the physician earned €500, up to a maximum bonus of €10,000 per year (equivalent to 20 points). Thus, publishing in higher-impact journals yielded larger rewards. The total annual budget allocated to the MBO bonus fund was €1 million. Only physicians employed as medical doctors (clinicians) without a university faculty affiliation were eligible for this scheme. Academic faculty physicians were excluded from MBO by definition. The goal of this program was to foster research activity among clinicians who traditionally focused on patient care and had lower research output. Second, in February 2018, the hospital achieved the status of IRCCS (*Istituto di Ricovero e Cura a Carattere Scientifico*, or Scientific Institute for Research and Healthcare). IRCCS is a special designation conferred by the Italian Ministry of Health to institutes that excel in biomedical and healthcare service-related research while also delivering high-quality healthcare. The designation process is rigorous and explicated by the law: the hospital had to submit documentation in early 2017 to the regional government to receive approval; once documentation was approved in August 2017, it underwent expert evaluations and on-site visits by the Ministry late in 2017. Upon recognition in 2018, the hospital was officially accredited as an IRCCS in two specialty research areas (Personalized Medicine and Innovative Biotechnologies). Even though IRCCS are an Italian particularity, there are several international comparable institutions that may be associated in features to IRCCS, as they too are nationally accredited to perform research. Institutes like IHU (*Instituts hospitalo-universitaires*, France), *Universitätskliniken* (Germany), AMC (*Academic Medical Centers*, U.S.), IIS (*Institutos de Investigaciòn Sanitaria*, Spain) are alike to IRCCSs, albeit regulated to a lesser extent.

As an IRCCS, the hospital under matter gains access to dedicated public research funds – notably an annual "Current Research" fund available exclusively to IRCCS institutions, and preferential direct channels to compete for "Targeted Research" grants (which, otherwise, are granted to other hospitals only by means of being appointed by their Regions of reference, upon being awarded the funds themselves in public tenders). The IRCCS status also carries obligations: the hospital must maintain high research standards, and it is subject to re-evaluation by experts every three years. Concurrent with IRCCS recognition, the hospital management compiled an official list of researchers (the IRCCS research staff "perimeter") who would be eligible to use the newly available public research funds (updated every year). This IRCCS staff list consisted mostly of academically affiliated physicians (university-employed

doctors working at the hospital), but it was also made by some non-academic hospital physicians who were active in research. Importantly, all academic physicians, whether on the IRCCS list or not, were not part of the MBO bonus program (as noted above), and were appointed by the university. Although such appointment was of course based on preemptive characteristics, it was not a matter of self-selection. Meanwhile, non-academic physicians on the IRCCS list were still eligible for the MBO bonus (because they were hospital-employed clinicians). In other words, there was a subset of "double-treated" individuals: a few non-academic doctors who qualified for the IRCCS research staff list (thus benefiting from the IRCCS public funding) while also being eligible for the MBO monetary incentive. We summarize the groups and timing of these interventions below. In 2017 (the introduction of MBO), the relevant physician groups can be categorized as follows: 0 = those with no MBO and not included in the IRCCS perimeter (pure control group), 1 = those who eventually only benefit from IRCCS funding (academic physicians not eligible for MBO), 2 = those who receive the MBO incentive only (non-academic physicians not included in IRCCS list), and 3 = those who receive both MBO and later IRCCS (non-academic physicians who were included in the IRCCS research staff list). In 2018 (the IRCCS recognition), we can instead categorize individuals as: 0 = neither IRCCS nor MBO (pure control), 1 = MBO-only, 2 = IRCCS-only, and 3 = both IRCCS and MBO. In the empirical analysis we leverage such categorizations to define treatment and control groups for each policy.

MBO-levels of treatment		IRCCS levels of treatment	
	Description		Description
Policy	<i>Management-By-Objective (2017)</i>	Policy	<i>IRCCS recognition (2018)</i>
0	Full control (neither MBO nor IRCCS)	0	Full control (neither MBO nor IRCCS)
1	IRCCS-treated only	1	MBO-treated only
2	MBO-treated only	2	IRCCS-treated only
3	MBO+IRCCS double treatment	3	MBO+IRCCS double treatment

Table 1: Description of the two sets of treatment tiers.

### 3 Data

We construct a panel dataset of the professionals working in the hospitals, joint with their research output spanning 2012 to 2022. The core personnel data come from the hospital's administrative records, which include all individuals employed in a professional capacity (physicians, researchers, and other healthcare staff) from 2017 onward. From that, we identify our population of interest: physicians and medical

researchers continuously employed at the hospital between 2017 and 2022. For each person, the HR data provide demographic information (age, gender, place of birth), employment details (job title or role, department and unit, contract type, and hiring date), and indicate whether the individual has an academic affiliation (university-employed physician) or is a hospital-only employed clinician. We exclude medical interns, residents, PhD students, and external collaborators, as consistent data on these categories were not available. Because the administrative records begin in 2017, we supplemented them by benchmarking the administrative information on the starting dates with hire dates and career details from online CVs and institutional websites, enabling us to infer each physician's presence at the hospital in earlier years. Using the hire date information, we retrospectively extend the panel back to 2012 for all individuals who were active within the hospital between 2012 and 2017. We resort to this choice as we are not being able to identify workers who were active before 2017 and left the hospital before that year, hence not allowing us to have any information about their career. Thus, in order not to build an unbalanced panel with attrition focused only on the post-treatment part of the sample, we keep only the units who are balanced for the whole longitudinal dataset. The resulting balanced panel covers 579 physicians and researchers, each observed annually from 2012 through 2022, yielding 6,369 person-year observations. This balanced structure ensures that we can track research output for a fixed cohort over time, including several years before and after two policy changes. We create indicators for each person's group status (academic vs. non-academic, IRCCS list member vs. not, etc.) which remain fixed for analysis, reflecting their initial category (e.g., if a physician became an academic or left during the period, we handle such cases in robustness checks by excluding switchers). To measure research output, we gathered data on scientific publications from Clarivate's Web of Science database. We retrieved all publications (articles, reviews, etc.) from 2012 to 2022 for which at least one author was affiliated with the hospital (using standardized affiliation names). We accurately crafted such process by ensuring the recovered authors were the ones affiliated to the hospital, by matching names and surnames, fields of research, and by double checking we were not dealing with homonyms. This removes the portion of sampling errors that other studies, like Jacob and Lefgren, 2011, have to account for due to their matching based only on researchers' surnames. Each publication record provides the publication year, journal, number of co-authors, and citations (as of the time of data collection), and some other bibliometric details. We then matched authorship on these papers to the individuals in our hospital staff panel by name. From this, we computed the number of publications per person per year, which is our main outcome variable. In addition, we recorded the total citations each person's publications

received (to evaluate research impact).

	Mean	SD	Min	Max
Age	51.184	7.859	31	70
Publications	2.187	4.579	0	102
Times cited (WoS)	57.867	204.886	0	7037
Times cited (All outlets)	61.494	218.530	0	7509
<i>Gender</i>				
Female	0.418	0.493	0	1
Male	0.582	0.493	0	1
<i>Role</i>				
Medical Director (hospital only)	0.409	0.492	0	1
Healthcare Professions' Manager	0.019	0.135	0	1
Sanitary Director	0.050	0.218	0	1
Faculty Member with Clinical Functions	0.523	0.500	0	1
<i>Statistics per treatment group</i>				
MBO-only treated publications	0.452	1.039	0	12
IRCCS-only treated publications	4.978	6.919	0	102
Double-treated publications	2.258	2.776	0	20
Pure control publications	0.848	1.683	0	23
MBO-only treated citations	8.127	32.219	0	1009
IRCCS-only treated citations	150.415	359.049	0	7509
Double-treated citations	66.689	145.036	0	2102
Pure control citations	15.540	37.684	0	485
				<i>Total</i>
MBO-only treated units				177
IRCCS-only treated units				194
Double-treated units				66
Pure control units				162
Total units of the balanced panel				579
Observations				6369

Table 2: Descriptive Statistics of the panel of the balanced units, 2012–2022

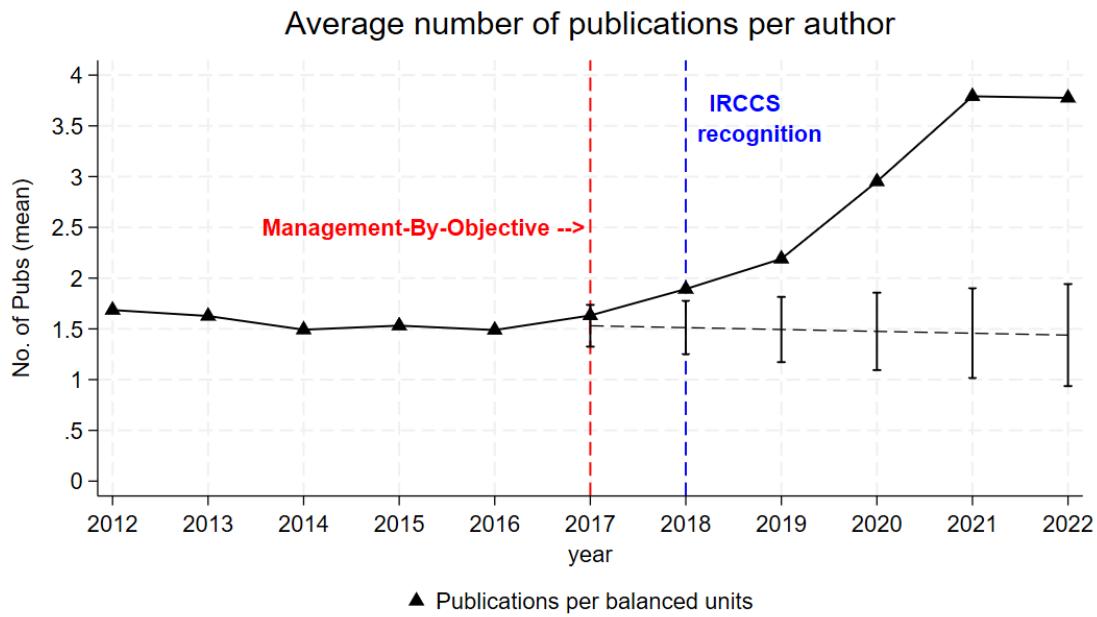
[Table 1](#) reports the summary statistics for the selected sample. About 52% of the 579 researchers in the panel are academic physicians (as in, they hold a university faculty appointment while also performing clinical activity), while the remainder are non-academic hospital physicians or other research/healthcare staff. On average, academic physicians in the IRCCS perimeter have higher scientific output than non-academics, reflecting different incentive structure and attitudes towards research. The average number of publications per person per year is around 5 versus less than 0.5 publication for non-IRCCS, non-academics. Their joint subset is instead in the middle way with respect to prolificacy, reporting almost 2.3 average yearly publications. The pure control group, which includes academic physicians and other professionals non included in either policy, presents a higher mean annual publica-

tion number (0.85). If we look at the times the works published in a given year were reportedly cited in October 2024, the period when data were collected, the statistics follow, intuitively, the same pattern. The variability of the reported variables is quite substantial in all groups. With respect to the size of the groups, the individuals ever being officially included into the IRCCS research staff list starting from 2018<sup>1</sup>, and not belonging to the non-academic MD category, are 194 individuals (approximately 33.5% of the sample), of whom a less than half but still relevant subset (66 individuals) are non-academic physicians (thus eligible for MBO as well). The non-academic non-IRCCS physicians are instead 177 (30.6% of the sample), while the pure control group is made by 162 individuals (28%). Graphs (a) and (b) in [Figure 1](#) plots the average number of publications per physician counted at the end of each year, from 2012 to 2022. We observe a relatively flat trend in the years up until 2017, followed by a noticeable increase starting around 2018. This timing aligns with the introduction the IRCCS recognition, suggesting, at a first glance, a possible aggregate effect of these policies on the overall research productivity at the hospital, mediating by the supposed lagged effect that would involve the execution of a scientific work before achieving publication. It must be noted that the time-lapse between submission to acceptance/publication are usually quite short in the medical field compared to other fields, like the economic one. Whereas to publish in the best ranked economic journals the lag has been longer than 2 years for decades, even reaching more than a 40 months-lag from submission to publications for papers ranked in the upper quantiles of the distribution ([Yohe, 1980](#), [Hadavand et al., 2024](#)), such time-lapse is on average 8 months or in the medical field, lowering to even a coupe-of-months time lengths for systematic reviews or literature reports ([T.-A. Chen et al., 2024](#)). For instance, a top-journal in the medical field like *JAMA Network Open*, states in its official address to authors that the median time from submission to publication is 94 days, i.e. 3 months ([JAMA, 2025](#)).

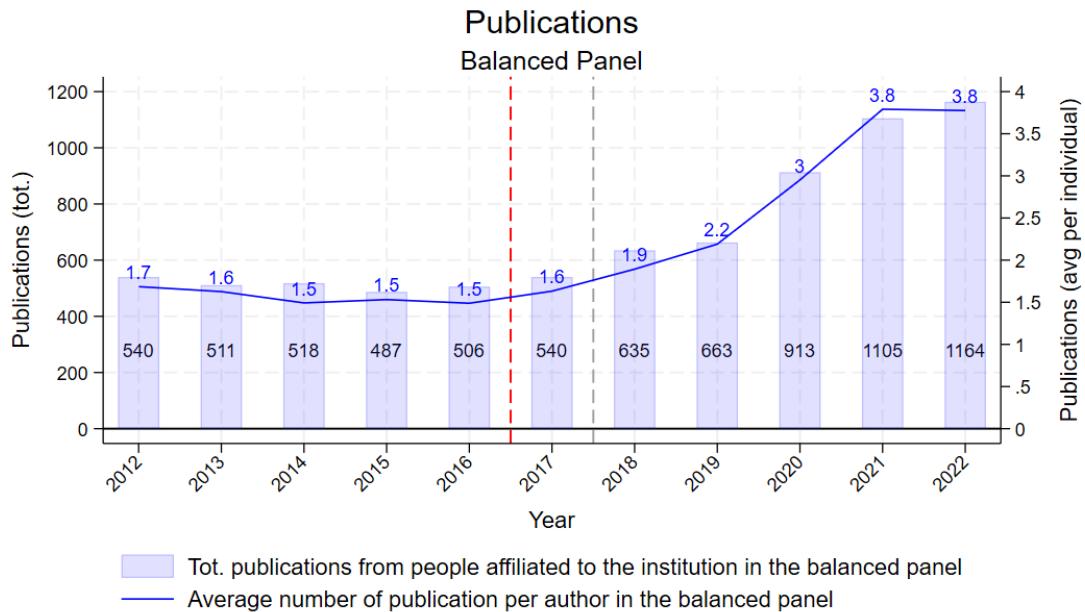
While other secular trends could also contribute to the increase (e.g., growing institutional emphasis on research), the positive, diverging pattern beginning in 2018 is quite suggestive. In the analysis below, we exploit the individual-level variation in exposure to the policies to identify their causal impact.

---

<sup>1</sup>As discussed later in the paper, the IRCCS list is actually updated every year. However, the updating is just a formal recognition of individuals already involved in the IRCCS' activities due to their productivity, even without having official access to the funds until they are acknowledged by the Ministry as formally taking part to the research group. In reality, it would be more accurate to consider such group of individuals as featured by a time-varying size pattern. We take this into account in the staggered adoption robustness check.



a) Average publications per researcher per year (2012–2022).



b) Total number of publications attributed to the employees of the institution for the balanced sample (2012–2022).

Figure 1: productivity: average trend and total output (2012–2022).

In defining treatment and control groups for the analysis, we use the categorizations described in Section 2. For the MBO policy, the "treated" group consists of non-academic physicians (those subject to the MBO bonus) and the control group consists of those not eligible (academic physicians and other staff). For the IRCCS policy, the treated group consists of those included in the IRCCS researcher list (primarily academic physicians) and the control group is those not on the list. Additionally, we will examine subgroups such as the double-treated individuals. The next sections outline our empirical strategies in detail.

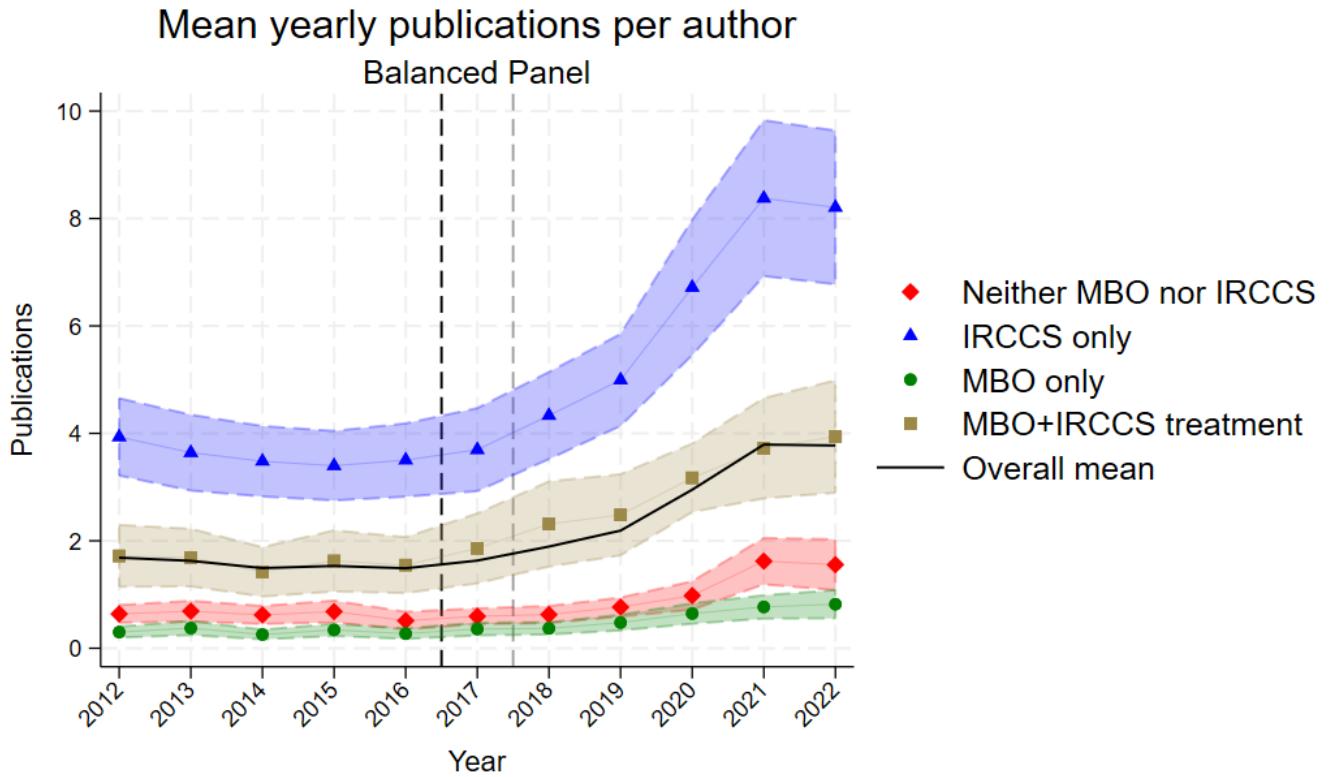


Figure 2: Average yearly publications across different groups of healthcare professionals, defined according their treatment status after 2017 and 2018.

A sketch of the average evolution of publication patterns across groups is presented in Figure 2. A first visual inspection seems to suggest us that the bulk in the jump of yearly publications is led by researchers affected by the IRCCS recognition, with apparently no impact brought about by the performance-based policy, as the only MBO-subject individuals who evidently display a trigger in productivity overall, are the ones who are double-treated.

## 4 MBO Policy

The first policy we analyze is the performance-based monetary incentive (MBO) implemented in 2017. The goal is to estimate the causal effect of this incentive on research output (publications) of the targeted physicians. We employ a twofold TWFE difference-in-differences (DiD) framework. First, we perform a set of estimates comparing the change in publications across different groups of individuals within the same model specification, before and after 2017 and according to the increasing “intensity” of the policy treatment, where the lowest indicators are 0 and 1 for the control groups (pure control, and never MBO-treated researchers who become part of the IRCCS perimeter in 2018), and 2 and 3 for the treated ones (MBO-treated only, double treated after 2018). Then, we repeat the DiD analysis with a standard binary treatment indicator, by subsetting the dataset each time in order to compare different groups. We explain this identification strategy better in the next subsection.

### 4.1 Empirical Strategy

Our first baseline specification for the MBO policy is a standard static specification for a two-way fixed effects DiD model, estimated via OLS:

$$Y_{it} = \alpha + \sum_{j=1}^3 \beta_j Post2017_t \times \mathbb{I}(MBO_i = j) + \delta_i + \tau_t + \varepsilon_{it}, \quad (1)$$

where  $Y_{it}$  is the research output of individual  $i$  in year  $t$  (measured as the number of publications),  $Post2017_t$  is an indicator for the post-intervention period (2017 and later),  $MBO_i$  is an indicator for different degree of the MBO-policy treatment. The range of  $MBO_i$  is expressed by  $MBO_i = \{0, 1, 2, 3\}$ , where 0 := pure control; 1 := MBO-control units who receive IRCCS recognition in 2018; 2 := MBO-treated units only; 3 := units treated with MBO in 2017 and with IRCCS recognition in 2018. Such indicator allows us to compare the effect on different categories of units with respect to the baseline category, as in those who do not receive any kind of funding or incentive ( $j = 0$ ).  $\delta_i$  are individual fixed effects,  $\tau_t$  are year fixed effects. The coefficients  $\beta_j$ , with  $j = \{1, 2, 3\}$ , on the interaction terms  $Post2017_t \times \mathbb{I}(MBO_i = j)$  capture the DiD estimates of the MBO incentive’s intensity of the effect on annual publications. By including  $\delta_i$ , we control for the time-invariant differences between physicians, while  $\tau_t$  for common shocks in each period. Standard errors are clustered at the individual level. The key identification assumption is that, absent the MBO policy, the research output of the treated and

control groups would have followed parallel trends. While such assumption is not directly testable, we present event-study evidence to assess whether the different groups exhibited similar trends prior to 2017<sup>2</sup>. However, as many academic doctors were, around the same time, exposed to the IRCCS intervention in 2018, we recover that in our baseline DiD, their post-2017 effect might partly reflect anticipation or effects of IRCCS. We address this potential overlap in two ways: first, by reporting dynamic estimates for the event-study, which also displays the immediate effect of the treatment before the IRCCS implementation of 2018. Second, by explicitly accounting for IRCCS in a triple-difference model (presented in Section 4.3). In addition, we implement a more granular approach that exploits the multiple treatment groups defined earlier. While in the Equation (1), the MBO eligibility was dealt with a four-level categorical variable, in a second set of OLS estimations, we perform various binary DiD comparisons between specific groups to directly assess the magnitude of treatment effects. Such estimations are structured as follows:

$$Y_{it} = \alpha + \beta Post2017_t \times MBO_i + \delta_i + \tau_t + \varepsilon_{it}, \quad (2)$$

Equation 2 mirrors Equation 1, with a slight variation; as in, the interaction of interest involves no longer four-level categorical indicator, but the binary dummy  $MBO_i$ , which assumes the value 0 or 1 depending on the appurtenance of unit  $i$  to the designed treatment (1) or control (0) group. Specifically, the estimation is performed several times with different compared units, by re-classifying the treatment and control groups in five different ways: (1)  $MBO_0$ : double-treated (MBO+IRCCS) vs. MBO-only; (2)  $MBO_1$ : double-treated vs. IRCCS-only; (3)  $MBO_2$ : double-treated vs. neither (full control); (4)  $MBO_3$ : MBO-only vs. IRCCS-only; and (5)  $MBO_4$ : MBO-only vs. full control. For each comparison, we restrict the sample to only the units involved in that very comparison. This provides insight into, for example, whether the MBO effect was larger than the IRCCS effect by comparing the different categories. We interpret these results alongside the main specification.

To corroborate our causal claim, as anticipated, we estimate event-study models which allow for the treatment effect to vary in each year relative to the policy change, by substituting the  $Post2017 \times MBO_i$  with a series of interactions of  $MBO_i$  with time-dummies for each year before and after 2017 (by excluding the one immediately before the treatment year to avoid multicollinearity issues). We do this by estimating the following equation:

---

<sup>2</sup>To avoid ambiguity in the interpretation of the results, in the Event-Study estimates we separate the groups between individuals treated with MBO-policy, irrespective of their potential double-treated status, and a control group, made by all individuals who are either academic physicians or non-physicians at all.

$$Y_{it} = \alpha + \sum_{h=2}^H \beta_H (Lag\ h)_{it} + \sum_{g=1}^G \beta_g (Lead\ g)_{it} + \delta_i + \tau_t + \varepsilon_{it} \quad (3)$$

The reference period being  $t := 2017$ .  $(Lag\ h)_{it}$  are the pre-treatment period dummies associated to the treated units  $(Lag\ h)_{it} = \mathbb{1}[t = 2017 - h] \text{ for } h \in \{2, \dots, H\}$ . The post-period dummies interacted with the treatment binary variable are instead indicated by  $(Lead\ g)_{it}$   $(Lead\ g)_{it} = \mathbb{1}[t = 2017 + g] \text{ for } g \in \{1, \dots, K\}$ . This yields coefficients for the differential trend between treated and control groups in each year, which we plot to visually inspect pre-trends and dynamic effects. We present these event-study results as figures accompanying the main tables. Note that, while in the comparison across multiple groups we follow the same subsetting approach as the one mentioned for mbo'did'2, our baseline event-study does not interact the period indicators with the multi-level intensity variable based on four categories, as its coefficients in the test would yield results hard to interpret. When we perform the baseline dynamic estimate on the whole sample, all MBO-treated units are considered as treated, while all non-MBO treated units are accounted as control, the distinction made regardless of the units' IRCCS status.

## 4.2 Results

[Table 2](#) presents the difference-in-differences estimates of the effect of the MBO monetary incentive on publications. The table includes different specifications, according to the different strategies described above (Cols. (1-6)). Overall, we observe a positive and statistically significant effect of the MBO policy only when the recipients are also part of the IRCCS perimeter. (double-treated, i.e.  $DT$ ). The baseline DiD intensity estimate (comparing non-academic to academic physicians, in Col. (1)) indicates that after 2017, IRCCS-related academic physicians (*IRCCS*) increased their annual publications by approximately more than 100% of the pre-treatment mean of the groups affected by any kind of policy, compared to the pure control (*PC*). On the other side, MBO-eligible MDs who were never part of the IRCCS perimeter (*MBO*) did not experience any significant effect after the policy, compared to non-eligible, non-IRCCS individuals (*PC*). MBO-physicians who were also involved in the IRCCS policy increase their publications by more than 50% of the pre-treatment average, compared to *PC*. When we differentiate among the four categories, we find that the post-2017 jump in the publications of the *DT* group is positive and significant only with respect to the *MBO* group (+78%) and the pure control (+72%), turning negative when compared to the IRCCS researchers (-73%).

	2012-2022					
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications
Level 1 (IRCCS)	2.08984 ***					
	[0.28890]					
Level 2 (MBO)	-0.16094					
	[0.12954]					
Level 3 (DT)	1.02864 ***					
	[0.23702]					
DT (vs. IRCCS)		-1.17824 ***				
		[0.34327]				
DT (vs. MBO)			1.25519 ***			
			[0.23172]			
DT (vs. PC)				1.15212 ***		
				[0.23798]		
MBO (vs. IRCCS)					-2.66348 ***	
					[0.31424]	
MBO (vs. PC)						-0.21831 [0.13316]
Observations	6,369	2,695	2,600	2,418	3,944	3,674
R-squared	0.77158	0.76431	0.59666	0.58012	0.79368	0.43801
Individual FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	Full	DT and IRCCS	DT and MBO	DT and PC	MBO and IRCCS	MBO and PC
Mean	1.916	1.600	1.600	1.600	0.309	0.309

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3: Impact of MBO Incentive on Annual Publications (Difference-in-Differences).

The *DT* group seems, quite counterintuitively, performing better compared to the *MBO* units than the control although, by looking at the clustered Standard Errors, the difference does not appear to be statistically significant. By contrast, *MBO* substantially underperform compared to the *IRCCS*-group (Column 5, with a striking  $-880\%$ ), while displaying no relevant differential in comparison to the pure controls.

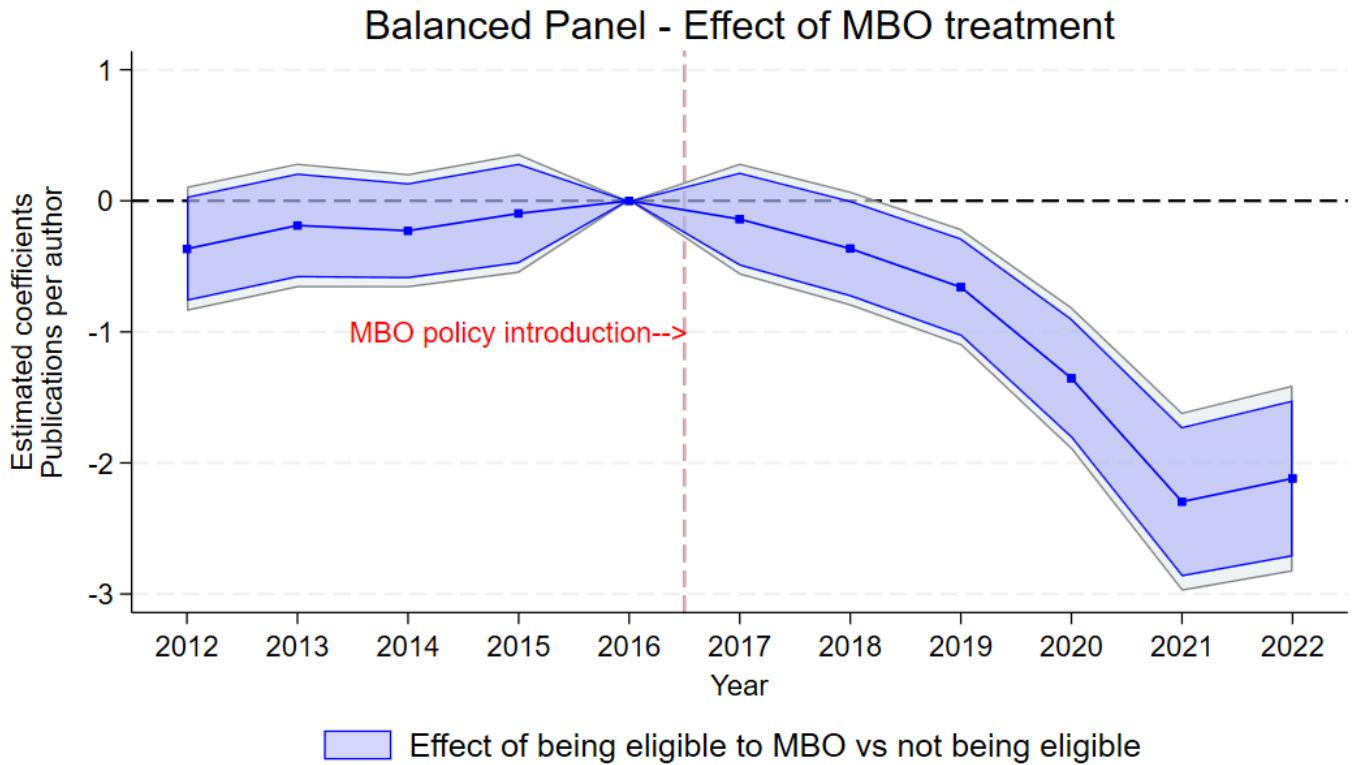
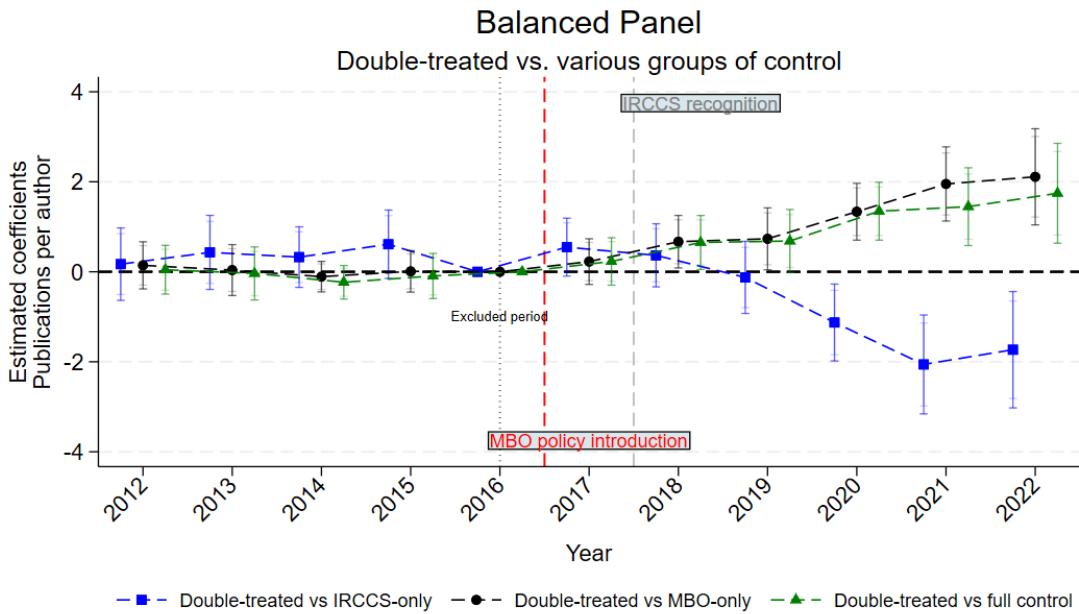
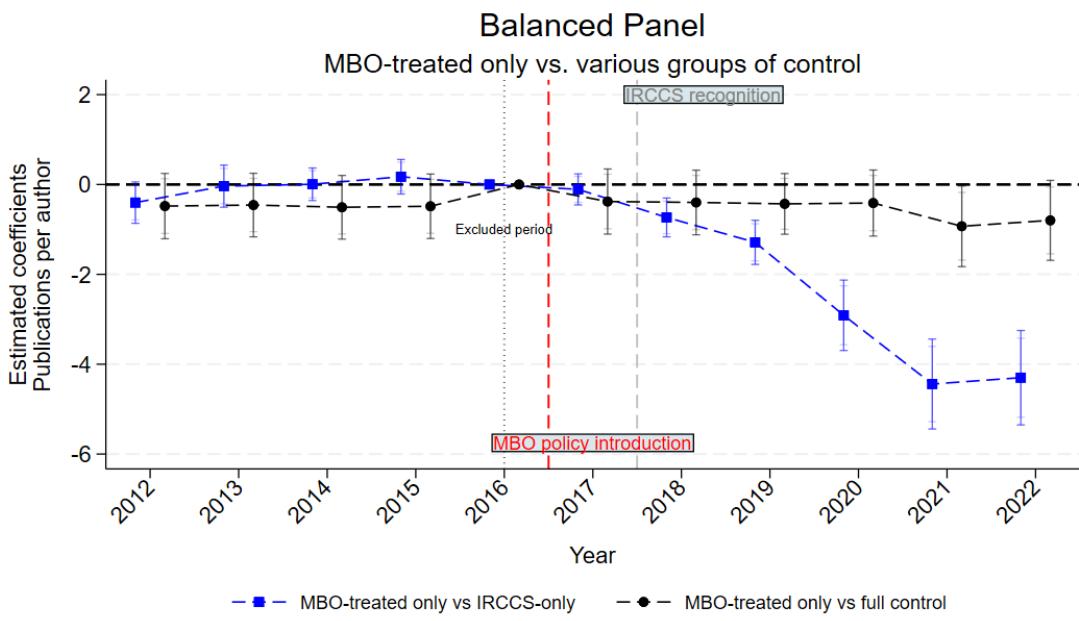


Figure 3: Event-study of MBO policy effect on publications. The persistent decline is sensibly driven by the confounding effect of the subsequent IRCCS recognition.

Coming to the dynamic setting, Figure 3 depicts the event-study estimates for the MBO policy's effect on publications, by subdividing the groups just according to the inclusion or not into the MBO-policy scheme. The figure plots the yearly coefficients for the interaction of MBO eligibility with each time period (taking 2016 as the reference year). The pre-2017 coefficients are close to zero and statistically null, indicating no divergent trend between the two groups, supporting the parallel trends assumption. The coefficients stay non significant in 2017, start decreasing over time only after 2018, showing that the publication gap between MBO and non-MBO physicians widened negatively and significantly in the post-policy period.



a) Event-study of MBO policy effect on publications of the double-treated units compared to the various control groups. The persistent decline is sensibly driven by the confounding effect of the subsequent IRCCS recognition.



b) Event-study of MBO policy effect on publications of the MBO-only treated units compared to the various control groups. The persistent decline is sensibly driven by the confounding effect of the subsequent IRCCS recognition.

Figure 4: Event-study of MBO policy effect on publications for comparisons across different groups.

The trough is reached in the immediate post-COVID period, as in 2021 only the treated group was publishing on average about  $-2.2$  fewer papers per person than they would have without the incentive in 2016. The event-study suggests that the MBO was substantially ineffective and the persisting decline displayed by the graph is sensibly driven by the compound impact caused by IRCCS recognition. We address the sensitivity to the boost in research activity triggered by the pandemic crisis in the Robustness checks. Given that in the latter dynamic specification the identification was not clear-cut, as some MDs were also involved by the compound scope of the IRCCS recognition starting from 2018, while the outcome of the control group was evidently shifted up heterogeneously by the second policy, we estimated a set of different dynamic models (referencing to Equation 3) which compare first the double-treated units (Graph a) of Figure 4), and then the *MBO*-only individuals (Graph b) of Figure 4), to the various control groups. While both graphs show the validation of the common trend assumption for all the groups of comparison, the evidence that the MBO policy affects individuals only joint with IRCCS funding is corroborated even more, with the negative and significant gap between double-treated individuals and MBO-treated only ones widening significantly over time, starting from 2018 only (the IRCCS recognition year), and stabilizing one year after the pandemic outbreak. For the double-treated MDs as well, the effect starts kicking in 2018 only, thus concomitantly with the IRCCS recognition.

### 4.3 Triple-Difference Analysis

Through the results above we attempted at disentangling the effect of the MBO policy under reasonable assumptions and across different groups. However, the almost simultaneous recognition of the hospital as IRCCS in 2018 raises concerns of compound treatment: hence, we also try to address the issue of interaction between the two policies. We therefore estimate a differences-in-differences-in-differences (DiDiD) model which embeds both policies.

$$Y_{it} = \alpha + \beta_1 Post2017_t + \beta_2 IRCCS_i + \beta_3 (Post2017_t \times MBO_i) + \\ + \beta_4 (Post2017_t \times IRCCS_i) + \beta_5 (Post2017_t \times MBO_i \times IRCCS_i) + X'_{it}\gamma + \varepsilon_{it}, \quad (4)$$

In Equation 4  $IRCCS_i$  is binary indicator assuming value 1 if  $i$  would ever be part of the IRCCS staff at a certain point after 2018, to which the associated coefficient to is  $\beta_2$ . By contrast,  $\beta_3$  captures the effect of the MBO policy on those neither in the IRCCS staff (double-treated include) nor the pure control, as it interacts the  $Post2017_t$  time dummy with the binary for being subject to the MBO.  $\beta_4$  is the coefficient linked to the effect of the IRCCS recognition on those not subject to MBO nor part of

the control group, starting from when the MBO is implemented (as the dummy  $IRCCS_i$  is interacted with  $Post2017_t$ ). Eventually, the triple interaction coefficient  $\beta_5$  bears major relevance: it recovers the additional effect for individuals who are considered as double-treated (MBO-eligible and included in the IRCCS perimeter after); it hence helps us to disentangle the compound impact of monetary incentives and public funding on the affected individuals. Note that while  $Post2017_t$  assumes unitary value in any year after 2017 included, the IRCCS change kicks in only in 2018. Not to avoid the rising of collinearity issues, we avoid including two way fixed effects, which we account for via a set of time-varying controls:  $X_{it}$  (age, age squared and dummies for gender, age cohort, birth province, and department) aims at accounting for differential, individual trends correlated with observables. The triple-difference strategy relies on the assumption that, absent the policies, the difference in outcomes between, say, non-academic and academic doctors would have evolved similarly over time for IRCCS-listed and non-listed groups. We also estimate further specifications, including some of the interactions with or without controls, including two way fixed effects when covariates are ruled out. Eventually, we integrate this set of estimates with an additional “complete” and covariate-less triple DiD, where TWFE are included at the expense of the terms  $Post2017$  and  $IRCCS_i$ , necessarily omitted to avoid collinearity. This approach could appear a bit complex; however, provided with the parallel trend checks we performed in the previous checks, we are somewhat able to assume that no other unobserved shocks differentially affected the subgroups’ trends.

Findings are reported in Table 4, with the most complete specifications in Cols. (7) and (8). Column (7) excludes time-varying controls, the  $Post2017_t$  and the  $IRCCS_i$  terms, but includes units and year fixed effects, while Column (8) uses individual controls but omits the fixed effects due to including the treatment dummies mentioned above. First, we observe how, in Col. (8), there is no significant effect after the policy implementation within the pure control group. Then, the second row of the same column show that the individuals affiliated to the IRCCS (both the IRCCS-researchers only and the compounded treated units) have, on average, higher publications than the other two groups (pure control and MBO-only eligible doctors), across the whole sample (thus even before the IRCCS introduction). However, it must be noted that this framework does not account at all for unobservable heterogeneities, which may not be tackled effectively by conditioning on our set of time-varying regressors. Moving on, the coefficient on  $Post2017 \times MBO$  results negative and not statistically differing from 0 in both models, confirming our previous hypothesis that MBO-eligible physicians did not increase their publication activity after the performance-based reform, absent a following affiliation with the IRCCS staff. On the other hand, both

columns display that the estimate of  $Post2017 \times IRCCS$  is large, strongly significant and comparable across both specifications. This validates the intuition that the IRCCS recognition had a substantial and positive effect on publication activity for the IRCCS-staff, which in such coefficient includes also the double-treated physicians.

	2012-2022; SEs clustered at individual level							
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications	(7) Publications	(8) Publications
Post2017								-0.26760
IRCCS								[0.25256] 2.96665 ***
Post2017 × MBO	-0.88737 *** [0.28646]			-0.98037 *** [0.17424]			-0.16086 [0.12953]	[0.37963] [0.32706] 2.10048 1.79566 ***
Post2017 × IRCCS								
MBO × IRCCS		0.61781 [0.39551]			-0.96832 * [0.57456]			[0.30140] 0.23336 -1.81489 ***
Post2017 × MBO × IRCCS			0.82784 ** [0.37408]			0.24562 [0.22767]	-0.92516 ** [0.67519]	[0.44998] -0.75093 [0.38099] [0.46327]
Observations	6,228	6,228	6,228	6,369	6,369	6,369	6,369	6,228
R-squared	0.34216	0.33938	0.33964	0.76263	0.75998	0.75987	0.77159	0.41766
Individual FE	NO	NO	NO	YES	YES	YES	YES	NO
Year FE	NO	NO	NO	YES	YES	YES	YES	NO
Covariates	YES	YES	YES	NO	NO	NO	NO	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced
Mean	2.187	2.187	2.187	2.187	2.187	2.187	2.187	2.187

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 4: Triple Diff-in-diff, comparing different groups.

We observe then how the interaction, time-invariant effect of being a non-academic Medical Doctor (thus eligible for the MBO bonus) and a IRCCS-staff member at the same time, has apparently a negative effect on scientific productivity with respect to the other groups (pure control and especially IRCCS-only researchers). However, such effect is huge and significant only when accounting for time-varying controls, which however are not able to capture time-invariant unobservable characteristics, nor average time trends. When including unit and time Fixed Effects however, the effect reverts its sign, although getting non-significant, possibly suggesting that the coefficient gets its relevance mostly from unobservable heterogeneities across individuals. Eventually, the triple interaction term  $Post2017 \times MBO \times IRCCS$ , which captures the differential effect for individuals simultaneously exposed to both policies (double-

treated), is negative in both complete estimations. While they are somehow comparable across the two specifications, the term appears significant only in the TWFE model, therefore highlighting the potential presence of either colliding bias in the included covariates, or the ineffectiveness of capturing time-varying heterogeneities of the fixed effect. However, given the absence of any sort of controls for general time trend in the last model, we are keen on tending towards the former hypothesis. Although the coefficient does not reach conventional significance levels, its negative sign suggests that the effect of IRCCS on research output may be attenuated when combined with MBO, potentially due to competing incentives or time constraints. To sum up, our findings confirm that a positive shift in research productivity is strongly linked to the accreditation of the hospital to receive IRCCS funding, while highlighting the weakness of implementing management-by-objective scheme alone. The absence of an additional positive interaction effect in the double-treated group suggests that the joint implementation of MBO and IRCCS does not yield additive benefits, but this shall be taken into account carefully, as the effect is with respect to a control group composed by all the groups included in the framework (and not only a selected one). In any case, it shows that what matters the most is the funding policy.

The most important thing to underline after presenting such estimates is noting that, in the DiDiD model, the post-2017 indicator picks up any overall time shock after 2017 which is common to all groups hence including, certainly, the *Post2018* IRCCS recognition). While the inclusion of year fixed effects  $\tau_t$  rules out accounts common shocks in Col. (7), it does not the same in the final specification. On top of that, even accounting for that, this does not allow to separate the effect heterogenously across individuals, especially due to the fact that the IRCCS recognition occurs one year immediately after the MBO policy. However, to further validate the estimates from the latter framework, we perform a robustness check by subsetting the sample only between 2016 and 2018, in order to rule out the bulk of the IRCCS' recognition impact, which only confound the estimates to the extent through which they are triggered immediately. Table A1 in the Appendix shows the results of such analysis, where we can observe how such results are way lower in magnitude and some of them gets non-significant, although the directions of the effects stay the same.

## 5 IRCCS Funding Analysis

Given the almost contemporaneous overlapping of the two policies, since the second one is introduced just one year after the first one, we cannot really disentangle the effectiveness of either discontinuity

without focusing on the second one in details as well. So we now turn to the IRCCS recognition, which triggered an allocation of greater funds for research, and which began in 2018. Such event mainly affected the hospital’s academic-affiliated researchers, allowing them to receive dedicated public resources and, in addition, by enhancing the prestige of the institution under the scientific point of view. We use a difference-in-differences approach to estimate the effect on academic physicians’ research output, in doing this exploiting a parallel identification strategy with respect to the assessment of the performance-based bonus.

## 5.1 Empirical Strategy

Mirroring what we did for the MBO-framework, to identify the IRCCS recognition’s effect, We estimate a baseline specification for a standard two-way fixed effects model, via OLS estimators:

$$Y_{it} = \alpha + \sum_{j=1}^3 \theta_j Post2018_t \times \mathbb{I}(IRCCS_i = j) + \delta_i + \tau_t + \varepsilon_{it}, \quad (5)$$

where  $Post2018_t$  equals 1 if  $t \geq 2018$  and 0 otherwise,  $IRCCS_i$  indicates if individual  $i$  is in the IRCCS perimeter (i.e., directly eligible for IRCCS funding), and the other included terms are the same as previously defined. In such regards, we need to make a clarification: while the hospital updates the IRCCS perimeter at the end of every year, we still perform a standard  $2 \times 2$  TWFE DiD estimation, by considering as treated those who are *ever*, sooner or later, included in the IRCCS perimeter. This is done in order to allow a more clear-cut identification of the ATT. As a matter of fact, even though solely IRCCS-only individuals should be formally granted access to IRCCS funding, a structured research group, within the personnel working at the institution, takes shape immediately after the recognition of the scientific excellence standards of the hospital under study. Such crowd can be accounted as the set of physicians who actively take part to most research activities, and collaborate in order to obtain the privileged funding which is granted to the hospital due to its IRCCS status. In any case, the perimeter is not re-updated yearly from scratch, but there exists a consistent core of researchers which stay constant over time, which is integrated yearly by those fellows who are “*de facto*” IRCCS researchers as they collaborate to the same funded projects and thus have access to the same resources, even though their status is officially recognised only after few periods; they hence basically self-select into the treatment after the IRCCS recognition. This notwithstanding, self-selection is not even the proper expression, as the inclusion into the IRCCS perimeter is always established by the institution (top-down). In addition

to that, reflecting the same settings of the MBO policy, the treatment as assumed as an absorbing one. However, and differently from the MBO-policy which is absorbing by construction, the inclusion in the IRCCS perimeter is not constant until the end of the time-span, as some individuals apparently “switch off” the treatment since are excluded from the list by the management at some point. There is no reason not to maintain the absorbing treatment assumption nonetheless, because while papers are written in a relatively short time frame, the exploitation of accessed funds can be actually prolonged in time, so researchers who started projects (or were involved in some) would feasibly keep working to such projects by using the same means even after their IRCCS status has expired. To check for these issues, albeit holding the absorbing assumption as always true, besides the standard test for parallel trends, we also perform additional estimates accounting for the staggered inclusion of some researchers to the core of the IRCCS perimeter, and controlling for the “switchers” across groups.

The coefficients  $\theta_j$  capture the average effect of the intensity level of the IRCCS recognition treatment, with the intensity depending on the degree of compoundedness of the policy, on the research output of the 3 treated groups with respect to the pure control, after 2018. In such setup we compare IRCCS-listed researchers (by subdividing such category into the sub-group of double-treated MDs and that of academics) to non-listed individuals (including both MDs, which is the first tier of treatment in this case, and the pure control group), before *and* after 2018. As before, the main identification assumption is that, without the IRCCS acknowledgment, the scientific productivity across groups would have followed parallel trends in the post- period. While TWFE account for overall temporal trends and time-invariant heterogeneities, we test the assumption through event-study estimations, as we did for the MBO policy. In addition, we integrate the analysis with the staggered adoption framework we described above, to assess the degree of self-selection into treatment. Since in such framework the issue of self-selection is major compared to the MBO-policy setting, we also provide the main event-study estimates (in both the main analysis and the staggered one) with additional validity checks which exploit the recent methodology of Rambachan and Roth, 2023. Their so-called *honestDiD* methodology allow for estimates with potential violations of the parallel trend assumption (which are plausible in case of selection), by computing confidence sets granting the robustness and significance of estimated dynamic coefficients in the period after the treatment *even with deviation from the common trends*, up to a certain point which is the threshold-level computed by the methodological algorithm itself; more on this later in the paper.

As for the MBO analysis, we implement two different strategies here. In 2018, the four categories identified categories (0, 1, 2, 3) are defined slightly differently (0 = control, 1 = MBO-only, 2 = IRCCS-

only, 3 = both, as reported already in Table 1). First, we estimate the effect of the intensity of such treatment. Then, we incorporate these estimates with a methodology to assess the differential effects across different groups. This second strategy entails the following comparisons: (1) double-treated vs. IRCCS-only; (2) double-treated vs. MBO-only; (3) double-treated vs. pure control; (4) IRCCS-only vs. MBO-only; (5) IRCCS-only vs. pure control; (6) IRCCS-only vs. double-treated (specular to (1)). Such comparisons are obtained by estimating a set of equations constructed as follows, by changing the comparison groups for every regression:

$$Y_{it} = \alpha + \theta Post2018_t \times IRCCS_i + \delta_i + \tau_t + \varepsilon_{it}, \quad (6)$$

Where  $\theta$  is different for each estimate and captures a different comparison, according to which group is the dummy  $IRCCS_i$  identifying each time (depending on the six comparisons mentioned before).

After this, we compute a set of dynamic coefficients for the IRCCS intervention by means of event-studies for the latter set of models (not for the intensity one, in order to avoid comparability ambiguities), interacting  $IRCCS_i$  with year dummies around 2018, according to the compared groups involved. The baseline comparison does not really reflect Equation 3. As a matter of fact, while in the baseline comparison of Equation 3 the dummy  $MBO_i$  takes unitary values if the individual is treated with the incentive policy, irrespective of their double-treatment status (so also IRCCS-individuals are included in the treatment), in (the one which follows) the approach is more “conservative”, as the binary variable takes value 1 only if  $i$  is included in the IRCCS perimetete *only*, without being subject to the MBO-policy as well (as in, is an academic doctor); the control group consists in the set of groups 3, 1, and 0 combined. Such equation enables the visualization of potential diverging pre-2018 patterns. Here is the estimated equation (via OLS):

$$Y_{it} = \alpha + \sum_{h=1 \wedge h \neq 2}^H \theta_h (Lag h)_{it} + \sum_{g=1}^G \theta_g (Lead g)_{it} + \delta_i + \tau_t + \varepsilon_{it} \quad (7)$$

Note that, while it is common practice to exclude the time immediately *before* the treatment to avoid multi-collinearity issues (as in,  $h = 1$ ), in the latter estimates we opt to exclude  $h = 2$  instead (which is  $t = 2016$ ), which is the same excluded period in the event-studies for the MBO-policy treament. We acted like this in order to allow greater comparability between the event-studies for the different policies.

## 5.2 Results

We report the DiD estimates from Equations 5 and 6 in Table 5. The estimates show a positive and significant impact, confirming the pattern we already observed previously, showing the importance of IRCCS recognition compared to the performance-based policy. Col. (1) report the coefficients for all three interactions  $Post2018_t \times \mathbb{I}(IRCCS_i = j)$  described in Equation 5. The other columns display instead the coefficients for the single interaction terms  $Post2018_t \times IRCCS_i$ , for different groups' comparisons as indicated in the table itself (DT vs. MBO, DT vs. IRCCS etc.). The second coefficient in Col. (1) implies that, on average, being listed as an IRCCS researcher-only (IRCCS) is linked with a growth of almost 2.5 additional publications per year (significant at 1% in the intensity-based model) compared to the pure control and accounting for the other levels. Such increase is quite outstanding, as it accounts for more than 125% of the average mean of individuals treated to some extent (non pure-control, as in, non PC) before the IRCCS recognition. The effect more than twofold in absolute terms when compared to the MBO impact for non-academics; however, the latter is still strong (a bit less than 1.2 publication per year, as in 59% of the pre-treatment mean for the non-PC). A negative, small and statistically non-distinguishable from zero effect is found for the MBO-only treated units. When breaking down the effect by category, results are consistent. The units treated compoundedly publish almost 1.26 more works compared to the MBO-treated only (+76%), 1.18 fewer works compared to the IRCCS-only researchers (-72%), and 1.15 more works compared to the pure control (+70%). Such findings are consistent with the stylized fact that IRCCS contributed substantially both academics and non-academic MBO-treated MDs who had the opportunity to be subject to both policies, although the researchers involved in scientific activities primarily are the ones who responded the most from the recognition, as they display higher publication rates even when compared to those who are selected into the IRCCS list and allowed to get the performance bonus. We observe this in the final columns of Table 5: in Col. (5), the effect of IRCCS recognition on IRCCS vs. MBO is almost +2.7 publications (+74%); in Col. (6), IRCCS publish almost 2.6 works more than the pure control (+71%). In Col. (7), we find the same coefficient retrieved in Col. (3) in module, but with reverted sign: as in, IRCCS publish 1.18 (+72%) more papers than the double-treated (DT).

	2012-2019; SEs clustered at individual level						
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications	(7) Publications
Level 1 (MBO)	-0.20198						
Level 2 (IRCCS)		[0.13028] 2.45771 ***					
Level 3 (DT)		[0.31442] 1.14568 ***					
DT (vs. MBO)			1.25519 ***				
DT (vs. IRCCS)				-1.17824 ***			
DT (vs. PC)					[0.34327] 1.15212 ***		
IRCCS (vs. MBO)						2.66348 ***	
IRCCS (vs. PC)							[0.31424] 2.55316 ***
IRCCS (vs. DT)							
Observations	6,369	2,600	2,695	2,418	3,944	3,762	2,695
R-squared	0.77615	0.59666	0.76431	0.58012	0.79368	0.78407	0.76431
Individual FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced
Mean	1.934	1.643	1.643	1.643	3.609	3.609	3.609

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 5: Impact of IRCCS recognition on Annual Publications (Difference-in-Differences).

The baseline dynamic specification for the IRCCS effect is represented in the event-study in Figure ???. As above mentioned, the definition of treated unit here is quite conservative, as we are only comparing academic doctors who receive IRCCS-funding only with the rest of the personnel, included those who receive the performance payment and have access to IRCCS-funding too. Prior to 2018, the publication raw numbers for the IRCCS-only group were relatively flat non-diverging with respect to the rest of the units. While their scientific output was way higher on average (since they were selected for being productive, see Table 2), we notice that, in terms of time-varying pattern, there is no significant difference from the slope of the publication output of non-treated ones until 2018, thus highlighting the relevant impact of the access to the IRCCS funding and the absence of evident pre-trends. In the graph, we observe a small and slightly significant jump in 2018, reasonably due to the, albeit small but existent, lag in the publishing process. The coefficient keeps increasing in 2019 and displays a greater-than-100% jump in 2020, which could be credited to both the passing of time which reduces the effect of the aforementioned lag, and to the boost in research activity triggered by the COVID outbreak. The effect seems to stabilize in 2022, after a further substantial growth in 2021. It appears that the IRCCS recognition's impact had already materialized few lags after its implementation, and researchers had possibly already optimized the resources in order to produce output in a timely fashion: so not only new projects have been ushered, but some of them have already been completed.

As for the MBO dynamic specification likewise, we estimated dynamic coefficients based on an identification which was quite conservative, but not really clear-cut: indeed, some IRCCS' researchers included in the control group were non-academic, hence they were enjoying the performance-based bonus. On top of that, although we already showed that the MBO-policy effect was quite neutral, it surely contributed to bring about some bias in the baseline estimates, especially for the double-treated individuals. Therefore, we performed a number of diverse dynamic models' estimation again, in doing this showing the graphic representations of the estimates foreseen by Equations 7 in Figure 6). In the latter, Graph a) reports the comparisons between double-treated units and the other groups, while Graph b) the comparison between IRCCS-only researchers and the other individuals.

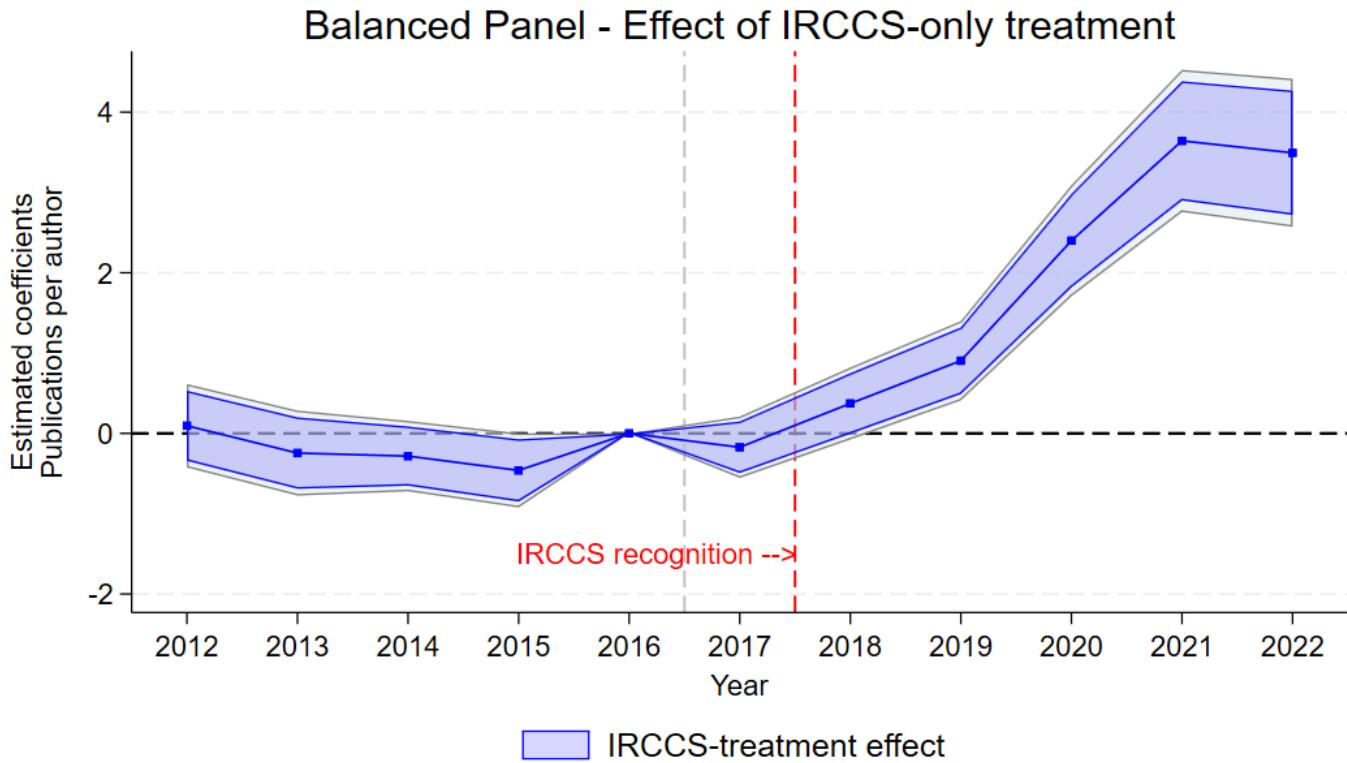
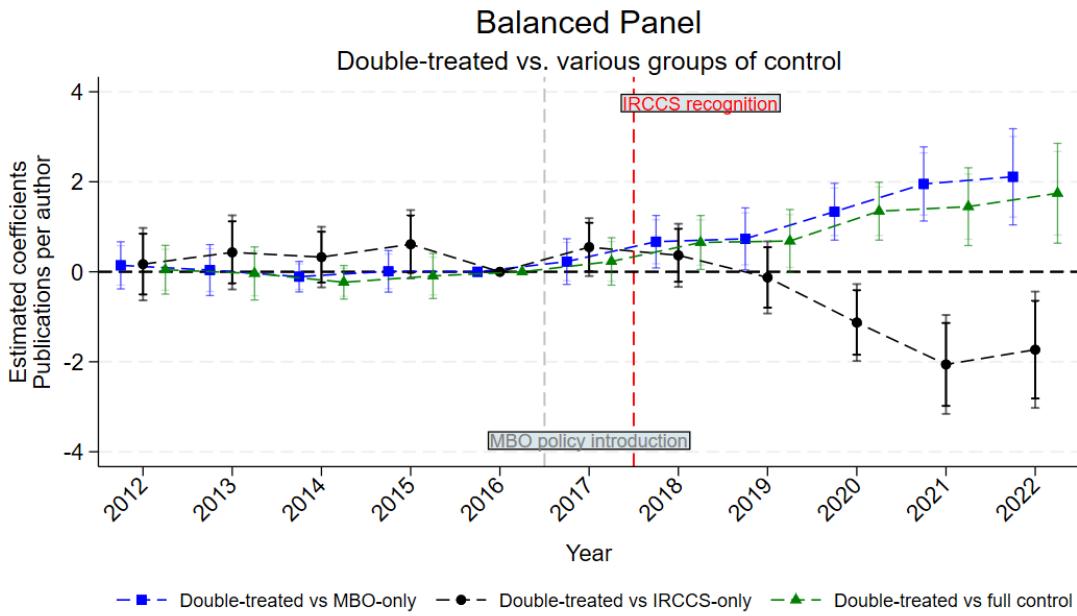


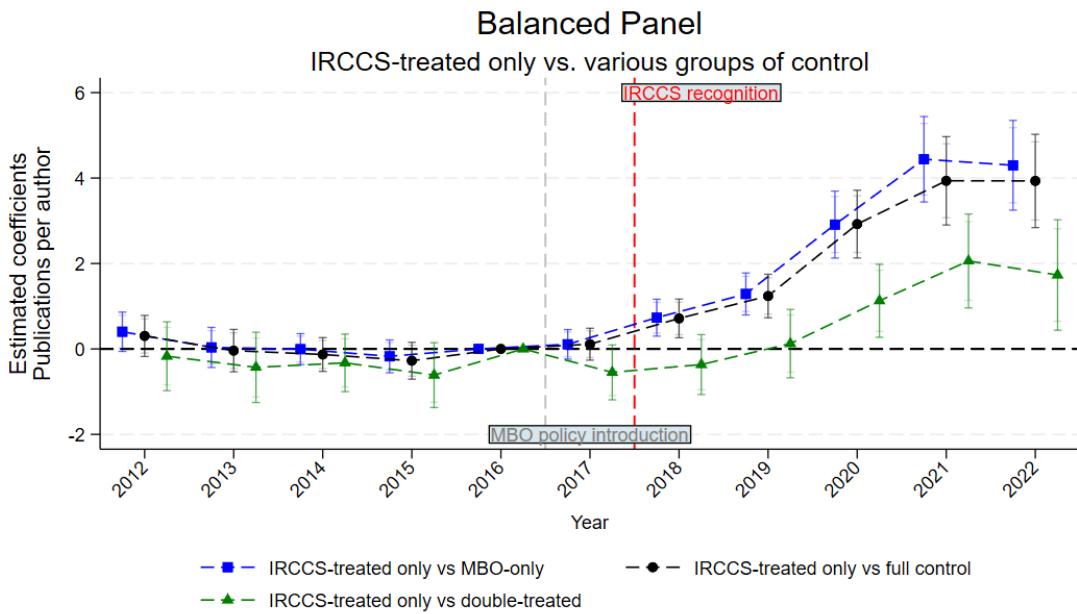
Figure 5: Event-study of IRCCS recognition effect on publications.

Both graphs keep validating the hypothesis that that the MBO policy does not really bias the estimates as its impact is almost non-existent; again, it slightly affects the research outputs only by shifting up the pattern of publications for the double-treated units; however, such pattern is more than offset by the IRCCS funding granted to academic researchers, whose trends follow along both in comparison with the pure controls and the MBO-treated only individuals.

It is however evident how the compound effect of the performance-based policy and the IRCCS access affects the dynamic comparison of IRCCS-listed academic doctors and the double treated. Although non-significant, we observe an immediate shift downwards in the publication output of IRCCS-only treated in comparison to the compounded ones in 2017, hence immediately after the MBO implementation (c, Graph b)). Anyway, such difference starts reverting back already the year after, with the IRCCS recognition occurrence, after which the IRCCS-only physicians start regaining their comparative advantage with respect to non-academic doctor, with their outcome boosted by the access to new resources.



a) Event-study of IRCCS recognition effect on publications of the double-treated units compared to the various control groups.



b) Event-study of IRCCS recognition effect on publications of the IRCCS-only treated units compared to the various control groups.

Figure 6: Event-study of IRCCS policy effect on publications for comparisons across different groups.

Overall, the IRCCS funding intervention appears to have had a substantial, almost immediate and persistent effect on research output, fed up by the pandemic and reinforced, only for some individuals, by the performance-based scheme. The estimates for the IRCCS-only group suggest that the heterogeneity in incentive structures across the different groups can only modestly be offset by rewarding productivity, if such reward is not paired with additional resources. It must also be noted that such additional resources already impact positively those who are selected into the perimeter of those who can actually access them, who are chosen by the management due to their better supposed performance in research-related activities; hence, such individuals possibly already embed a slightly difference incentive structure, if compared to the non-academic doctors for instance. In any case, the presence of academics as well in the pure control groups confirms that the heterogeneity in pre-existing incentive behavioral structure is not only present across groups, but also within them. To sum up, it looks like that academic researchers, already incentivized by intrinsic and career motivations, react substantially when provided with more resources, whereas individuals not attracted by the same motifs or similar career progression incentives, show no incremental response to the monetary scheme, or do so only in presence of extra funding. Given the relevance of such results, we next provide with a range of robustness checks to ensure that these findings are not driven by other confounding factors. The first one is included in this very same section, and address the concern of potential bias driven by the violation of the parallel trend assumption.

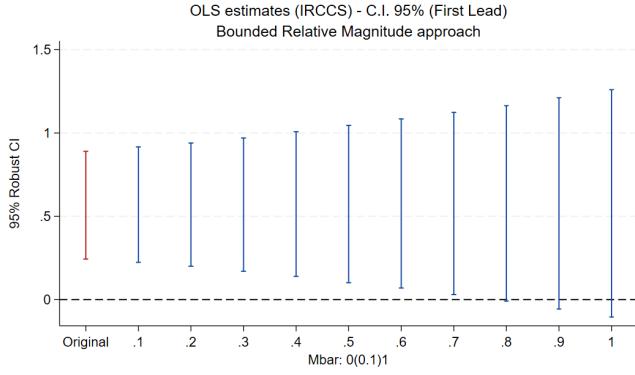
### 5.2.1 Parallel Trends

While the main estimates seem to display the absence of deviation from the parallel trends in all the provided specifications, we give a greater focus on the pre-trends' issue, given the significance and magnitude of our results. This check is also required due to the fact that the treatment is endogenous and its assignment certainly depends on existing characteristics, which are heterogeneous across groups and units. First of all, notwithstanding the overall absence of statistical significance for the pre-treatment lags interacted with the treatment dummies, we still observe that, especially in our baseline graph in Figure 5, there is a slightly decreasing (albeit non-significant) trend for the IRCCS staff in the periods prior to 2016, which seems to become also statistically significant at 95% in 2015. Although a visual inspection would not assess such deviation as able to undermine the validity of the identification, and even if the trend is falling down (which, if anything, would lead our estimated post-treatment effects to be downward biased), we follow the *honest* methodology by Rambachan and Roth, 2023 to account

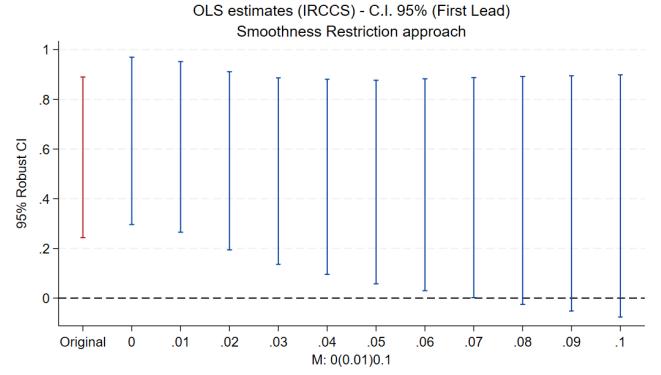
for plausible deviations from the common trend. In accordance to their approach, we can undertake sensitivity checks even by hypothesizing relevant deviations from parallel trends. We do so by computing *bounds on relative magnitude* (*BM*): as in, we replicate their methodology by estimating the entity of the divergence from the parallel patterns after the treatment, up to the size that could be able to invalidate the significance of the post-treatment outcomes, with respect to the assumption of the absence of pre-trends. Such potential *post-* deviation is assumed to have the same size as the one in the period reporting the greatest difference in trends among all of our pre-treatment coefficients (in our baseline case, 2015). We also estimate a second bunch of sensitivity checks, where the assumption regarding the entity of the post-treatment deviation is not made according the maximum size of one of the lag-related differences, but due to a linear extrapolation of the hypothesized pre-trends which would lead the estimated results to be non-consistent (*smoothness restriction - SR*). We provide evidence for the *honest* confidence bounds for the baseline estimates (Figure 7), using both approaches (*BM* in graphs a) and c); *SR* in graphs b) and d)), and by allowing variation in the common trends able to invalidate the significance of the coefficient on the first lead estimation (graphs a) and b)) or the overall average across all post-treatment coefficients (graphs c) and d)). Results are reported, as said before, in Figure 7. In each graph, the y-axis represent the robust 95% confidence interval for the target estimated coefficient. The x-axis, on the other side, represents the varying parameter of interest,  $Mbar$  in the bounded relative magnitude method and  $M$  in the Smoothness one. For the *BM* approach, it represents the percentage of the size of the post-treatment violation (based on the maximum pre-treatment difference) for which the estimate is allow to be biased before becoming non significant. For the *SR* approach, the x-axis represent the percentage of the deviation in the slope from the linear extrapolation of the pre-trend divergence (based on the post-treatment evolution) that would make the estimate invalid. In all graphs, the baseline confidence interval (as in, the one with no deviation from common trend assumed) is in correspondence of  $Mbar = 0$  (or  $M = 0$ ) and it is indicated by a vertical red C.I. line.

Graphs a) and b) display the robustness confidence set of the estimated coefficient in the first lead for the baseline event-study. According to the *BM* approach (graph a, the x-axis ranging from 0 to 100%), the effect maintains its significance up to assuming an actual 80% deviation from the maximum divergence measured in the pre-treatment period. The smoothness restriction approach instead tells that the first lead would be invalid only if the slope divergence in the pre-treatment period from the linear extrapolation obtained from the leads' pattern overcame 7%. Considering that we are referring to the first lead (when the effect is lower compared to the rest of the post-treatment period, and it has not

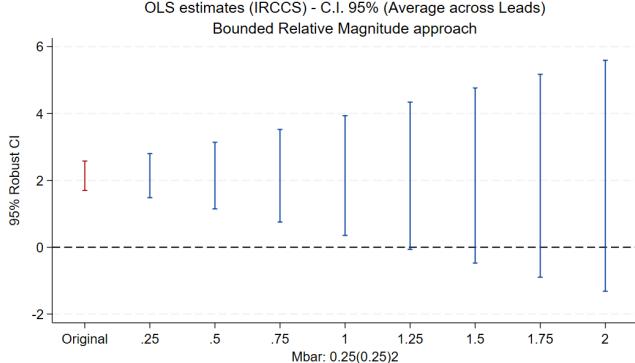
fully kicked in already), we can consider such results as robust. If we look at the average effect across all leads (as in, the ATT), we observe that the estimated coefficients are even more robust: indeed, they maintain their validity up to a 125% bounded relative magnitude violation, and up to a greater than 12.5% difference in the linearly extrapolated slopes. This confirms the validity of the parallel trend assumption in our framework. We also report the *honestdid* estimations for the various comparisons across group in the Appendix, in Figures from [A1](#) to [A6](#).



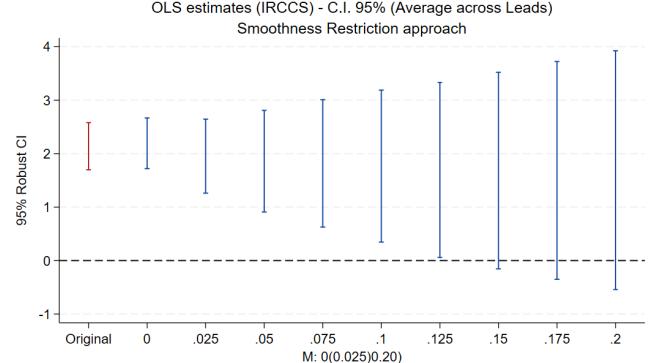
a) BM approach for the significance of the first lead.



b) SR approach for the significance of the first lead.



c) BM approach for the significance of the average across all leads.



d) SR approach for the significance of the average across all leads.

Figure 7: Honest DiD robust confidence sets for overall IRCCS effect estimated with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).

## 6 Robustness Checks

In this section, we present further robustness tests to verify whether the estimated effects are driven by the policies or by other possible confounders. Our main and first check is the performance of alternative

estimation techniques for staggered treatments, put in place due to the potential problem of self-selection into treatment that occurs due to the fact that some researchers are formally included in the IRCCS perimeter some time after the IRCCS recognition in 2018. We then address potential bias from the COVID-19 shock, control for spillover effects in publications, and account for individuals who changed status and for outliers.

## 6.1 Validity of the TWFE $2 \times 2$ DiD identification

We already acknowledged that the IRCCS recognition happened at once at the institution level. However, even though there exists a bulk of core researchers who stay in the IRCCS list for the whole post-treatment time-span, the perimeter is re-updated annually by the hospital management. Therefore, the inclusion of some individuals in the IRCCS research staff could be considered as a staggered adoption process: as a matter of fact, while in our main specification we considered all IRCCS *ever-listed* individuals as “treated” in 2018, only some of them were formally included in the perimeter during that year, as others should have sensibly be considered as not yet treated. While people hired after 2017 is not a concern, as we only keep 2012-2022 balanced units, some faculty members or even some non-academic MDs (the double-treated) may get added to the list in a following period, when the perimeter is upgraded. We show the evolution of the “aggregation” of new units into the IRCCS perimeter in Figure A7, in the Appendix. As consistently displayed in the bar graph, the “core” of IRCCS researchers stays quite constant in the first 3 years after the recognition (152, 140, 141), jumping to 181 units in 2021 only, and even overcoming the threshold of 200 in 2022, possibly due to the research-boost triggered by COVID. If the oscillation around the “main core” of the IRCCS’ researchers kept within the magnitude of the deviations displayed between 2018 and 2020, there would be no reason to suspect that self-selection was upward biasing our estimates. However, the jump observed after 2020 requires us to motivate more specifically our decision to implement a TWFE DiD as our main estimation strategy even in presence of a (mild) form of staggered adoption. First, we observe that, notwithstanding such approximation, the all basic parallel trend tests show no anticipatory assumption nor divergence in the pre-trends. Second, such tests are all corroborated by using the *honest* methodology of Rambachan and Roth, 2023 to check for the validity of the effects in presence of potential deviation from the pre-trends (which we do not actually observe). Finally, to confer more validity to such simplification, we remove from the sample all the (few) “late-adopters”, as in those who are not part of the initial core of IRCCS researchers, to assess whether the effect is driven by them or if it is biased up by selection issues.

Table A2 in the Appendix proves us that is not the case, as the effect are still present (and, indeed, magnified) by excluding those who are selected into the research perimeter after the IRCCS' recognition. The dynamic estimates derived from such sensitivity checks are also reported in the Appendix, in Figures A8 and A9, and both corroborate the validity of our simplification procedure by showing the absence of pre-trend. The absence of an effect for the Double-Treated individuals is reasonably due to the lowered statistical power of estimates caused by the subsetting of the dataset. As a matter of fact, the bulk of the “late-IRCCS adopters” is indeed made by non-academic physicians later included into the IRCCS perimeter due to their productivity.

To finally benchmark our baseline estimation with the identification strategy that would be formally required in our framework, we estimate a Standard two-way fixed effects DiD with staggered adoption to assess such effect. However, TWFE estimators can be biased when treatment when dealing with heterogenous timing, above all when dynamic effects may not be homogenous over time, or if there is endogeneity in the adoption (de Chaisemartin and D'Haultfoeuille, 2020, Callaway and Sant'Anna, 2021, Sun and Abraham, 2021, Borusyak et al., 2024). Even if most of the treatment happens in one period only (2018), to be safe we apply the estimator by de Chaisemartin and D'Haultfoeuille, 2020, which re-weights the estimates through the comparisons between joiners and untreated units, and leavers and already treated units. This shall help getting unbiased average treatment effects in staggered designs with heterogeneous effects. We estimate event-study models for the IRCCS effect using both the standard TWFE method and the de Chaisemartin and D'Haultfoeuille (2020) weighted method. We only report, however, dynamic estimates, and our main specification is then:

$$Y_{it} = \delta_i + \tau_t + \sum_{k \neq -1} \beta_k \cdot D_{it}^k + \varepsilon_{it}, \quad (8)$$

where  $Y_{it}$  is the outcome of interest for unit  $i$  in year  $t$ ,  $\delta_i$  and  $\tau_t$  are unit and time fixed effects.  $D_{it}^k$  is a binary variable taking unitary value if  $i$  is  $k$  years away from treatment in  $t$ , 0 otherwise. We normalize the coefficient for the period immediately before treatment ( $k = -1$ ) to zero, so that all  $\beta_k$  coefficients are interpreted relative to the year before the IRCCS inclusion. The estimate is integrated by the approach by de Chaisemartin and D'Haultfoeuille, 2020. Appendix Figure A10 shows the dynamic effect of IRCCS using TWFE and the corrected dCDH estimator. The patterns are very similar, although the TWFE estimates significantly overstate some of the late coefficients. This is reasonable, as the weights constructed by the corrected estimation methodology are based on the comparison between “joiners” (as

in, late adopters) with the never-treated units. We replicate the estimates for key subgroup comparisons (e.g., double-treated vs. MBO-only in a staggered framework), to observe the same results (Appendix Figures ?? to ??). This suggests that selective inclusion into IRCCS (which indeed exists and is due on past performance and productivity) does not compromise our main findings. In summary, all robustness checks confirm the credibility of our results. Note that further recent estimation techniques, such as Callaway and Sant'Anna, 2021, would provide with the presence of pre-trends, driven by the late-adopters, due to selection into treatment credited to recognition of past productivity. This is acknowledged and self-explanatory, but does not invalidate the significance of our estimates<sup>3</sup>

## 6.2 Excluding COVID-19 Years

It seems clear that the COVID-19 pandemic outbreak in 2020 substantially triggered upwards the pattern of publications. The main concern would be if such event could have affected research productivity, independently on the granting of IRCCS funding to the hospital. The main confounder would be the re-allocation of time due to the pandemic: as a matter of fact, time and efforts could have been devoted to medical assistance at the expense of research, especially for non faculty doctors (Mantellini et al., 2020, Ahrendt et al., 2022, Franzoni et al., 2025). In such case, the diverging pattern should not be credited to the take-off of the funding policy, but to the reduced research activity by physicians. On the other hand, the postponement of elective activities might have given some doctors more time for writing; in addition to that, due the joint national effort put in place in the struggle to counteract the pandemic, COVID-related research at the time could have been easy to perform, with lower publishing lags. While the latter two hypotheses are not necessarily a concern, since they would explain a feasible amplification of a still existing IRCCS effect, the first threat to validity would be a major confounder, able to substantially bias the effect in the reported estimates. First, such concern shall be mitigated by the fact that, notwithstanding that MDs were on the frontline in the struggle against the pandemic, also the assistance activity of faculty members was increased due to the emergency. However, to ensure that our results are not driven by anomalies during the pandemic, we re-estimate the main models using data only up to 2019<sup>4</sup>. We thus truncate the panel at 2019, and re-run the MBO and IRCCS difference-in-differences analyses. Tables A3 and A4 in the appendix show the estimates for the pre-COVID sample, both for the MBO and the IRCCS framework. In both cases, the results are similar in direction to the

---

<sup>3</sup>Estimates of the effects using the methodology of Callaway and Sant'Anna, 2021 can be provided upon request.

<sup>4</sup>Note that already by looking at the main event-studies we can retrieve that the effect kicks in before the pandemic started, which is already a good validation of our main results.

full-sample results, although the size of the coefficients is, as expected, reduced by the shortening of the time span. We avoid re-performing the event-studies with the reduced sample, as the full dynamic specifications already show that the first leads already hint the overall patterns described insofar even without Covid occurring.

### 6.3 Spillover Effects

A further concern for our identification strategy is the potential violation of SUTVA (*Stable Unit Treatment Value Assumption*), as in, that the policies might have indirect effects on those who are not directly targeted, mostly due to collaboration. For instance, IRCCS physicians who do not receive MBO-bonuses might however “spill” their gains towards non-academic MDs or even the double-treated ones, by increasing their output and at the same time co-authoring with the performance-based remunerated doctors. Since we observe that the IRCCS-funded researchers are the ones responding more significantly to the shock, if there was such a spillover effect, it would indeed be biasing our estimation downwards, in doing this reducing the differential across groups. However, we test for spillovers by conducting the following two exercises. First, we remove from the sample the double-treated units; then, we identify all publications involving a collaboration between an MBO-eligible physician and an IRCCS-listed researcher. We call publications of *Type A* those which, in the panel, can be flagged as authored by a MBO-treated individual in collaboration with at least one member of the IRCCS perimeter. *Type B* publications are instead those identifiable as authored by IRCCS staff and co-authored with at least a non-academic physician. Then, we perform two Diff-in-diff estimations for each exercise. In the first we compare MBO-treated to the rest of the sample (as in Equation 2, but without the compound treatment), by first excluding authors of *Type A* papers, hence the non-academic MDs who *ever* collaborates with an *even* non-MBO IRCCS member. Then, we do the same for the IRCCS-treated (mirroring Equation 6 while excluding double-treated again), this time excluding *Type B*, as in IRCCS researchers who *ever* co-author with MBO-treated units. This quite stringent exclusion of *all* authors of cross-written papers ensures that the treatment effects only stem from either within-group or individual publications. Such approach eliminates entire collaboration links between the two groups. We avoid removing both groups from the sample at the same time in order not to lose too much statistical power for the estimates. The results of these two exercises seem to remain consistent with the main findings (Table A5 in the Appendix). The Management-by-Objective impact remains negative and statistically significant when both MBO-treated and IRCCS-treated authors of cross-group joint papers are removed. As a matter of

fact, the absolute value of the negative effect is magnified when *Type A* researchers (non-academic MDs collaborating with IRCCS individuals) are removed, while the drop is weakened when excluding *Type B* researchers from the regression (academic IRCCS collaborating with MBO-treated individuals). The same, but in the opposite direction, is retrieved when testing the consistency of the effect of the IRCCS recognition on the productivity of empowered researchers, even when excluding the network groups. The evidence suggests that the observed effect may be actually dampened by the collaboration with IRCCS researchers sought by MBO-only doctors; absent such behavior, our estimates could be actually showing even greater divergent patterns. The related dynamic estimates of such robustness checks are reported in the Appendix as well, in Figure A13, Graphs a) and b) respectively.

## 6.4 Controlling for outliers and switchers

Our main estimates have been performed, so far, by considering group status as fixed and absorbing, based on the initial policy definitions at the moment of the implementation of the given policies. In reality, there could be some variations over time due to changes across different roles. For instance, some physicians shift from non-academic positions to academic role at a certain point in time post-treatment, exiting from the group of those who are eligible in the MBO. On the other side, non-prolific academics may lose their tenure and become MDs without faculty appointment. In either case, the effect of productive non-academic MDs considered as “fixed” MBOs at the implementation of the MBO switching into the academic group, as the opposite pattern indeed, should again bias down our estimates, as the publications of these more (less) productive individuals shall be accounted for in the outcome of the control (treatment) group in the baseline analysis<sup>5</sup>. To guarantee that these ”switchers” are not affecting our results to a great extent, we perform a robustness check where we exclude any unit whose status changed during the observation treatment window. To be precise, we allow for three scenarios: 1) drop any non-academic physician in 2017 who becomes, at a given point, a faculty member after 2017 (shifting from MBO-eligibility to non-eligibility); 2) any faculty member who left academia after 2017, to sort into non-academic medical practice; 3) any individual who undertook either of such changes. Cross-position switchers account for barely 3.5% of the dataset. The findings (Appendix,

---

<sup>5</sup>Note that one might argue, for the baseline comparison, that switching out of the MBO-treatment group should be accounted as a violation of the absorbing treatment assumption, as MDs who become academics lose the eligibility to the performance-based payment. However, it is reasonable to keep such individuals in the MBO-treatment group, as the purpose of the performance-based payment is exactly that of triggering an incentive mechanism for which individuals start publishing after receiving the money in the first place. If after that they keep publishing even after losing the eligibility to the MBO, and if they adjust their incentive structure to such a point that they even switch to an academic position, the policy could be considered as effective.

Table A6) are extremely similar in all scenarios. If anything, and quite surprisingly, the MBO negative effect decreases slightly in absolute terms (suggesting the switchers to academia were not so prolific as expected); the same happens to the IRCCS effect, although in the opposite direction, confirming the mentioned unexpected finding. While such results suggest that dynamic transitions are not relevant enough to create distortions in the effects' estimates, the fact that the estimates slightly diminish after removing switchers suggest that those who cross groups from MD to academia and vice versa were not less (more) productive on average. Hence, to check whether the different groups have different incentive structures moving them for real, we need to identify those individuals who may shift upwards the overall publication numbers of their respective group, especially with regards to the IRCCS-affiliated members. Hence, we re-perform the analysis by removing: 1) outliers (as in, the last percentile of the annual publication distribution); 2) the top 10%; 3) the upper quartile of the annual publication distribution<sup>6</sup>. We do this and re-estimate both the baseline MBO and IRCCS specifications: they are reported in the Appendix, in Tables from A7 to A9 for the MBO-treatment estimates, and in Tables from A10 to A12 for the IRCCS ones. As expected, the negative shift after the MBO-implementation for treated units, as for the positive effect of the IRCCS recognition decrease in size when always more units in the upper parts of the publications distribution are removed from the sample. However, the directions of the effects stay the same even when eliminating one fourth of the sample; in addition, even if some coefficients lose some degrees of significance, the ones who were statistically different from zero in the main analysis performed on the whole sample never lose such feature. Hence, it is evident that the huge effect we observe is driven but not entirely made by a sub-group of scientific superstars among IRCCS researchers.

## 7 Research Quality and Networks

It is valuable to assess whether the policies had a significant impact on the nature of the research collaborations stemmed after their implementation, and the impact/quality of the output. We therefore perform two sets of additional estimates in order to tackle two questions mostly: (1) whether and to what extend the policies affected research collaborations across the two groups (which so far we have used only as an excluded group to perform robustness checks), and (2) how our results relate to research quality of the members of the institution under study, proxied here by citations.

---

<sup>6</sup>Note that the threshold for the upper quartile is just 3 annual publication, so we drop all those who have published at least 3 publication in any year, but we keep those who have published exactly 3.

## Cross-groups Co-authorship

We begin this section by defining all papers listing, among the co-authors, both a non-academic doctor (subject to MBO) and an IRCCS researcher of the same institution as “networked” publications; this is a decent representation of network effects between the two categories of hospital employees. We make use of the same categorization we employed when controlling for the validity of *SUTVA*, by subdivide networked publications according to the groups’ perspective. *Type A* are those for which an MBO-eligible individual is an author (MBO-group outcome count) and at least one co-author is a designated researcher. The opposite holds for *Type B* papers: a IRCCS physician is an author (IRCCS-group outcome count), and at least one co-author comes from the non-academic physicians’ group. Therefore, *Type A* and *Type B* label the very same publications, but they are attributed to different groups for our assessments’ purpose. In doing this, we consider the number of *Type A* publications for each non-academic individual and that of *Type B* works for the IRCCS units as different outcomes. Table 6 displays the DiD estimates for the effect of the policies on the collaborations as defined above; Cols. (1) and (2) report coefficients for *Type A* and *Type B* outcomes in the MBO-based framework. Cols. (3) and (4) do the same for the IRCCS-based setting. For collaborations of MBO-treated physicians with IRCCS researcher (*Type A*), a significant increase in cross-group collaborations post-2017 (MBO) relative to non-treated MDs is retrieved (23.7 p.p., which amounts to 76.6% of the pre-treatment mean). On the other side, the effect on *Type B* publications, as in those included in the count for IRCCS-researchers (which are double-treated here, as they are control group in the MBO-based DiD) and co-authored with the treated units, is negative and significant (-18.5 p.p., -60%). By contrast, after the IRCCS recognition, we observe a significant decrease in *Type A* publications for the treated (which, again, are double-treated, as in this setting the MBO-only treated individuals are in the control group), which amounts to -12.8 p.p.. However, such decrease is quite irrelevant relative to the mean of the treatment group prior to the IRCCS recognition (-3.5%). On the other side, the increase in *Type B* papers for IRCCS-treated individual amounts to 38.2 p.p., a bit greater coefficient but still quite negligible relative to the mean (+10.6%). Such findings suggest that non-academic doctors who were prompted to do research, apparently sought out more partnerships with the members of the IRCCS perimeter. Similarly, for the academic IRCCS researchers (*Type B* outcomes), there is an increase in collaborations with non-faculty MDs after 2018, compared to non-IRCCS academics. While the increase in co-authorship between MBO-treated and IRCCS members is quite substantial relative to the mean of the non-academic physicians, the incentive to collaboration for the IRCCS’s individuals is sensibly lower

in percentage compared to their pre-2018 average. While the results may suggest that the policies led to better integration between the two groups, we could also foresee the emergence of strategic behavior among MBO-treated individuals aimed at improving their performance indicators by aggregating their activity to that of IRCCS-based researchers.

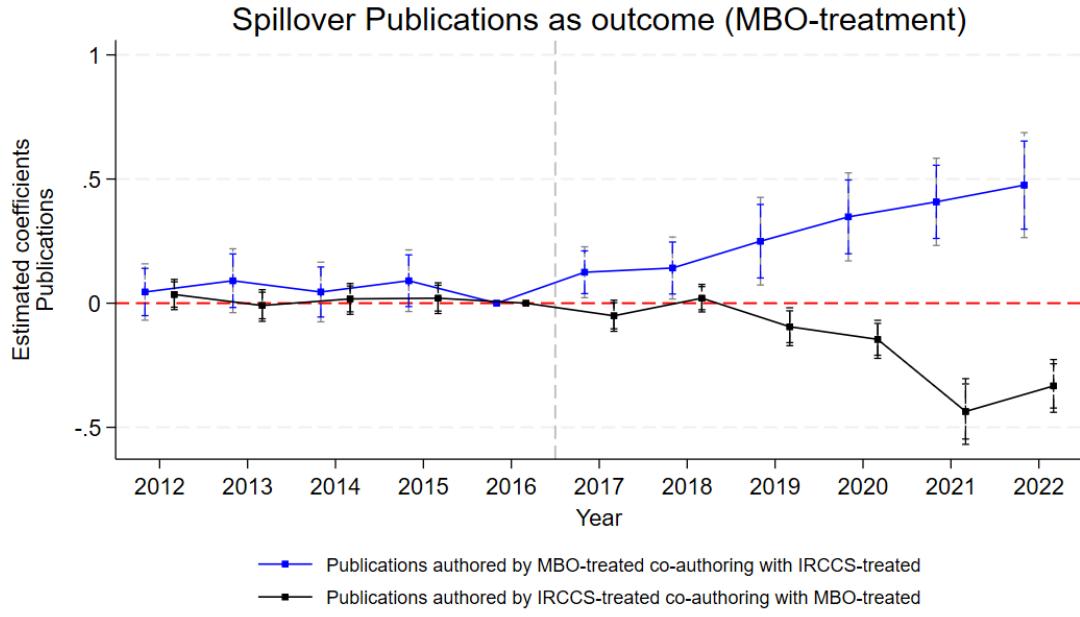
	2012-2022; SEs clustered at individual level			
	(1) Publications by MBO-treated co-authored with IRCCS (Type A)	(2) Publications by IRCCS-treated co-authored with MBO (Type B)	(3) Publications by MBO-treated co-authored with IRCCS (Type A)	(4) Publications by IRCCS-treated co-authored with MBO (Type B)
Post 2017*MBO	0.23717 *** [0.04766]	-0.18498 *** [0.02874]		
Post 2018*IRCCS			-0.12843 *** [0.02778]	0.38241 *** [0.05308]
Observations	5,691	5,691	5,691	5,691
R-squared	0.48304	0.49938	0.47510	0.51355
Individual FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022
Panel	Full	Full	Full	Full
Mean	0.309	0.309	3.609	3.609

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

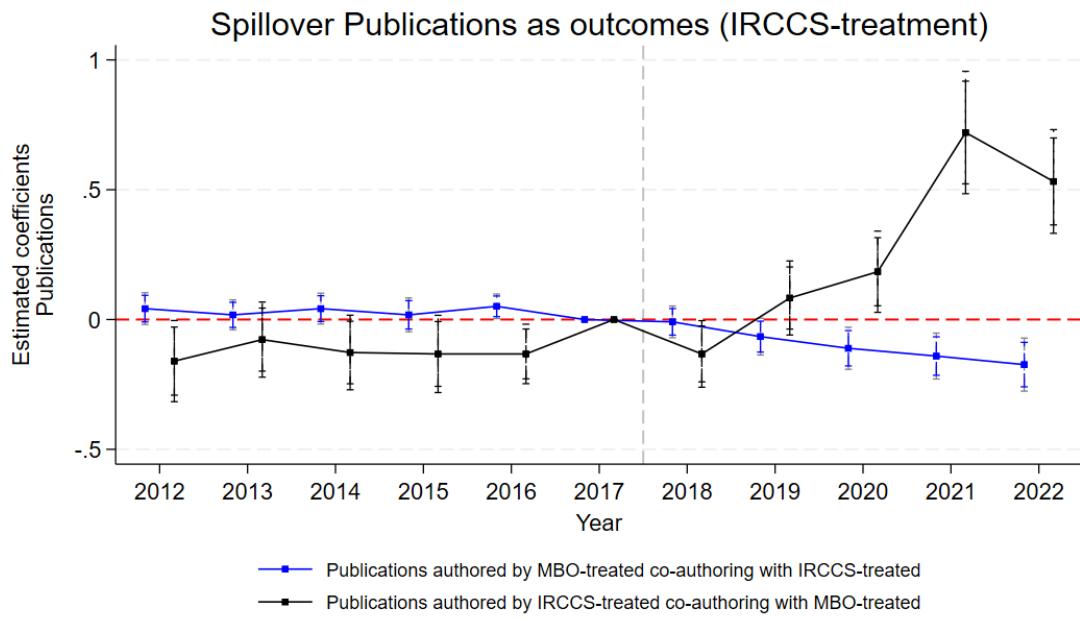
Table 6: Impact of MBO-policy and IRCCS recognition on annual cross-group publications (Difference-in-Differences).

Figure 8 reports visually the dynamic estimates for networked publications. Panel (a) shows the event-study for *Type A* (MBO group co-authorship) after 2017, while Panel (b) for *Type B ones* (IRCCS group co-authorship) after 2018. Both graphs indicate little to no difference in cross-group collaboration rates between the two groups prior to the policies. After the two policies implementations, the affected groups show substantial effects. In (a), all MBO physicians start collaborating way more with IRCCS-researchers in 2017, and such increases is persistent through 2018 and beyond. By contrast, double-treated physicians appear to collaborate less with non-academic MDs, although such pattern starts showing up in 2019 only. In (b), IRCCS academics are featured by a relative increase in collaborations with MBO doctors with a lag after 2018 (the effect seems to be small a negative and barely significant in the first post-treatment period). On the other side, as already showed in graph a) as well, double-treated physicians reduced their collaboration activities with other IRCCS researchers, by displaying a modest in relative terms but still persistent decline. While such effects would suggest an increased interdisciplinary externalities, which would surely benefit research by allowing a easier transfer from science to clinician practices, the fact that the effect is substantial only for MDs who eligible to the MBOs, negative for the double-treated overall and positive and significant but slightly relevant for the

IRCCS researchers could underline the presence of free-riding behavior in publications that, although undistinguishable in the presen analysis, could be better assessed by looking at the papers' impact.



(a) MBO physicians collaborating with IRCCS researchers (MBO-treatment effect)



(b) IRCCS researchers collaborating with MBO physicians

Figure 8: Event-study of policy impact on cross-group collaborative publications (IRCCS-treatment effect).

## Citation Impact

A further relevant question is whether a greater number of publications also triggered a positive shift in the influence and quality of such works, or it came at the expense of it. We hence analyze the policy effect on citations, employed as proxy of publications' impact. We do that through re-modeling the IRCCS equation (Equation 5) in order to adapt it to a set of different outcomes and weights. First, as discussed already in other studies such as the ones by Azoulay et al., 2011 or Waltman and van Eck, 2013, we cannot estimate the effect on citations without accounting for the publications' age. Thus, we build a yearly discounting term to control for the period when the citations were gathered, which is Fall 2024 ( $[2024 - t]^{-1}$ ). After doing that, we normalize each paper for time by multiplying its count of citations by the discount term, and build a yearly aggregate individual impact weight by summing up all yearly normalized citations for every researcher. This is the first normalized variable we obtain, which we use first as a Inverse Probability Weight in a regression of the treatment on the count of publications, and then as an outcome. The second outcome is obtained from the sum of all non-discounted yearly citations received by papers published by a given author in a given year; after aggregating them, we normalize such total by time, making use of the discount term described already. Eventually, we build an averaged measure of publications' quality, by dividing the number of total (discounted) citations of papers published in a given year by an author by the number of total papers published by the very same author. In these regressions, as in the latest ones, the IRCCS-treatment is assigned only to IRCCS researchers non-eligible to the MBO, while the MBO-treatment also considers the double-treated.

	2012-2022; SEs clustered at individual level					
	(1) Publications (IPW) (by no. of discounted 2024 citations)	(2) Sum of discounted 2024 citations	(3) 2024 Citations (over yearly publications)	(4) Publications (IPW) (no. of discounted 2024 citations)	(5) Sum of discounted 2024 citations	(6) 2024 Citations (over yearly publications)
Post2018*IRCCS	5.26598 *** [1.88683]	20.70414 *** [3.41649]	-142.44386 *** [23.02236]			
Post2017*MBO				-2.98268 * [1.53918]	-9.53533 *** [2.01651]	60.70126 *** [18.88641]
Observations	3,258	6,369	6,369	3,258	6,369	6,369
R-squared	0.85726	0.44474	0.27837	0.85488	0.43467	0.27060
Individual FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced
Mean	3.609	16.18	302.8	0.653	2.492	123.7

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 7: Impact of IRCCS recognition and MBO policy on Publication quality (Difference-in-Differences).

The results are presented in Table 7 for the IRCCS and MBO effect on citations. Apparently, by (inversely) weighting publications by the number discounted citations, the effect of the IRCCS' recognition is amplified for more structured researchers. By looking at the sum of (discounted) citations as outcome, IRCCS recognition seems to have boosted the number of quotes for IRCCS physicians, by a number of 20.7 citations (+127%). However, the IRCCS group's post-2018 average number of (discounted) citations per publications shows a major and significant decrease:  $-142.44$ , amounting to almost  $-50\%$  of the pre-2018 mean. By contrast, we observe a negative, slightly significant effect on the number of MBO-eligible individuals' publications in the IPW-weighted regression, while the overall drop in the sum of discounted citations is quite striking in relative terms, as it amount to  $-9.5$  citations, as in  $-382\%$ . Quite surprisingly, and specularly to what emerged for the IRCCS-affected individuals, the impact on the average number of discounted citations over yearly publications is quite substantial: almost +61 average citations per publications, which means an increase of  $\approx 50\%$  compared to prior to 2017. Such findings leads to the conclusion that, while the two concomitant policies have an established effect on physicians' research productivity, which is positive for the structured selected academics and basically null for the incentivized MDs, the quality impact is basically the opposite. The aggregate increase (decrease) in the IRCCS- (MBO-) treated individuals' normalized citations is clearly driven by the shift in the raw number of publications, while the average quality of the works seems to be neglected by the units who can directly access funding, and maximized by the MBO-incentivized MDs. While prior beliefs would suggest that, due to the embedded incentive structures, core researchers would focus on quality over quantity upon accessing funds, it looks like that, once achieved the possibility to increase their resources, they fuel their scientific output numbers at the expense of the influence brought about by their research. On the other side, MBO-eligible individuals seem not to carry over projects indifferently in bulks, but even without significantly raising their publication rates, they tend to focus on the quality of their research. This may be due to the fact that the performance-based rewards are not only based on quantity, but also on the impact factor of the journal where publications are issued.

Such findings, integrated by the evidence of absent driving pre-trends or anticipatory behaviors, are displayed in dynamic terms in the event-studies in Figure 9.

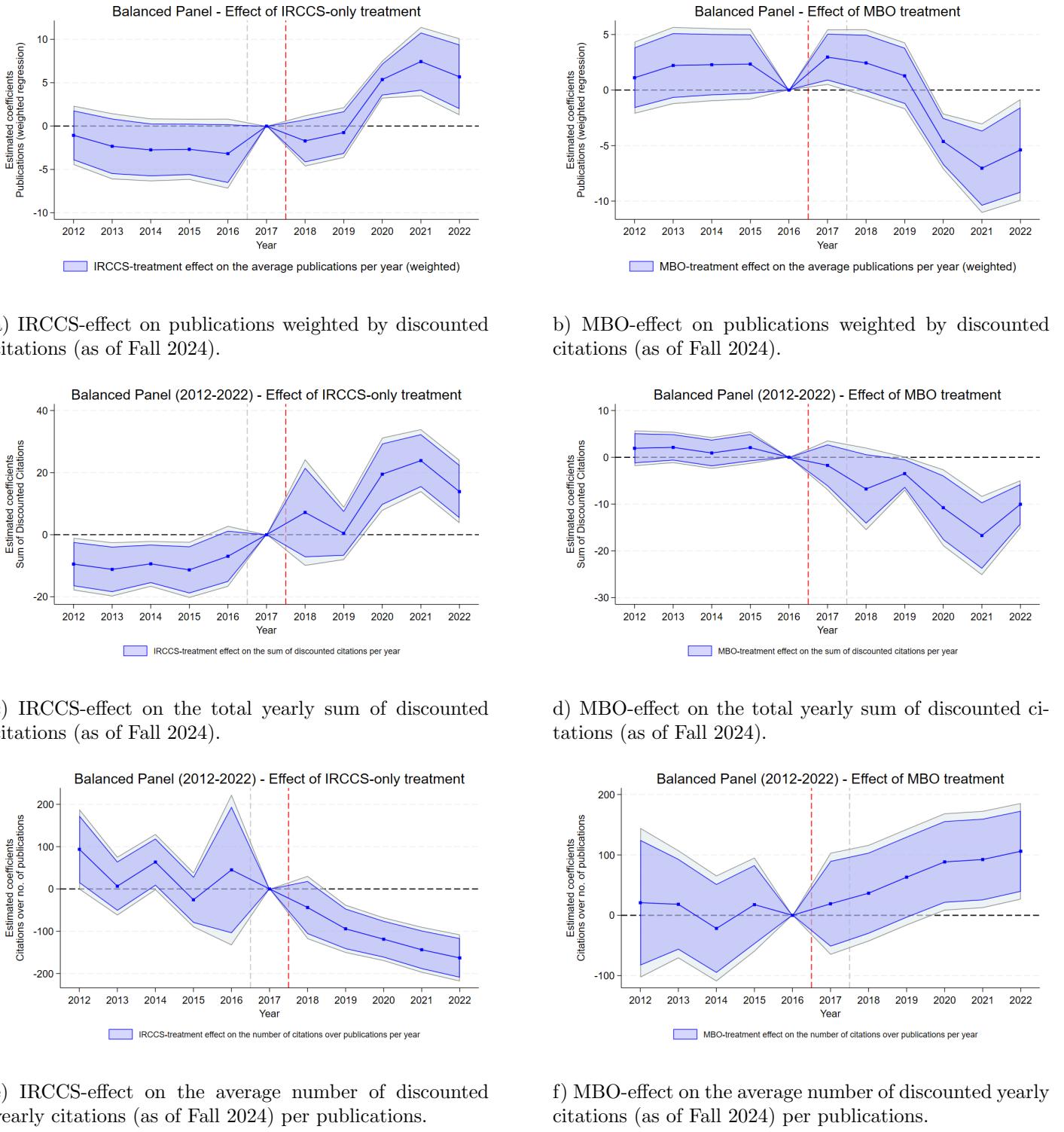


Figure 9: Event-Studies representing the dynamic estimates of the IRCCS recognition and MBO-policy impacts on the impact of published papers, proxied with yearly citations measured as of Fall 2024.

## 8 Conclusions

This paper examined two concurrent strategies to trigger research activity in a healthcare institution: direct monetary incentives based on performance for individual, non-academic, physicians versus increased public funding linked to institutional support. Using individual-level data from a major private Italian hospital, we found that the performance-based monetary reward (MBO) did not lead to significant improvements in research output of clinicians, who previously had limited publication activity, and were the main target of said intervention. The MBO seemed to spur a positive impact only on those MDs who also were involved in the research activity financed through the new accessible funding. As a matter of fuck, the increase of public research funds through IRCCS recognition produced a huge boost in the productivity of the academic physicians included in the research perimeter. Our findings highlight the relative importance of incentives in driving research productivity, which can be apparently improved through individual rewards only in presence of adequate, material resources. As non-academic doctors did not respond as strongly as expected, when incentivized to do so by monetary gains, our research shows instead how the different incentive structures embedded in the main groups who make up the hospital personnel under question play a major role in shaping research output patterns; indeed, the (substantially) increased number of works published by individuals already affiliated to the academic institution, seem to prove so. The findings seem to demonstrate that lack of proper funding was the main obstacle to academic research engagement, and that mere monetary rewards are not able to narrow the gap between the objectives of non-structured researchers and those who account on that to progress in their professional careers. We also document that the combination of resources and incentives may nonetheless yield interesting outcomes with respect of cross-group collaboration: as better funding allows scientific opportunities, individual incentives should ensure such opportunities to be exploited. However, relatively speaking, it looks like such opportunities have been mostly seized by less prolific researchers with respect to how structured scientists got involved in cross-group collaborations, possibly highlighting a scope for free-riding spillovers. However, the gain from such collaborations ought not to be underestimated in the long-run, as they could improve the scope of the synergies required to fostered the translational impact of medical research (i.e., clinical intuitions shaping academic inquiry and the opposite). Importantly, we find ambiguous effects on research quality. The increase in research output went along by a major boost in overall citation impacts for IRCCS researchers, and a substantial decrease for MBO-treated physicians, even in relevant terms. However, the direction this effect was evidently

driven by the shift in the raw numbers of scientific works: as a matter of fact, we observe a persistent decrease in the number of discounted citations per paper for IRCCS-researchers after the policy, and a positive, a bit lagged but significant trend in the average (discounted) impact of the publications of MBO-incentivized physicians, underlining their greater value for quality over quantity, which may be caused by either improvement of intrinsic attitudes towards valuable research, or by gaming strategies elaborated in concomitance with the bonus scheme designed by the management, which takes into account the impact factor of publishers. Moreover, the policies appeared to foster greater collaboration between clinicians and academics, which could further enhance the translational value of the research . This integration of efforts is particularly valuable in healthcare, where bridging the gap between frontline practice and research can accelerate innovation and implementation of findings.

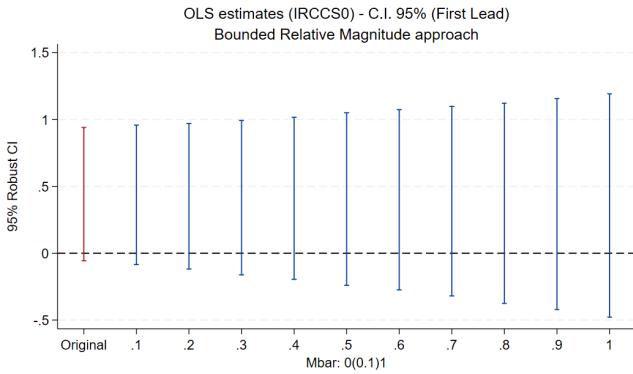
This paper contributes to policy discussions on how to stimulate innovation in public institutions and healthcare organizations. It also offers evidence that targeted performance incentives can only be an effective tool to unlock research potential, in presence of *actual* resources being provided to those who require them. For hospital administrators and policymakers, the results may imply that investing in a culture of research via both feeding long-term professional incentives relying on career progression and scientific curiosity (properly rewarded) may be more effective than short-run performance-based remuneration. It is indeed true that incentives ought to be aligned with multifaceted targets to avoid strategic behaviors which may possibly bring about detrimental societal repercussions. The main limitation of this study is its focus on a single (albeit large and relevant) institution, over a horizon which is long enough to observe individual behaviors, but not wide enough to capture the full unfolding of long-term research outcomes. Longer-term consequences on research quality, career trajectories, and patient outcomes, although beyond the current scope of this article, should be the main target for future related study. On top of that, while our quasi-experimental strategy helps disentangling the effects of the policies (and across strongly comparable individuals), broader studies able to better frame the external validity of our estimates to different conditions, such as public healthcare facilities or geographical heterogeneities, should be pursued. In any case, our results offers a compelling case that, within an imperfect market environment like the healthcare one, managerial triggers on incentives can only play an important role in fostering innovation if properly traded off with the necessities required by the need for proper funding.

# Appendix

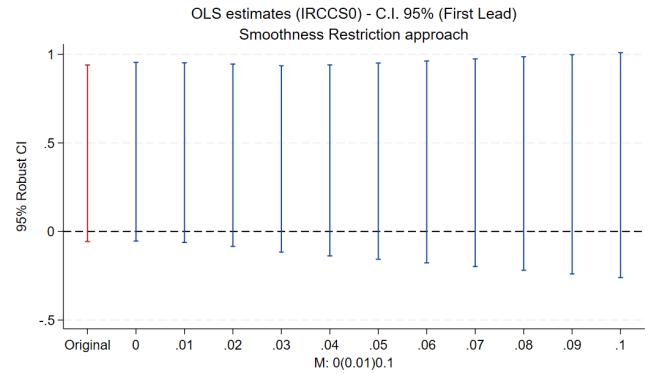
	2012-2022; SEs clustered at individual level							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Publications	Publications	Publications	Publications	Publications	Publications	Publications	Publications
Post2017								0.06839
IRCCS								[0.17762] 2.97093 ***
Post2017×MBO	-0.49848 *** [0.19260]			-0.10817			-0.00719	[0.40128] -0.11682
Post2017×IRCCS				[0.12208]			[0.09933] 0.41295 **	[0.27715] 0.28231
MBO×IRCCS	0.72602 * [0.42276]						[0.18272]	[0.24031] -1.83172 ***
Post2017×MBO×IRCCS		0.81178 ** [0.39721]			0.29829	0.03520	[0.24884] [0.30633]	[0.47509] 0.30768 [0.41861]
Observations	1,709	1,709	1,709	1,737	1,737	1,737	1,737	1,709
R-squared	0.37898	0.37871	0.37898	0.91201	0.91195	0.91212	0.91278	0.46120
Individual FE	NO	NO	NO	YES	YES	YES	NO	NO
Year FE	NO	NO	NO	YES	YES	YES	NO	NO
Covariates	YES	YES	YES	NO	NO	NO	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2016-2018	2016-2018	2016-2018	2016-2018	2016-2018	2016-2018	2016-2018	2016-2018
Panel	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced
Mean	1.671	1.671	1.671	1.671	1.671	1.671	1.671	1.671

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

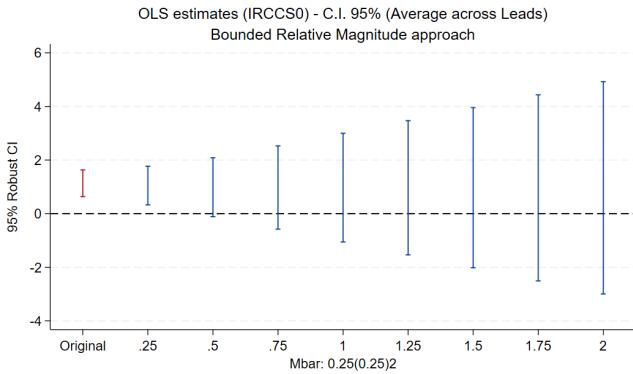
Table A1: Triple Diff-in-diff between 2016-2018, comparing different groups.



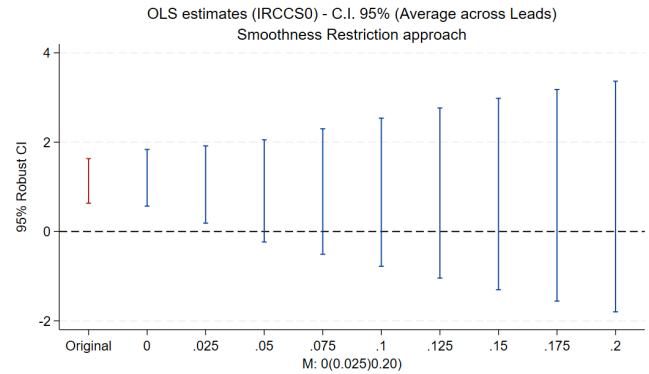
a) BM approach for the significance of the first lead.



b) SR approach for the significance of the first lead.

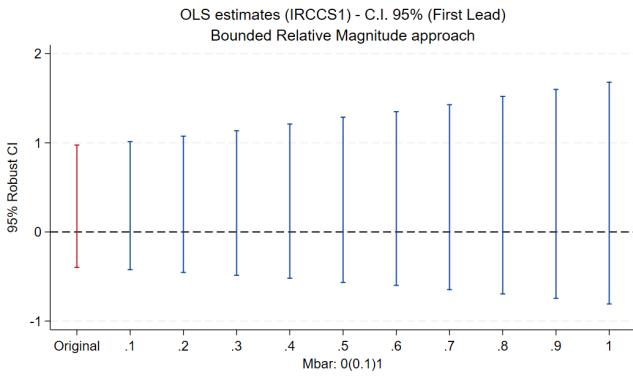


c) BM approach for the significance of the average across all leads.

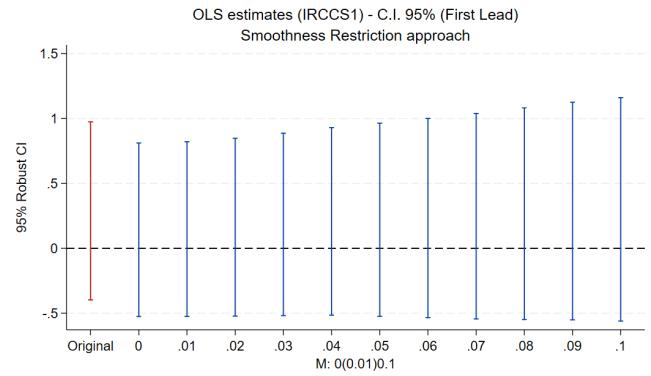


d) SR approach for the significance of the average across all leads.

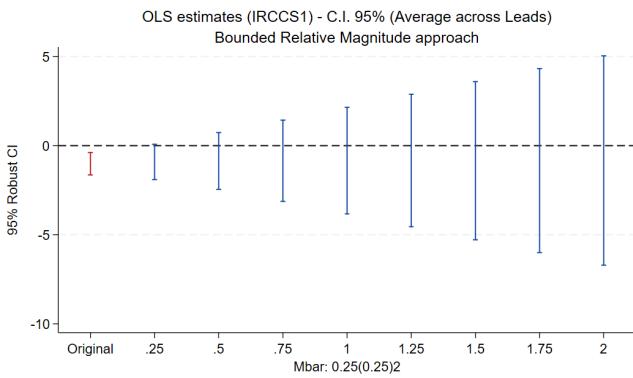
Figure A1: Honest DiD robust confidence sets in the comparison between double-treated and MBO-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



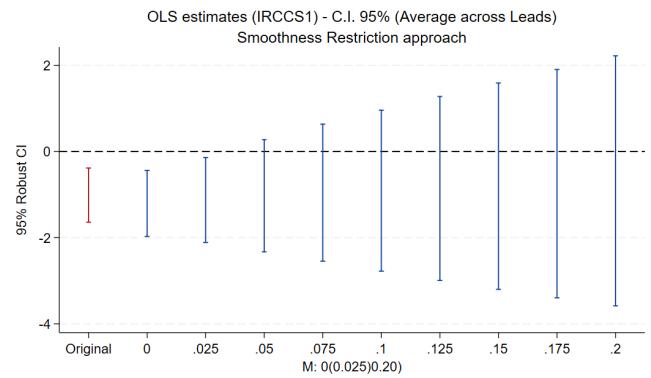
a) BM approach for the significance of the first lead.



b) SR approach for the significance of the first lead.

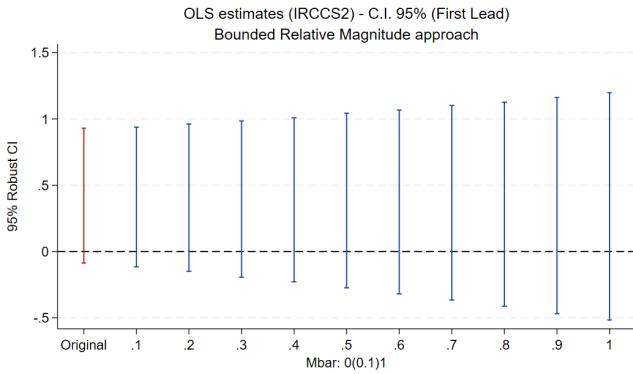


c) BM approach for the significance of the average across all leads.

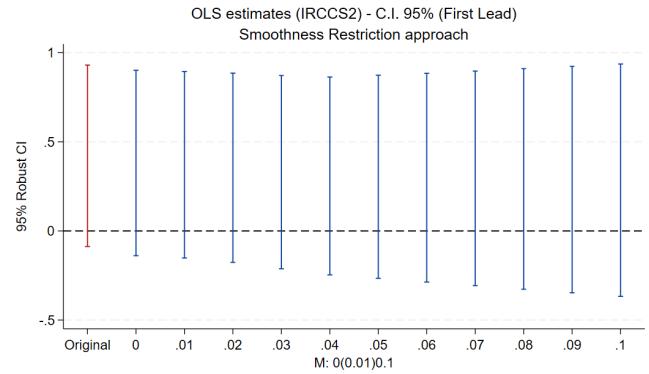


d) SR approach for the significance of the average across all leads.

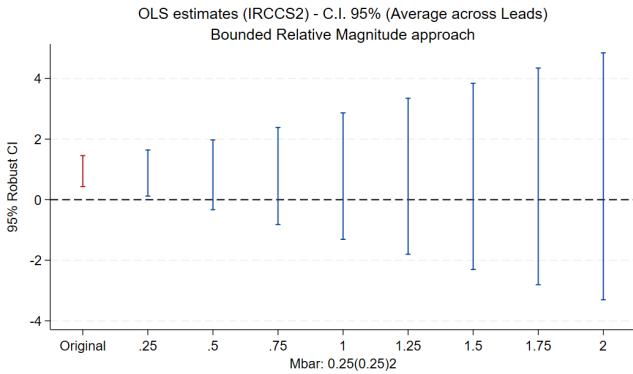
Figure A2: Honest DiD robust confidence sets in the comparison between double-treated and IRCCS-only units, estimated the overall IRCCS effect by adopting different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



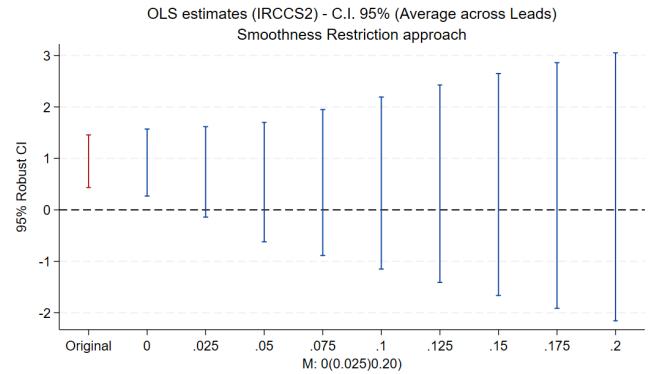
a) BM approach for the significance of the first lead.



b) SR approach for the significance of the first lead.

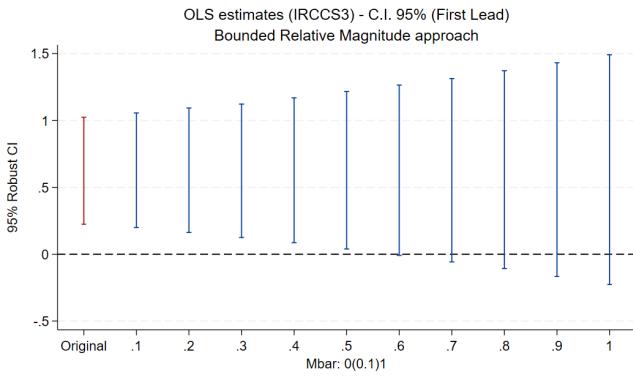


c) BM approach for the significance of the average across all leads.

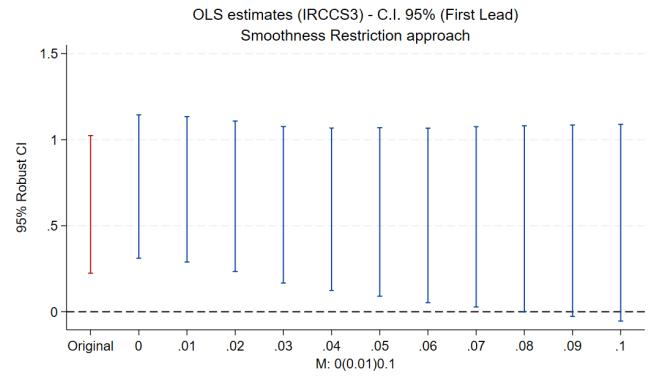


d) SR approach for the significance of the average across all leads.

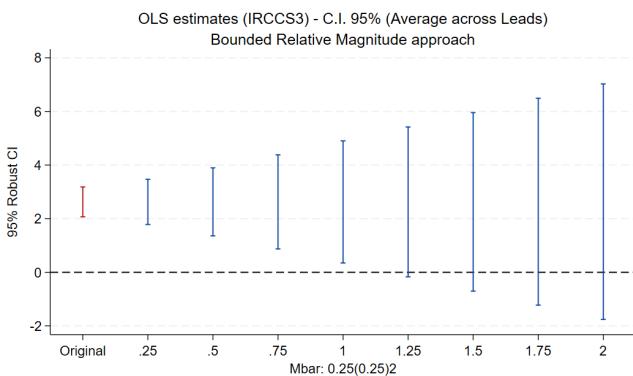
Figure A3: Honest DiD robust confidence sets in the comparison between double-treated and pure control units, estimated the overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



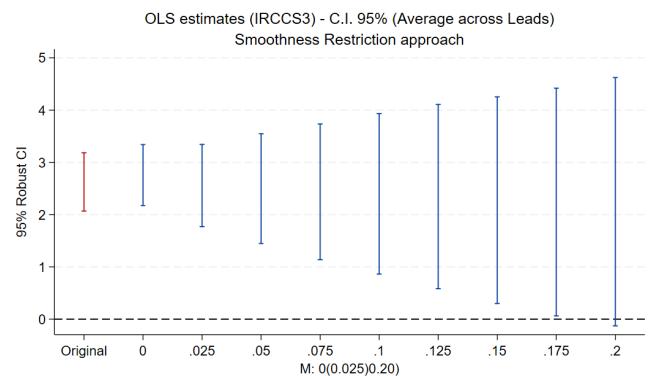
a) BM approach for the significance of the first lead.



b) SR approach for the significance of the first lead.

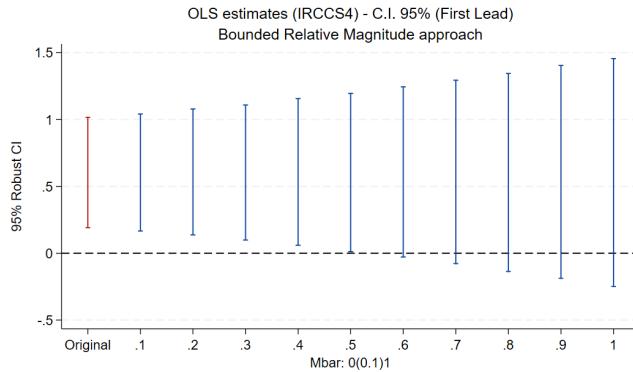


c) BM approach for the significance of the average across all leads.

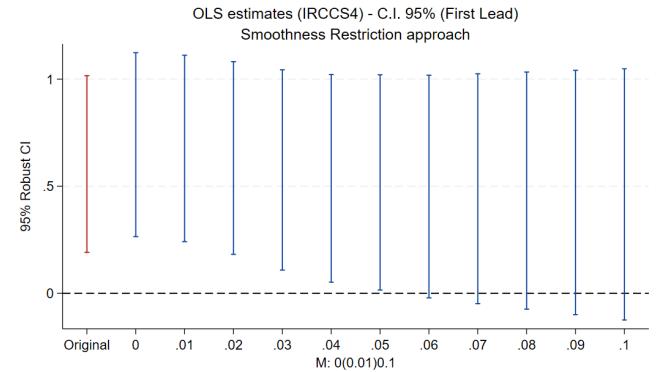


d) SR approach for the significance of the average across all leads.

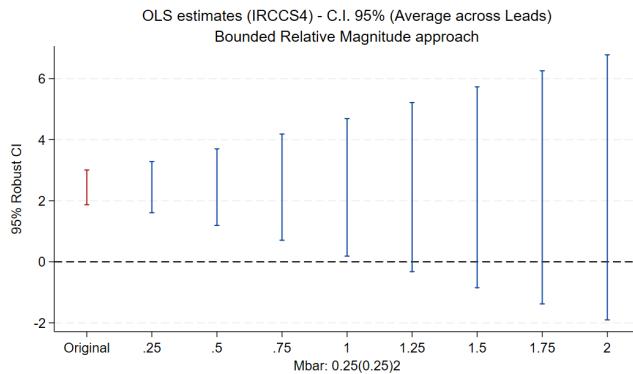
Figure A4: Honest DiD robust confidence sets in the comparison between IRCCS-only and MBO-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



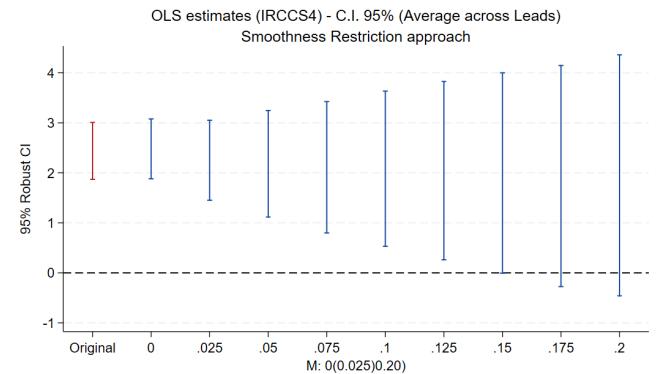
a) BM approach for the significance of the first lead.



b) SR approach for the significance of the first lead.

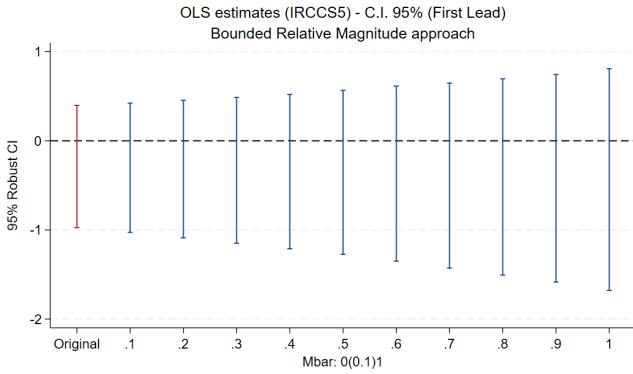


c) BM approach for the significance of the average across all leads.

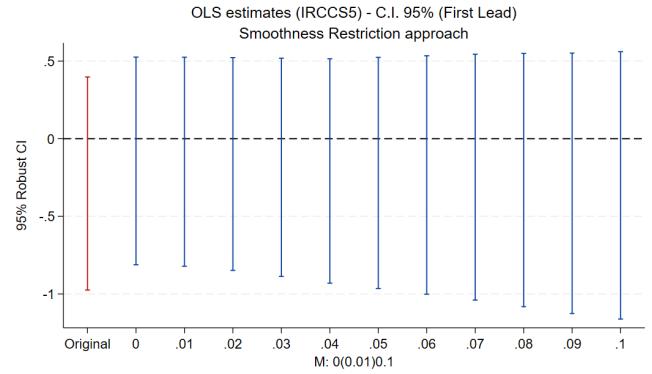


d) SR approach for the significance of the average across all leads.

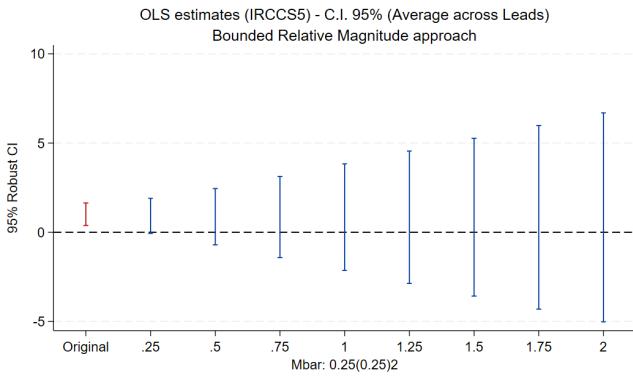
Figure A5: Honest DiD robust confidence sets in the comparison between IRCCS-only and pure control units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



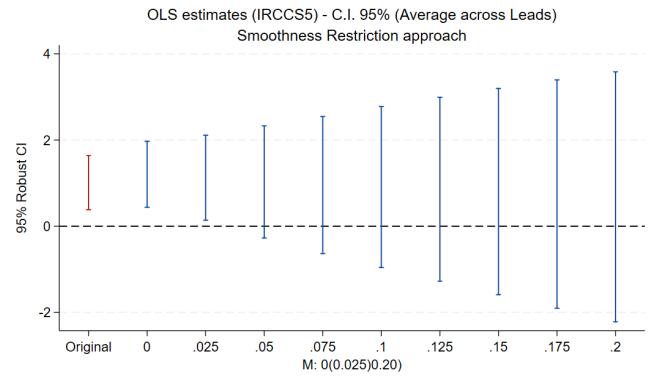
a) BM approach for the significance of the first lead.



b) SR approach for the significance of the first lead.



c) BM approach for the significance of the average across all leads.



d) SR approach for the significance of the average across all leads.

Figure A6: Honest DiD robust confidence sets in the comparison between IRCCS-only and double-treated units, estimated the overall IRCCS effect by adopting different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).

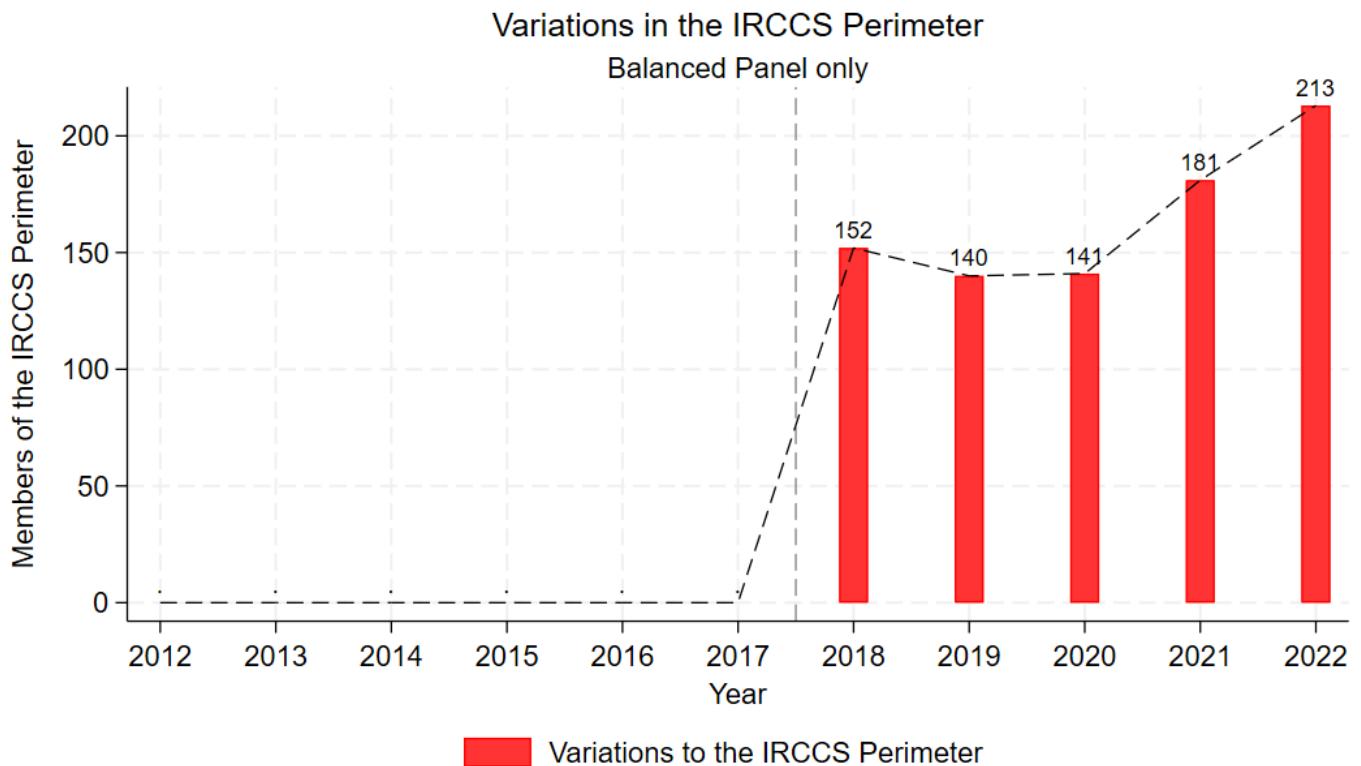


Figure A7: Evolution of the units “lately selected” into the IRCCS perimeter.

	2012-2019; SEs clustered at individual level						
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications	(7) Publications
Level 1 (MBO)	-0.20359						
Level 2 (IRCCS)	[0.13061] 3.56622 ***						
Level 3 (DT)	[0.49982] 1.98599 ***						
DT (vs. MBO)	[0.63832]	1.89804 ***					
DT (vs. IRCCS)		[0.66797]	-1.08438				
DT (vs. PC)			[0.76469]	1.81836 ***			
IRCCS (vs. MBO)				[0.67145]	3.79011 ***		
IRCCS (vs. PC)					[0.51433]	3.67938 ***	
IRCCS (vs. DT)						[0.51849]	1.08438
							[0.76469]
Observations	4,961	2,096	1,287	1,914	3,043	2,861	1,287
R-squared	0.79994	0.61952	0.77160	0.60170	0.81605	0.80583	0.77160
Individual FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No late inclusion	No late inclusion	No late inclusion	No late inclusion	No late inclusion	No late inclusion	No late inclusion
Mean	2.030	2.990	2.990	2.990	4.880	4.880	4.880

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A2: Impact of IRCCS recognition on Annual Publications in the sample obtained after excluding units later included in the IRCCS perimeter (Difference-in-Differences).

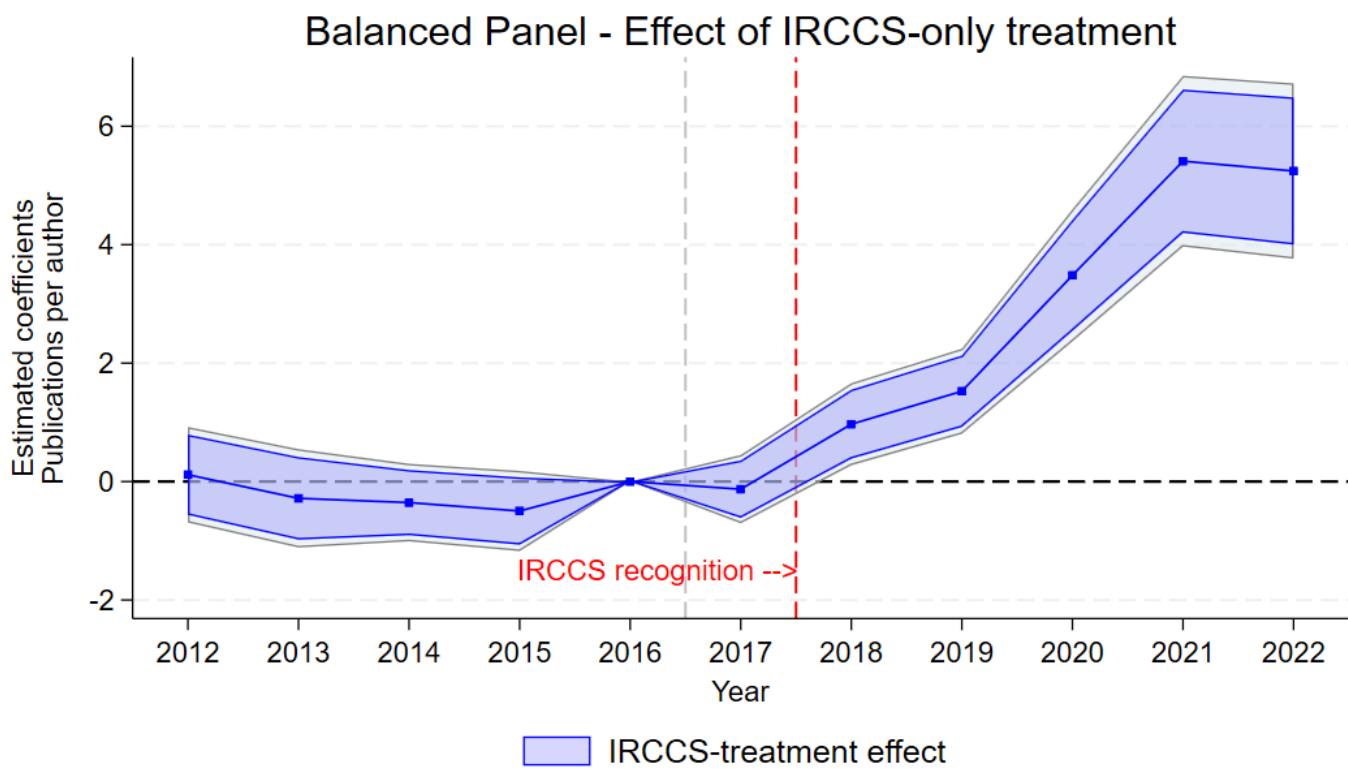
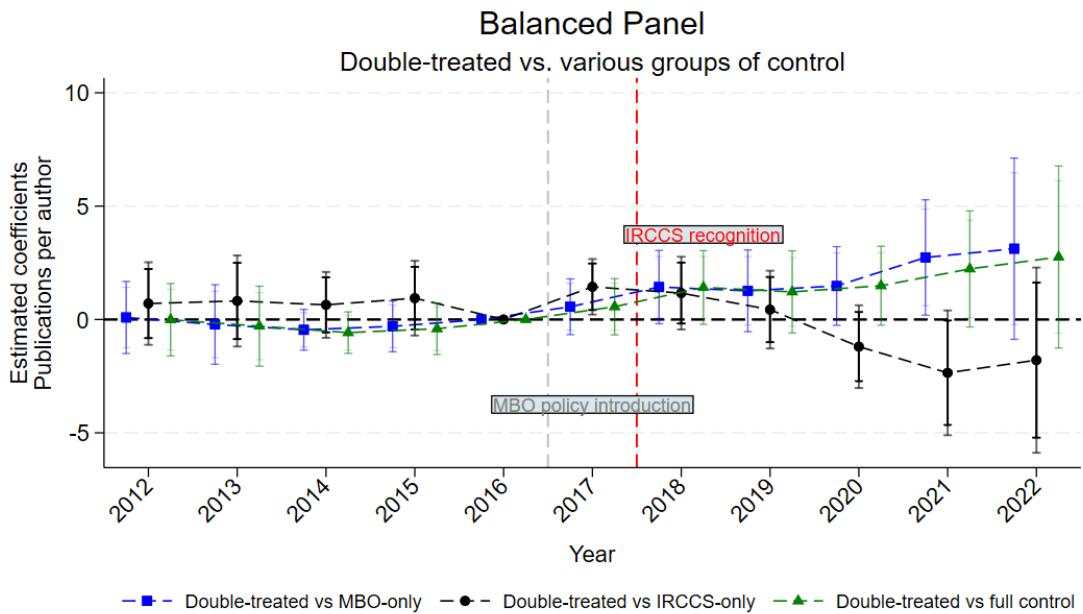
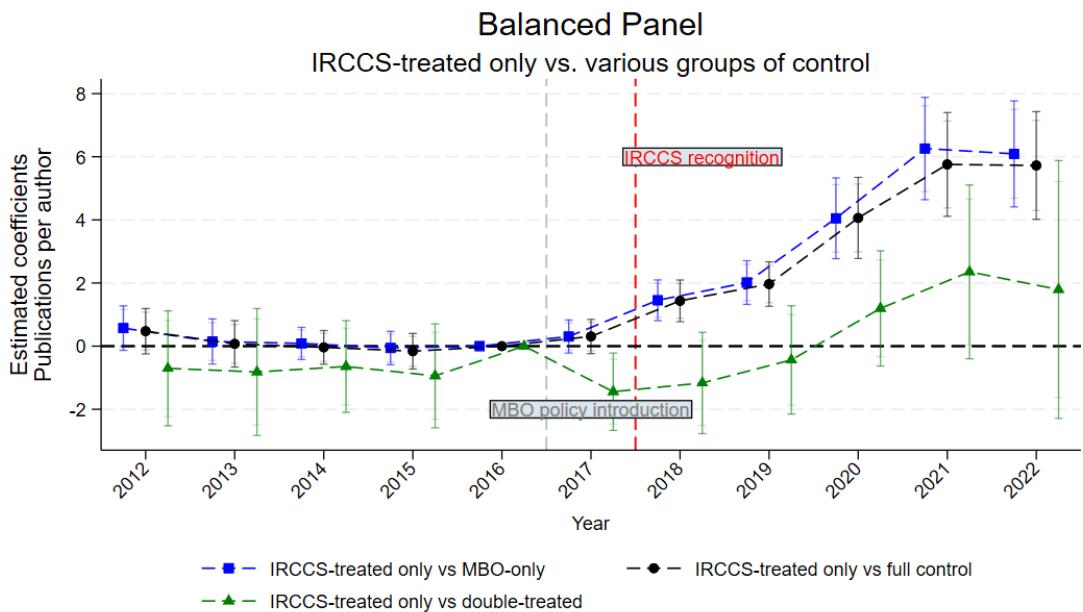


Figure A8: Event-study of IRCCS recognition effect on publications.



a) Event-study of IRCCS policy effect on publications of the double-treated units compared to the various control groups.



b) Event-study of IRCCS policy effect on publications of the IRCCS-only treated units compared to the various control groups.

Figure A9: Event-study of IRCCS policy effect on publications for comparisons across different groups. Units included in the IRCCS perimeter in any year after 2018 are excluded from the estimates.

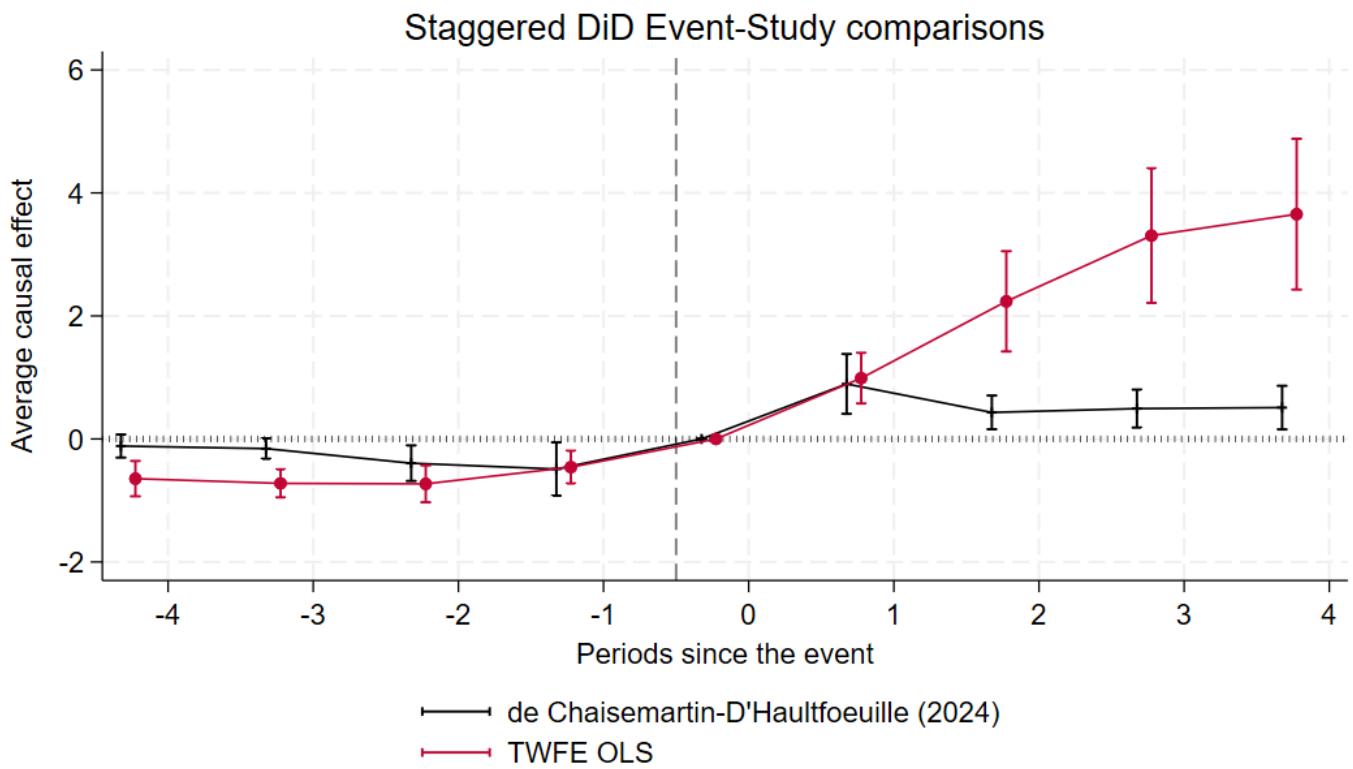


Figure A10: Staggered DiD estimates performed by accounting for the late inclusion of some units into the IRCCS perimeter

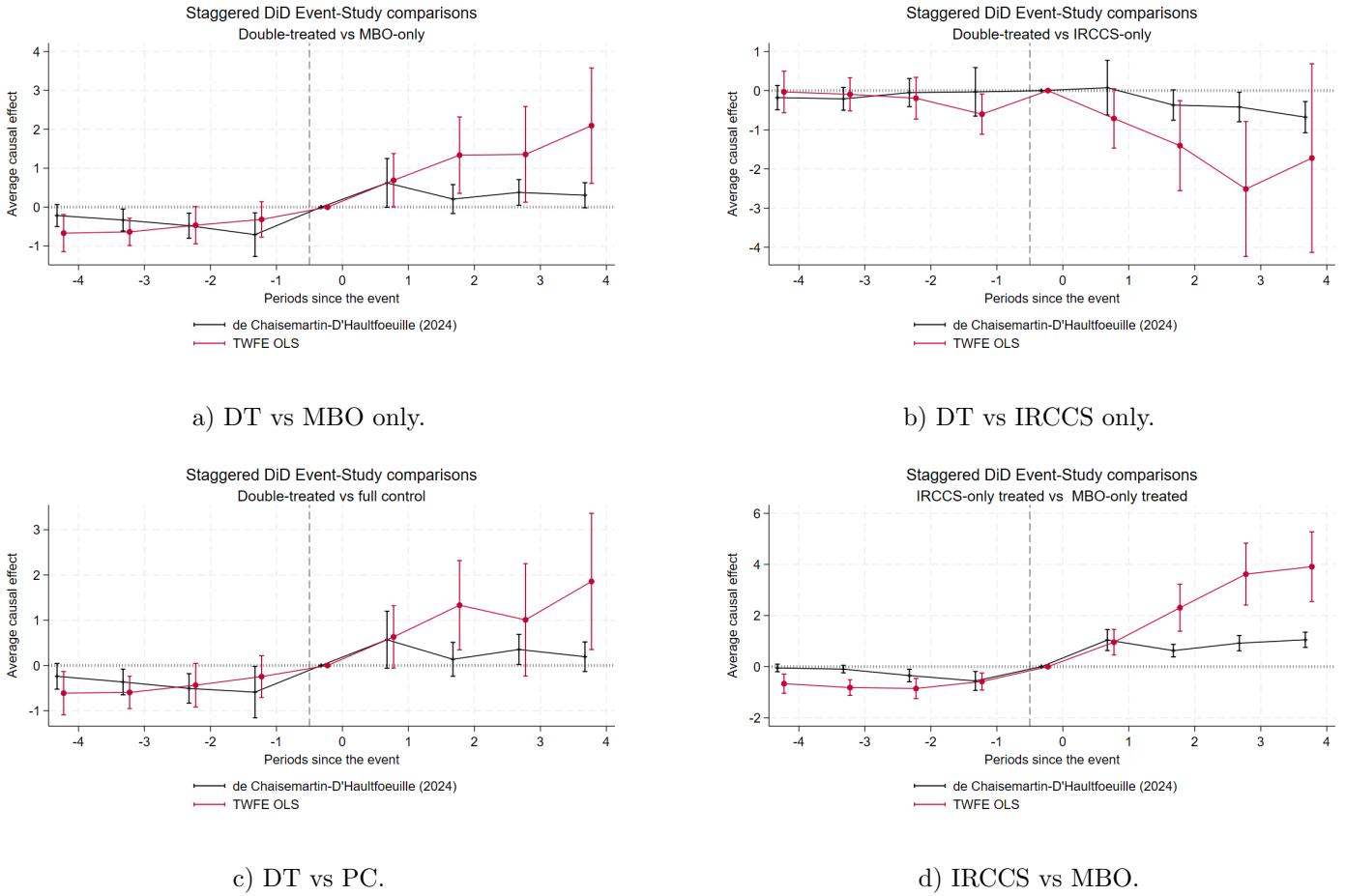
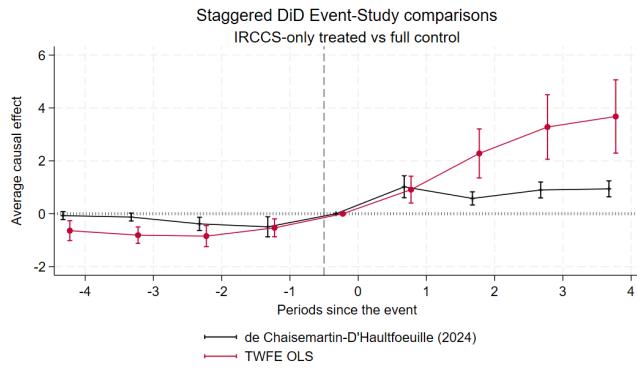
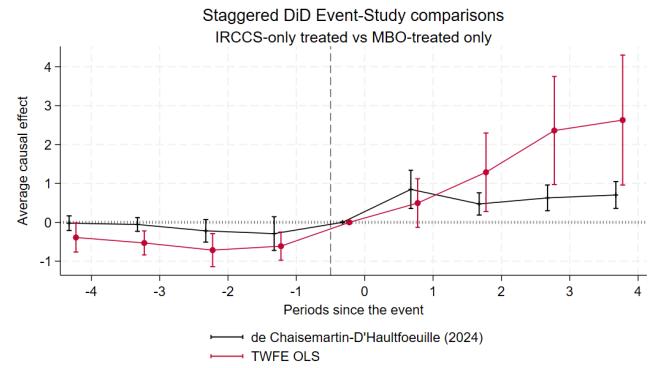


Figure A11: Staggered DiD estimates performed with different comparison groups



a) IRCCS vs PC.



b) IRCCS vs DT.

Figure A12: Staggered DiD estimates performed with different comparison groups (II)

	2012-2019; SEs clustered at individual level					
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications
Level 1 (IRCCS)	0.71621 *** [0.15851]					
Level 2 (MBO)	0.05909					
Level 3 (DT)	[0.06977] 0.58542 ** [0.23124]					
DT (vs. IRCCS)		-0.29912				
DT (vs. MBO)		[0.27422]	0.64915 *** [0.22714]			
DT (vs. PC)			0.68242 *** [0.22994]			
MBO (vs. IRCCS)				-0.94827 *** [0.16669]		
MBO (vs. PC)					0.03328	
						[0.07400]
Observations	4,632	1,960	1,920	1,776	2,856	2,672
R-squared	0.84352	0.83332	0.66112	0.62149	0.86498	0.46176
Individual FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2019	2012-2019	2012-2019	2012-2019	2012-2019	2012-2019
Panel	Full	Full	Full	Full	Full	Full
Mean	1.916	1.600	1.600	1.600	0.309	0.309

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A3: Impact of MBO Incentive on Annual Publications in the sample truncated to exclude Covid years (Difference-in-Differences).

	2012-2019; SEs clustered at individual level							
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications	(7) Publications	
Level 1 (MBO)	0.03328							
	[0.07392]							
Level 2 (IRCCS)	0.98154							
	***							
	[0.17022]							
Level 3 (DT)	0.68242							
	***							
	[0.22935]							
DT (vs. MBO)		0.64915						
		***						
		[0.22714]						
DT (vs. IRCCS)			-0.29912					
			[0.27422]					
DT (vs. PC)				0.68242				
				***				
				[0.22994]				
IRCCS (vs. MBO)					0.94827			
					***			
					[0.16669]			
IRCCS (vs. PC)						0.98154		
						***		
						[0.17039]		
IRCCS (vs. DT)							0.29912	
							[0.27422]	
hline	Observations	4,632	1,920	1,960	1,776	2,856	2,712	1,960
R-squared	0.84462	0.66112	0.83332	0.62149	0.86498	0.85431	0.83332	
Individual FE	YES	YES	YES	YES	YES	YES	YES	
Year FE	YES	YES	YES	YES	YES	YES	YES	
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	
Time Range	2012-2019	2012-2019	2012-2019	2012-2019	2012-2019	2012-2019	2012-2019	
Panel	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	
Mean	1.934	1.643	1.643	1.643	3.609	3.609	3.609	

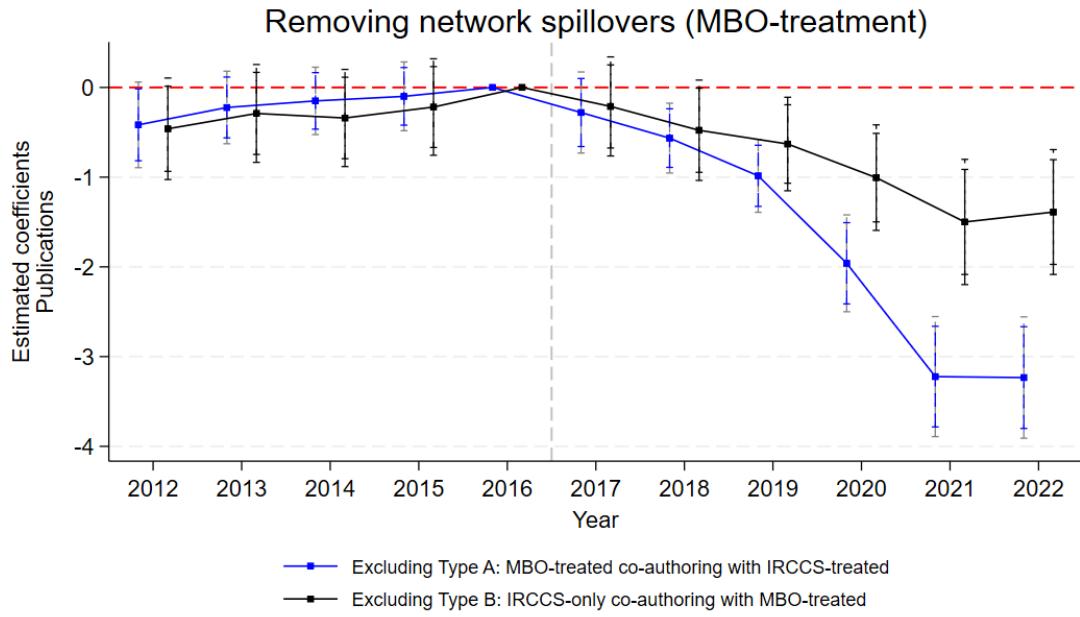
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A4: Impact of IRCCS recognition on Annual Publications in the sample truncated to exclude Covid years.

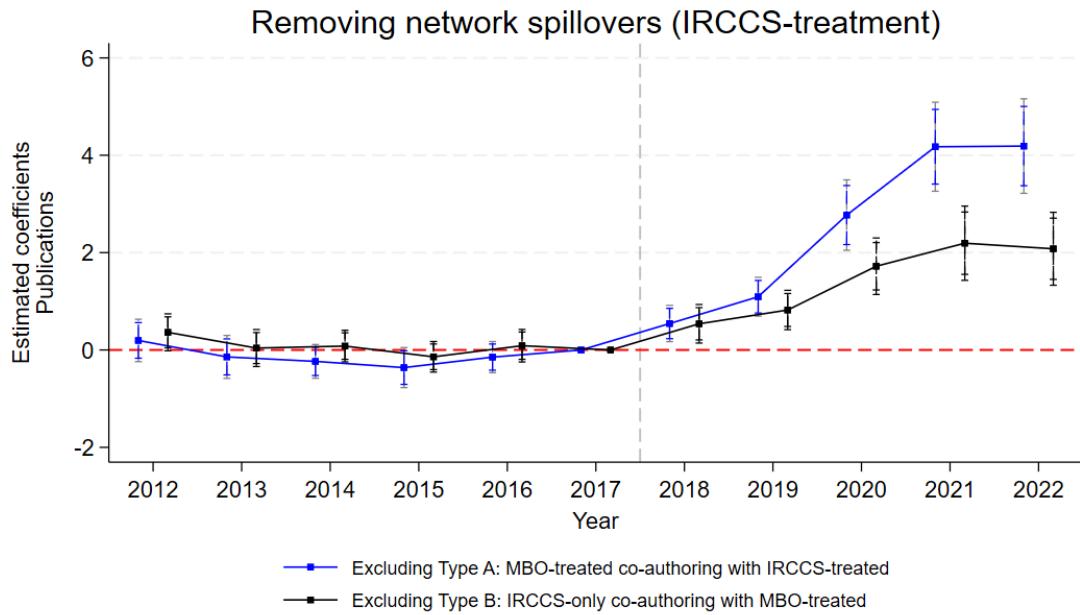
	2012-2022; SEs clustered at individual level			
	(1) Publications	(2) Publications	(3) Publications	(4) Publications
Post 2017*MBO	-1.46107 *** [0.16033]	-0.59558 *** [0.11642]		
Post 2018*IRCCS			2.65333 *** [0.29579]	1.30822 *** [0.17934]
Observations	5,410	5,176	5,410	5,176
R-squared	0.77285	0.69057	0.78582	0.70015
Individual FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No Type A	No Type B	No Type A	No Type B
Mean	0.0800	0.309	3.367	2.573

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A5: Impact of MBO-implementation and IRCCS recognition on Annual Publications in the sub-samples obtained by excluding collaborating units.



a) Event-study of the MBO policy effect on publications once excluded double-treated units and *Type A* researchers (MBO-treated units who *ever* co-authored with IRCCS researchers) in the first case, and double-treated units and *Type B* researchers (IRCCS-treated units who *ever* co-authored with non-academic physicians) in the second case.



a) Event-study of the IRCCS recognition effect on publications once excluded double-treated units and *Type A* researchers (MBO-treated units who *ever* co-authored with IRCCS researchers) in the first case, and double-treated units and *Type B* researchers (IRCCS-treated units who *ever* co-authored with non-academic physicians) in the second case.

Figure A13: Event-study of for the dynamic effects of the MBO policy (a) and the IRCCS recognition (b) effect on publications in different samples subset based on collaboration dynamics. Double-treated units are always excluded.

	2012-2022; SEs clustered at individual level					
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications
Post 2017*MBO	-0.99410 *** [0.18123]	-0.97558 *** [0.17501]	-0.98844 *** [0.18214]			
Post 2018*IRCCS				2.45143 *** [0.31386]	2.36095 *** [0.29975]	2.45662 *** [0.31666]
Observations	6,182	6,336	6,149	6,182	6,336	6,149
R-squared	0.76641	0.76227	0.76607	0.77908	0.77413	0.77879
Individual FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No switchers	No switchers	Without switchers	No switchers	No switchers	Without switchers
Panel	out of MBO	into MBO		out of MBO	into MBO	
Mean	0.547	0.653	0.547	3.609	3.591	3.591

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A6: Impact of MBO-implementation and IRCCS recognition on Annual Publications in the sub-samples obtained by excluding those who switch across groups.

	2012-2022; SEs clustered at individual level					
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications
Level 1 (IRCCS)	1.46589 *** [0.15798]					
Level 2 (MBO)	-0.05252					
Level 3 (DT)	[0.09045] 0.90403 *** [0.20576]					
DT (vs. IRCCS)		-0.60300 ** [0.25095]				
DT (vs. MBO)			1.12117 *** [0.21357]			
DT (vs. PC)				1.03322 *** [0.21886]		
MBO (vs. IRCCS)					-1.82093 *** [0.17815]	
MBO (vs. PC)						-0.10680 [0.10009]
Observations	6,116	2,464	2,573	2,385	3,724	3,652
R-squared	0.68023	0.63565	0.59311	0.55226	0.71289	0.45260
Individual FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No 99th pct	No 99th pct	No 99th pct	No 99th pct	No 99th pct	No 99th pct
Mean	1.451	1.529	1.529	1.529	0.303	0.303

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A7: Impact of MBO-implementation on Annual Publications in the sub-sample obtained by excluding authors whose number of annual publication figures at least once in the top 1% of the yearly publication distribution.

	2012-2022; SEs clustered at individual level					
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications
Levels 1 (IRCCS)	0.48534 *** [0.09321]					
Levels 2 (MBO)	0.04138					
Levels 3 (DT)	[0.06098] 0.56069 *** [0.12934]					
DT (vs. IRCCS)	0.12593					
DT (vs. MBO)	[0.16660]	0.58666 *** [0.14259]				
DT (vs. PC)		0.63605 *** [0.14441]				
MBO (vs. IRCCS)			-0.52654 *** [0.10861]			
MBO (vs. PC)				0.03384		
					[0.06237]	
Observations	4,290	924	2,208	1,889	2,398	3,366
R-squared	0.46219	0.44792	0.42087	0.43776	0.48700	0.38833
Individual FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No 10th dc	No 10th dc	No 10th dc	No 10th dc	No 10th dc	No 10th dc
Mean	0.489	0.788	0.788	0.788	0.269	0.269

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A8: Impact of MBO-implementation on Annual Publications in the sub-sample obtained by excluding authors whose number of annual publication figures at least once in the top 10% of the yearly publication distribution.

	2012-2022; SEs clustered at individual level					
	(1) Publications	(2) Publications	(3) Publications	(4) Publications	(5) Publications	(6) Publications
Level 1 (IRCCS)	0.23605 ** [0.10623]					
Level 2 (MBO)	0.01065					
Level 3 (DT)	[0.05342] 0.41681 *** [0.11213]					
DT (vs. IRCCS)	0.18883					
DT (vs. MBO)	[0.17036]	0.42702 *** [0.13893]				
DT (vs. PC)		0.43545 *** [0.14095]				
MBO (vs. IRCCS)			-0.27243 ** [0.12424]			
MBO (vs. PC)				0.00116		
					[0.05390]	
Observations	3,476	462	1,935	1,527	1,948	3,014
R-squared	0.36912	0.43849	0.35567	0.35567	0.38338	0.32899
Individual FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No 4th qt	No 4th qt	No 4th qt	No 4th qt	No 4th qt	No 4th qt
Mean	0.293	0.514	0.514	0.514	0.240	0.240

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A9: Impact of MBO-implementation on Annual Publications in the sub-sample obtained by excluding authors whose number of annual publication figures at least once in the top 25% of the yearly publication distribution.

	2012-2022; SEs clustered at individual level						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Publications	Publications	Publications	Publications	Publications	Publications	Publications
Level 1 (MBO)	-0.09982						
	[0.09994]						
Level 2 (IRCCS)	1.73328						
	***						
	[0.17903]						
Level 3 (DT)	1.04590						
	***						
	[0.21457]						
DT (vs. MBO)		1.12117					
		***					
		[0.21357]					
DT (vs. IRCCS)			-0.60300				
			**				
			[0.25095]				
DT (vs. PC)				1.03322			
				***			
				[0.21886]			
IRCCS (vs. MBO)					1.82093		
					***		
					[0.17815]		
IRCCS (vs. PC)						1.72533	
						***	
						[0.18492]	
IRCCS (vs. DT)							0.60300
							**
							[0.25095]
Observations	6,116	2,573	2,464	2,385	3,724	3,536	2,464
R-squared	0.68708	0.59311	0.63565	0.55226	0.71289	0.68649	0.63565
Individual FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No 99th pct	No 99th pct	No 99th pct	No 99th pct	No 99th pct	No 99th pct	No 99th pct
Mean	1.451	1.529	1.529	1.529	2.662	2.662	2.662

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A10: Impact of IRCCS recognition on Annual Publications in the sub-sample obtained by excluding authors whose number of annual publication figures at least once in the top 1% of the yearly publication distribution.

	2012-2022; SEs clustered at individual level						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Publications	Publications	Publications	Publications	Publications	Publications	Publications
Level 1 (MBO)	0.03580						
	[0.06219]						
Level 2 (IRCCS)	0.54436						
	***						
	[0.10734]						
Level 3 (DT)	0.63640						
	***						
	[0.14284]						
DT (vs. MBO)		0.58666					
		***					
		[0.14259]					
DT (vs. IRCCS)			0.12593				
			[0.16660]				
DT (vs. PC)				0.63605			
				***			
				[0.14441]			
IRCCS (vs. MBO)					0.52654		
					***		
					[0.10861]		
IRCCS (vs. PC)						0.57684	
						***	
						[0.11081]	
IRCCS (vs. DT)							-0.12593
							[0.16660]
Observations	4,290	2,208	924	1,889	2,398	2,079	924
R-squared	0.46499	0.42087	0.44792	0.43776	0.48700	0.48086	0.44792
Individual FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No 10th dc	No 10th dc	No 10th dc	No 10th dc	No 10th dc	No 10th dc	No 10th dc
Mean	0.489	0.788	0.788	0.788	1.024	1.024	1.024

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A11: Impact of IRCCS recognition on Annual Publications in the sub-sample obtained by excluding authors whose number of annual publication figures at least once in the top 10% of the yearly publication distribution.

	2012-2022; SEs clustered at individual level						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Publications	Publications	Publications	Publications	Publications	Publications	Publications
Level 1 (MBO)	0.00163						
	[0.05383]						
Level 2 (IRCCS)	0.27090						
	**						
	[0.11728]						
Level 3 (DT)	0.43137						
	***						
	[0.13869]						
DT (vs. MBO)		0.42702					
		***					
		[0.13893]					
DT (vs. IRCCS)			0.18883				
			[0.17036]				
DT (vs. PC)				0.43545			
				***			
				[0.14095]			
IRCCS (vs. MBO)					0.27243		
					**		
					[0.12424]		
IRCCS (vs. PC)						0.28409	
						**	
						[0.12608]	
IRCCS (vs. DT)							-0.18883
							[0.17036]
Observations	3,476	1,935	462	1,527	1,948	1,540	462
R-squared	0.37000	0.35567	0.43849	0.35567	0.38338	0.38755	0.43849
Individual FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Time Range	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022	2012-2022
Panel	No 4th qt	No 4th qt	No 4th qt	No 4th qt	No 4th qt	No 4th qt	No 4th qt
Mean	0.293	0.514	0.514	0.514	0.467	0.467	0.467

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A12: Impact of IRCCS recognition on Annual Publications in the sub-sample obtained by excluding authors whose number of annual publication figures at least once in the top 25% of the yearly publication distribution.

## References

- Acemoglu, D., & Restrepo, P. (2017). Secular Stagnation? The Effect of Aging on Economic Growth in the Age of Automation. *American Economic Review*, 107(5), 174–179.
- Ahrendt, D., Consolini, M., Mascherini, M., & Sàndor, E. (2022). *Fifth round of the living, working and covid-19 e-survey: Living in a new era of uncertainty*. (tech. rep.). Eurofound - Living conditions and family life.
- Aksoy, Y., Basso, H. S., Smith, R. P., & Grasl, T. (2019). Demographic Structure and Macroeconomic Trends. *American Economic Journal: Macroeconomics*, 11(1), 193–222.
- AMS. (2020). *Transforming health through innovation: Integrating the nhs and academia* (tech. rep.). The Academy of Medical Sciences.
- Azoulay, P., Graff Zivin, J. S., & Manso, G. (2011). Incentives and creativity: Evidence from the academic life sciences. *The RAND Journal of Economics*, 42(3), 527–554.
- Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar Extinction. *The Quarterly Journal of Economics*, 125(2), 549–589.
- Babina, T., He, A. X., Howell, S. T., Perlman, E. R., & Staudt, J. (2023). Cutting the Innovation Engine: How Federal Funding Shocks Affect University Patenting, Entrepreneurship, and Publications\*. *The Quarterly Journal of Economics*, 138(2), 895–954.
- Banal-Estañol, A., Jofre-Bonet, M., Iori, G., Maynou, L., Tumminello, M., & Vassallo, P. (2023). Performance-based research funding: Evidence from the largest natural experiment worldwide. *Research Policy*, 52(6), 104780.
- Barrenho, E., Gautier, E., Miraldo, M., Propper, C., & Rose, C. (2025). Innovation Diffusion Among Coworkers: Evidence from Senior Doctors. *Management Science*.
- Barrenho, E., Miraldo, M., Propper, C., & Walsh, B. (2021). The importance of surgeons and their peers in adoption and diffusion of innovation: An observational study of laparoscopic colectomy adoption and diffusion in England. *Social Science & Medicine*, 272, 113715.
- Baruffaldi, S. H., Marino, M., & Visentin, F. (2020). Money to move: The effect on researchers of an international mobility grant. *Research Policy*, 49(8).
- Benavente, J. M., Crespi, G., Figal Garone, L., & Maffioli, A. (2012). The impact of national research funds: A regression discontinuity approach to the Chilean FONDECYT. *Research Policy*, 41(8), 1461–1475.

- Bertoli, P., & Grembi, V. (2019). Malpractice risk and medical treatment selection. *Journal of Public Economics*, 174, 22–35.
- Bloom, D. E., Kuhn, M., & Prettner, K. (2022). Modern Infectious Diseases: Macroeconomic Impacts and Policy Responses. *Journal of Economic Literature*, 60(1), 85–131  
Cited By :75.
- Bloom, D. E., Chen, S., Kuhn, M., McGovern, M. E., Oxley, L., & Prettner, K. (2020). The economic burden of chronic diseases: Estimates and projections for China, Japan, and South Korea. *The Journal of the Economics of Ageing*, 17, 100163.
- Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does Management Matter in schools? *The Economic Journal*, 125(584), 647–674.
- Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2020). Healthy Business? Managerial Education and Management in Health Care. *The Review of Economics and Statistics*, 102(3), 506–517.
- Bloom, N., Propper, C., Seiler, S., & Van Reenen, J. (2015). The Impact of Competition on Management Quality: Evidence from Public Hospitals. *The Review of Economic Studies*, 82(2), 457–489.
- Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting Event-Study Designs: Robust and Efficient Estimation. *The Review of Economic Studies*, rdae007.
- Bosquet, C., Combes, P.-P., Henry, E., & Mayer, T. (2022). Peer Effects in Academic Research: Senders and Receivers. *The Economic Journal*, 132(648), 2644–2673.
- Brogaard, J., Engelberg, J., & Parsons, C. A. (2014). Networks and productivity: Causal evidence from editor rotations. *Journal of Financial Economics*, 111(1), 251–270.
- Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., Kokot, J., & Wiesen, D. (2024). A new look at physicians' responses to financial incentives: Quality of care, practice characteristics, and motivations. *Journal of Health Economics*, 94, 102862.
- Burgess, S., Propper, C., Ratto, M., & Tominey, E. (2017). Incentives in the Public Sector: Evidence from a Government Agency. *The Economic Journal*, 127(605), F117–F141.
- Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- Chandra, A., & Skinner, J. (2012). Technology Growth and Expenditure Growth in Health Care. *Journal of Economic Literature*, 50(3), 645–680.
- Chen, S., Cao, Z., Nandi, A., Counts, N., Jiao, L., Prettner, K., Kuhn, M., Seligman, B., Tortorice, D., Vigo, D., Wang, C., & Bloom, D. (2024). The global macroeconomic burden of Alzheimer's

disease and other dementias: Estimates and projections for 152 countries or territories. *The Lancet Global Health*, 12(9), e1534–e1543

Cited By :18.

Chen, S., Prettner, K., Kuhn, M., & Bloom, D. E. (2021). The economic burden of COVID-19 in the United States: Estimates and projections under an infection-based herd immunity approach. *Journal of the Economics of Ageing*, 20

Cited By :34.

Chen, T.-A., Lin, M.-H., Chen, Y.-C., & Chen, T.-J. (2024). The Time from Submission to Publication in Primary Health Care Journals: A Cross-Sectional Study. *Publications*, 12(2), 13.

Colussi, T. (2018). Social Ties in Academia: A Friend Is a Treasure. *The Review of Economics and Statistics*, 100(1), 45–50.

Dal Bó, E., Finan, F., & Rossi, M. A. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service\*. *The Quarterly Journal of Economics*, 128(3), 1169–1218.

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *The American Economic Review*, 110(9), 2964–2996.

Franzoni, F., Lucifora, C., Russo, A. G., & Vigani, D. (2025). In the eye of the storm. Long-term impact of the Covid-19 pandemic on healthcare utilization in Lombardy. *Economics & Human Biology*, 101497.

Ganguli, I. (2017). Saving Soviet Science: The Impact of Grants When Government R&D Funding Disappears. *American Economic Journal: Applied Economics*, 9(2), 165–201.

Gaynor, M., Moreno-Serra, R., & Propper, C. (2012). Can competition improve outcomes in UK health care? Lessons from the past two decades. *Journal of Health Services Research & Policy*, 17 Suppl 1, 49–54.

Gaynor, M., Moreno-Serra, R., & Propper, C. (2013). Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service. *American Economic Journal: Economic Policy*, 5(4), 134–166.

Ghandour, Z., Siciliani, L., & Straume, O. R. (2022). Investment and quality competition in healthcare markets. *Journal of Health Economics*, 82, 102588.

Ghirelli, C., Havari, E., Meroni, E., & Verzillo, S. (2023). The Long-Term Causal Effects of Winning an ERC Grant. *IZA Discussion Paper No. 16108*, (JRC133001).

- Goodall, A. H. (2011). Physician-leaders and hospital performance: Is there an association? *Social Science & Medicine* (1982), 73(4), 535–539.
- Grant, J., & Buxton, M. J. (2018). Economic returns to medical research funding. *BMJ Open*, 8(9), e022131.
- Gruber, J., & Owings, M. (1996). Physician Financial Incentives and Cesarean Section Delivery. *The RAND Journal of Economics*, 27(1), 99–123.
- Hacker, K. (2024). The Burden of Chronic Disease. *Mayo Clinic Proceedings: Innovations, Quality & Outcomes*, 8(1), 112–119.
- Hadavand, A., Hamermesh, D. S., & Wilson, W. W. (2024). Publishing Economics: How Slow? Why Slow? Is Slow Productive? How to Fix Slow? *Journal of Economic Literature*, 62(1), 269–293.
- Hao, Z., Cheng, H., Bärnighausen, T., & Chen, S. (2024). The effects of parental retirement on adult children's health: Evidence from China. *Health Economics*, 33(1), 12–20  
Cited By :1.
- Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, 95(9-10), 1168–1177.
- JAMA. (2025). *For authors* (tech. rep.). JAMA Network Open.
- Kessler, D. P., & McClellan, M. B. (2000). Is Hospital Competition Socially Wasteful?\*. *The Quarterly Journal of Economics*, 115(2), 577–615.
- Longo, F., Siciliani, L., Gravelle, H., & Santos, R. (2017). Do hospitals respond to rivals' quality and efficiency? A spatial panel econometric analysis. *Health Economics*, 26(S2), 38–62.
- Mantellini, P., Battisti, F., Sassoli de Bianchi, P., Armaroli, P., Battagello, J., Giorgi Rossi, P., Giubilato, P., Senore, C., Ventura, L., Zorzi, M., & Zappa, M. (2020). *Rapporto sulla ripartenza degli screening - settembre 2020* (tech. rep.). Osservatorio Nazionale Screening.
- Maynou, L., McGuire, A., & Serra-Sastre, V. (2024). What happens when the tasks dry up? Exploring the impact of medical technology on workforce planning. *Social Science & Medicine*, 352, 117014.
- Molitor, D. (2018). The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration. *American Economic Journal: Economic Policy*, 10(1), 326–356.
- Muralidharan, K., & Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119(1), 39–77.

- Propper, C., Burgess, S., & Green, K. (2004). Does competition between hospitals improve the quality of care?: Hospital death rates and the NHS internal market. *Journal of Public Economics*, 88(7), 1247–1272.
- Rambachan, A., & Roth, J. (2023). A More Credible Approach to Parallel Trends. *Review of Economic Studies*, 90(5), 2555–2591.
- Rousseau, S., Catalano, G., & Daraio, C. (2021). Can we estimate a monetary value of scientific publications? *Research Policy*, 50(1), 104116.
- Shurtz, I. (2013). The impact of medical errors on physician behavior: Evidence from malpractice litigation. *Journal of Health Economics*, 32(2), 331–340.
- Shurtz, I. (2014). Malpractice Law, Physicians' Financial Incentives, and Medical Treatment: How Do They Interact? *The Journal of Law and Economics*, 57(1), 1–29.
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
- Tang, B., Li, Z., Hu, S., & Xiong, J. (2022). Economic Implications of Health Care Burden for Elderly Population. *Inquiry: A Journal of Medical Care Organization, Provision and Financing*, 59, 00469580221121511.
- Waldinger, F. (2010). Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany. *Journal of Political Economy*, 118(4), 787–831.
- Waltman, L., & van Eck, N. J. (2013). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Journal of Informetrics*, 7(4), 833–849.
- Wu, P., & Wang, S. (2024). Examining the Economic Value of Investing in Health. *Journal of Systems Science and Information*, 12(5), 575–589.
- Ye, X., Wang, M., Xia, Y., He, P., & Zheng, X. (2023). Direct economic burden attributable to age-related diseases in China: An econometric modelling study. *Journal of Global Health*, 13 Cited By :6.
- Yohe, G. W. (1980). Current Publication Lags in Economics Journals. *Journal of Economic Literature*, 18(3), 1050–1055.