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

Gabriele Letta[a] Luca Salmasi[b] Gilberto Turati[b]

[a] Department of Economics and Finance, Università Cattolica del Sacro Cuore, Milan

[b] Department of Economics and Finance, Università Cattolica del Sacro Cuore, Rome

Abstract

We study the impact of two policy interventions on scientific research productivity among high-skilled workers in a major private Italian hospital. The first is a performance-based monetary incentive scheme (Management-by-Objectives, MBO) aimed at non-academic physicians. The second is the hospital's recognition as a Scientific Institute for Research, Hospitalization and Healthcare (IRCCS), which enhances the access to public funding to researchers. Using detailed panel data on physicians' publications over the period 2012–2022, we implement difference-in-differences strategies to evaluate the effects of both policies. We find that monetary incentives alone do not lead to a significant increase in research output among non-academic physicians. By contrast, IRCCS recognition generates a large and persistent increase in productivity among academic researchers and individuals exposed to both policies. The increase in output is primarily driven by access to stable institutional funding rather than competitive project-based grants, and is associated with a substantial expansion of research teams. This expansion works mostly through the intensive margin, with larger and more interconnected groups of existing collaborators, and is concentrated within the academic group, with limited cross-group spillovers. Consistently, the increase in research output is broad-based across fields, with little evidence of reallocation. Finally, we examine research quality using citation-based measures and healthcare indicators. While IRCCS recognition increases total citations through higher volumes, it is linked with a drop in citations per paper. In contrast, MBO-treated physicians exhibit improvements in average citation impact. The absence of evidence on gaming suggests these patterns reflect a quantity–quality trade-off. Finally, the increase in research activity is accompanied by a substantial expansion in the volume of high-complexity surgical procedures, with no systematic evidence of changes in healthcare quality outcomes as measured by PNE indicators.

*We profoundly acknowledge Matteo Ferraro for his support to the construction of the PNE-indicator dataset, and Armando Vitolo and Antonio Giulio de Belvis for their help in framing the institutional settings. We also kindly thank Paolo Berta, Vincenzo Carrieri, Elena Cottini, Joanna Kopinska, François Lafont, Matteo Lippi Bruni, Domenico Lisi, Francesco Nicolini, Alessandro Pompa Pacchi, Giuseppe Sorrenti, Rahmanda Thaariq, Stefano Verzillo, and the participants to the UPF Student Seminar Series, the III Sapienza Ph.D. Conference at DISSE, the 4^{th} APHEC Conference, the Workshop on IPR, Science and Innovation at DEP - UCSC, the XXXVII SIEP Conference and the 5th Bank of Italy Human Capital Workshop for their helpful comments and suggestions.

Keywords: High-skilled Workers, Knowledge Production, Management-By-Objective, Public Funding, Research Productivity.
JEL Classification: I10, I23, J24, O31.

1 Introduction

How should healthcare institutions allocate scarce resources when they aim at stimulating physician to engage in research? This question is becoming increasingly relevant in light of the rising costs of healthcare systems, the demographic transition, and the tightening of public research budgets. As a matter of fact, in hospital settings, research effort may be under-provided, especially among non-academic clinicians, as it requires time, resources, infrastructure, and organizational support, while its main returns accrue to institutions and society, rather than directly to individual doctors. This might determine a potential wedge between private and organizational objectives, which raises a crucial economic question: is it more effective to explicitly reward observable research output, or to expand the institutional resources available for research production?

In this paper we provide empirical evidence on two different approaches to foster research in healthcare: performance-based monetary bonuses (i.e., feeding individual incentives *ex post*), and government funding allocations (i.e., addressing infrastructure and research capacity). We exploit the setting of a large private Italian teaching hospital, where two related institutional shocks occurred in said direction. In 2017, the hospital introduced a performance-based Management-By-Objectives (MBO) scheme, financially rewarding non-academic physicians for publishing papers. After that, in early 2018, the hospital was granted the IRCCS status (Scientific Research Hospital designation), thus allowing for additional public funding, mostly benefiting academic doctors. Exploiting this institutional combination, we attempt to compare an explicit individual monetary reward to a wide institutional infrastructural intervention within the same organization and time period.

The present analysis speaks to a specific economic counterfactual, to be framed as part of a broader two-step decision process faced by healthcare institutions. First, hospitals must decide whether to deploy scarce resources to research activities at all, rather than to alternative uses such as patient care, staffing, protected time or clinical equipment. This first question concerns the general equilibrium and the desirability level of research investment within healthcare organizations, representing a major policy topic itself. While this paper does not directly address this margin, it is a key motivation for our broader research agenda. Second, conditional on a hospital's decision to devote resources

to research, organizations need to decide how said resources should be allocated. They may indeed rely on individual incentives rewarding observable output, or instead expand institutional research capacity through funding and support. We focus on this second aspect. By comparing the effects of publication bonuses with those of institutional research funding within the same organizational context, we study how different incentive structures interact with physicians' roles and constraints, shaping the effectiveness of scientific output-based policies.

The study can be framed within the debate on the transformation of research funding. In recent years, public research budgets have faced increasing pressure, raising concerns about the sustainability of traditional funding models. In the U.S., proposed budget cuts for federal research agencies such as the NIH and NSF have triggered active discussions about the future of scientific funding (Peel et al., 2025; Garisto, 2025). At the same time, private actors, like philantropic foundations, have increasingly stepped in to sustain research activities (Glenza, 2025; Shekhtman et al., 2024). The European research environment is all but alien to these concerns. In fact, fiscal constraints and institutional reforms have triggered discussions about how to trade-off public funding, private participation, scientific excellence and independence (Highman et al., 2023; Marchandot and Morel, 2025). These dynamics reflect wider shifts in the organization and governance of scientific research.

Healthcare is a particularly salient sector in this debate. Aging populations and the rising prevalence of chronic diseases are increasing the demand for healthcare services and putting pressure on health systems worldwide (Tang et al., 2022; D. E. Bloom et al., 2020; Ye et al., 2023). At the same time, innovation and biomedical research are widely recognized as key drivers of long-run improvements in health outcomes and system efficiency (Grant and Buxton, 2018; Chandra and Skinner, 2012). Sustaining research activity within healthcare organizations therefore requires designing effective institutional environments and incentive structures that support scientific production.

At the hospital level, research-active institutions may not only improve clinical standards, depending on innovation diffusion and translationality (Barrenho et al., 2021, 2025), but also attract high-quality professionals (AMS, 2020; Maynou et al., 2024). Such institutes could also develop comparative advantages, ultimately favoring patient outcomes (N. Bloom, Propper, et al., 2015, N. Bloom et al., 2020, Ghandour et al., 2022). In academic environments, physicians and researchers are inherently motivated to publish by career-oriented motives (tenure, promotions, prestige, Checchi et al., 2021), as well as altruistic reasons or due to scientific curiosity (Rousseau et al., 2021). Yet, outside academia, such *academic incentives* are missing, and physicians are more oriented towards clinical task and other activities, thus facing weaker incentives to conduct research. While this may

avoid short-term concerns about research tasks potentially crowding out the quality of care, this also dampens the effectiveness of translating research innovations into clinical practices and immediate interlinks between academia and patient care, therefore raising the policy question of how to effectively address the topic.

Beyond contributing to the thick debate on hospital management and incentive schemes in imperfect markets, as the one of education (Muralidharan and Sundararaman, 2011, N. Bloom, Lemos, et al., 2015) and healthcare (N. Bloom, Propper, et al., 2015, N. Bloom et al., 2020, Goodall, 2011), our study aims at shedding some light on whether performance-based pay schemes can effectively stimulate research productivity among non-academic physicians in comparable settings, and to what extent may institutional public funding generate greater or diverse effects. We also try to explore potential spillovers between groups of physicians—whether incentives targeting one group indirectly affect the research output and collaborations of the other.

Our work contributes to three strands of literature. First, we relate to the economics of management and incentives in public services (Goodall, 2011, Dal Bó et al., 2013, Burgess et al., 2017), by evaluating a hospital-level policy that rewards research rather than clinical performance. As a matter of fact, there is a wide literature strand on fiscal incentives to physicians' activity, although focused on care tasks only (Gruber and Owings, 1996, Shurtz, 2013, 2014, Molitor, 2018, Bertoli and Grembi, 2019, Brosig-Koch et al., 2024). Second, we contribute to the literature on research funding and scientific productivity, assessing how public funds versus private incentives shape outputs in medical research. Such literature involves mixed findings, as some scholars display how additional funding often correlates with higher research productivity, even causally (Azoulay et al., 2011, Benavente et al., 2012, Ganguli, 2017, Baruffaldi et al., 2020, Babina et al., 2023, Ghirelli et al., 2023); by contrast, others find weak or non-existent effects of grants on individual publications and productivity (Jacob and Lefgren, 2011, Banal-Estañol et al., 2023). Third, we engage with the literature on knowledge spillovers and peer effects in publications (Azoulay et al., 2010, Colussi, 2018, Bosquet et al., 2022), exploring whether incentives applied to one group of physicians trigger collaborative responses across groups.

Previewing our results, we find that the MBO scheme does not significantly increase research productivity among non-academic physicians, whereas IRCCS recognition and the associated public funding raise the publication output of academic doctors by about 60%. The two policies are complementary: non-academic physicians respond to monetary incentives only when they also have access to institutional research funding. The increase in research output is broad-based across fields, with no evidence of systematic reallocation toward specific specialties. Instead, gains appear to reflect an overall expansion in research activity rather than shifts in the composition of output.

We show that these effects are primarily driven by stable funding for routine research activities and are accompanied by a strengthening of within-group collaboration, with no evidence of increased cross-group interactions. Finally, we document heterogeneous effects on research quality. While IRCCS recognition increases total citations through higher publication volumes, it is associated with a decline in citations per paper. In contrast, MBO-eligible physicians experience improvements in citation impact despite no increase in output, suggesting a quantity–quality trade-off rather than strategic gaming of the incentive scheme. Consistent with these findings, we observe a substantial expansion in the volume of high-complexity surgical procedures following IRCCS recognition. This increase is not accompanied by systematic changes in healthcare quality outcomes, as measured by PNE indicators.

The remainder of the paper is structured as follows. Section 2 describes the institutional framework and the two policies. Section 3 presents data sources and sample construction. Section 4 presents dynamic estimates of both schemes, while Section 5 focuses on the impact of the IRCCS recognition. Section 6 reports additional analyses, in particular heterogeneity and robustness checks. Potential mechanisms are discussed in Section 7, while Section 8 treats the impact on networks, spillovers and research quality. Section 9 concludes.

2 Institutional Framework

The study is set in a leading private hospital in Rome, Italy, affiliated with a major 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 75 of its affiliated researchers are included in the top 2% ranking of scientists globally (according to Stanford's standardized citation metrics). These characteristics make it an ideal environment for studying 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 about €1 million. The bonus also varies depending on the authorship structure of the publication: first-authors of papers get the whole computed bonus if there are not internal collaborators in the publication (100%), and 60% of it if they co-author with internal affiliates. Non-first author researchers who co-author with external colleagues only get 50% of the calculated reward, which goes down to 40% if the publication involves at least another internal affiliated member. A comprehensive scheme of the rewarding policy is presented in Table A1, in the Appendix. 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 practitioners 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 formally regulated by the law: after preparing the process over the previous years, 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 is an Italian particularity, there are several international comparable institutions that may be associated in features to IRCCS, being nationally accredited to perform research. Institutes like IHU (*Instituts hospitalo-universitaires*, France), *Universitatskliniken* (Germany), AMC (*Academic Medical Centers*, U.S.), IIS (Institutos de Investigaciòn Sanitaria, Spain) are alike to IRCCSs, albeit regulated to a lesser extent.

As IRCCS, the hospital under question 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 occasional re-evaluation by experts. The quality standards foreseen by the

government in order to obtain the IRCCS recognition are presented in Table A2.

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)[1]. This IRCCS staff list consisted mainly of academically affiliated physicians (university-employed doctors working at the hospital), but it also involved some non-academic hospital physicians who were active in research. Importantly, all academic physicians, whether included in 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 might have been based on existing characteristics, it was not a matter of direct 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 who will not be 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. Said definitions are summed up in Table 1. The empirical analysis leverages such categorizations to define treatment and control groups for each policy.

3 Data

We construct a panel dataset of the professionals working in the hospitals, combined with their research output spanning 2012 to 2022. The core personnel data come from the hospital's administrative records, which include all individuals structurally employed in a professional capacity (physicians, researchers, and other healthcare staff management)

[1]As discussed later in the paper, the IRCCS list is updated every year. However, the updating is a formal recognition of individuals already involved in the IRCCS' activities due to their productivity and research activity, 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 when we discuss the validity of our approach.

from 2017 onward.

Code	MBO	IRCCS	Physician type	Policy exposure
0	No	No	Academic physicians *not in IRCCS list*	Full control: ∗ no monetary incentives ∗ no public funding
1	Yes	No	Medical Directors [No IRCCS] - Non-academic physicians *not in IRCCS list*	∗ Eligible for monetary incentive ∗ No access to IRCCS public funding
2	No	Yes	Academic physicians *in IRCCS list*	∗ Access to IRCCS public funding ∗ Not eligible for monetary incentives
3	Yes	Yes	Medical Directors [IRCCS] - Non-academic physicians *in IRCCS list*	Exposure to both (*double-treated*): ∗ Eligible for monetary incentives ∗ Access to IRCCS public funding

Notes: The Management-By-Objectives (MBO) scheme was introduced in 2017 and provides performance-based monetary bonuses for publications to non-academic physicians only. IRCCS recognition in 2018 granted access to dedicated public research funding to physicians included in the official IRCCS research staff list, updated annually. Academic physicians are never eligible for MBO incentives by institutional design. Treatment codes are used to define difference-in-differences comparisons relative to the introduction of each policy.

Table 1: Classification of physicians by type and exposure to monetary incentives (MBO) and public research funding (IRCCS).

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. Data on medical interns, residents, PhD students, and external collaborators were not available, and they are therefore excluded by construction. 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, institutional websites, LinkedIn pages and the official website of Italian physicians (FNOMCeO), 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, in our main analyses we keep the units who

are balanced for the whole longitudinal dataset only. However, for the sake of robustness, we also constructed an unbalanced panel by collecting online information[2]. The resulting balanced panel covers 584 physicians and researchers, each observed annually from 2012 through 2022, yielding 6,424 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 (medical directors vs. faculty members, MBO vs. non-MBO, IRCCS list member vs. not, etc.) which remain fixed in the main empirical strategy, reflecting their initial category. Some individuals (3.4%) switch position in the time series under consideration (e.g., a physician became an academic or vice-versa during the period); we handle such cases in robustness checks by excluding switchers. In addition, some individuals are aggregated to the IRCCS perimeter with one or few years' lag compared to the initial timing (2018). While this would make the setting more adequate to a staggered adoption analysis, in the main estimates we considered such *ever-treated* individuals as treated from the start, due to their preemptive involvement with research activity and with the offical IRCCS-research group, thus considering them as informally treated individuals expecting formal recognition at some point. We later corroborate the validity of such design by keeping in the sample only those who receive the IRCCS-status from the beginning, and by performing some Staggered DiD estimates as well, confirming the robustness of our main estimates.

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 the authors in our panel (and also on those in the unbalanced dataset). 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

[2]Retrieving online information was necessary to construct the unbalanced panel. As a matter of fact, the attrition in the panel after 2017 was easily obtainable by observing which professionals were yearly absent or present in the official records provided by the institution. But to trace their affiliation before 2017, while for most of them the hire date was an available information, for those no longer or not yet employed in either 2020 or 2023, such variable was missing. We hence collected online CVs, LinkedIn self-reported information, or data from institutional websites to reconstruct the affiliation. In some cases, especially for the oldest doctors, no CV or online info was available. We thence looked at their profile on the official online portal of Italian physicians (*FNOMCeO*), which reports date and place of birth, and the dates and places of their medical degree, their habilitation as physicians, their registration in the local medical books, and (if they did it), of their medical residency. We assume that MDs completing their last residency in the university affiliated to the hospital under study, or getting their habilitation immediately after completing their degree in said university if they did not specialize themselves, could be considered as enrolled in the institutional ranks that same year, as for most of our attrition subjects these events were reported happening decades before the starting year of our panel. In the few cases of inconsistency (no official CV information online, no LinkedIn profiles, a degree/residency completed in another university or hospital), we sought for additional online information via scraping from Italian websites for reviewing physicians (*TopDoctors*, *IlMioDottore*, etc.) or online articles. The source material of this operation has been stored and may be made available upon request.

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 on researchers' surnames only. Each publication record provides the publication year, journal, number of co-authors, citations (as of the time of data collection, which is Fall 2024), and other bibliometric information. We match authorship information from these records to the individuals in our hospital staff panel using names, surnames, and fields of specialization, carefully verifying potential homonyms. From this matched dataset, we construct our main outcome variable, namely the number of publications per person per year. We also record the total citations received by each individual's publications in order to observe research impact. In addition, the bibliometric information allows us to construct several measures of research collaboration, including the number of co-authors per publication, the number of internal collaborators affiliated with the hospital, and the entry of new collaborators into researchers' co-authorship networks. These variables are used later in the paper to explore potential mechanisms behind the estimated effects and to perform robustness checks.

Table 2 reports the summary statistics for the selected sample. About 52% of the 584 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. In terms of treatment groups, it is clear how the units differentiated by status (IRCCS-only, MBO-only, double-treated) do not sum up to the total number of individuals in the sample: this is due to the switching units, which are taken care of in the robustness checks. 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 0.46 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 slightly higher mean annual publication number than doctors eligible for MBO only (0.85).

If we look at the times the works published in a given year were reportedly cited in Fall 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, and not belonging to the non-academic MD category, are 198 (approximately 34% of the sample). Within the IRCCS group, 69 individuals are non-academic physicians (thus eligible for MBO as well). The non-

academic non-IRCCS physicians are instead 178 (30.5% of the sample), while the pure control group is made by 162 individuals (27.7%).

	Mean	SD	Min	Max	N
Sample composition					
Non-IRCCS listed Medical Directors (MBO-treated)					178
IRCCS-listed academics (IRCCS-only)					198
IRCCS-listed Medical Directors (double-treated)					69
Non-IRCCS listed academics (pure control)					162
Total units of the balanced panel					584
Workers' statistics					
Age	51.127	7.857	31	70	
Publications	2.186	4.564	0	102	
Times cited (WoS)	57.833	204.140	0	7037	
Times cited (All outlets)	61.463	217.740	0	7509	
Female	0.420	0.494	0	1	
Male	0.580	0.494	0	1	
Medical Director (hospital only)	0.409	0.492	0	1	
Healthcare Professions' Manager	0.018	0.134	0	1	
Sanitary Director	0.050	0.218	0	1	
Faculty Member with Clinical Functions	0.523	0.500	0	1	
Publications' statistics					
Publications of non-IRCCS listed Medical Directors (MBO-treated)	0.461	1.054	0	12	
Publications of IRCCS-listed academics (IRCCS-only)	4.929	6.880	0	102	
Publications of IRCCS-listed Medical Directors (double-treated)	2.258	2.768	0	20	
Publications of non-IRCCS listed academics (pure control)	0.848	1.683	0	23	
Citations of non-IRCCS listed Medical Directors (MBO-treated)	8.537	33.919	0	1009	
Citations of IRCCS-listed academics (IRCCS-only)	148.799	356.584	0	7509	
Citations of IRCCS-listed Medical Directors (double-treated)	66.691	144.069	0	2102	
Citations of non-IRCCS listed academics (pure control)	15.540	37.684	0	485	
Observations	6424				

Table 2: Descriptives

Figure 1 plots the average number of publications per physician per year between 2012 and 2022 for the balanced panel. Additional descriptive figures showing total publications and trends under different degrees of attrition are reported in Appendix Figures A1 and A2. We observe a relatively flat trend in the years up until 2017, followed by a noticeable increase starting around 2018. This pattern is basically invariant in the comparison between the balanced panel, and the various possibilities of attrition-flawed datasets that we can build with the available information (Figure A2). The timing also aligns with the IRCCS recognition, suggesting, at a first glance, a possible aggregate effect of these policies on the overall research productivity of 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 less in the medical field, lowering to even a couple-of-months time length for systematic reviews or literature reports (Chen et al., 2024). In addition, 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).

On the other side, two 2017-published papers with two citations each as in October 2024 (if using the WoS citation metrics), are "Reverse Time-Dependent Effect Of Alphafetoprotein And Disease Control On Survival Of Patients With Barcelona Clinic Liver Cancer Stage C Hepatocellular Carcinoma" (Ponziani et al., 2017, *World Journal of Hepatology*, first received in August 2017 and published in December 2017) and "The chromosome analysis of the miscarriage tissue. Miscarried embryo/fetal crown rump length (CRL) measurement: A practical use" (D'Ippolito et al., 2017, *PLos One*, first received in March 2017 and published in June 2017). They display similar if not even shorter publication lags. In addition to that, one can even argue that the effect observed in a timely fashion with respect to the two policy implementations may reflect the fact that some *pipeline projects*, already on-going before 2017, may have been strategically accelerated or finalized in order to obtain the bonus or enabled by the IRCCS additional funds.

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.

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 made by 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.

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.

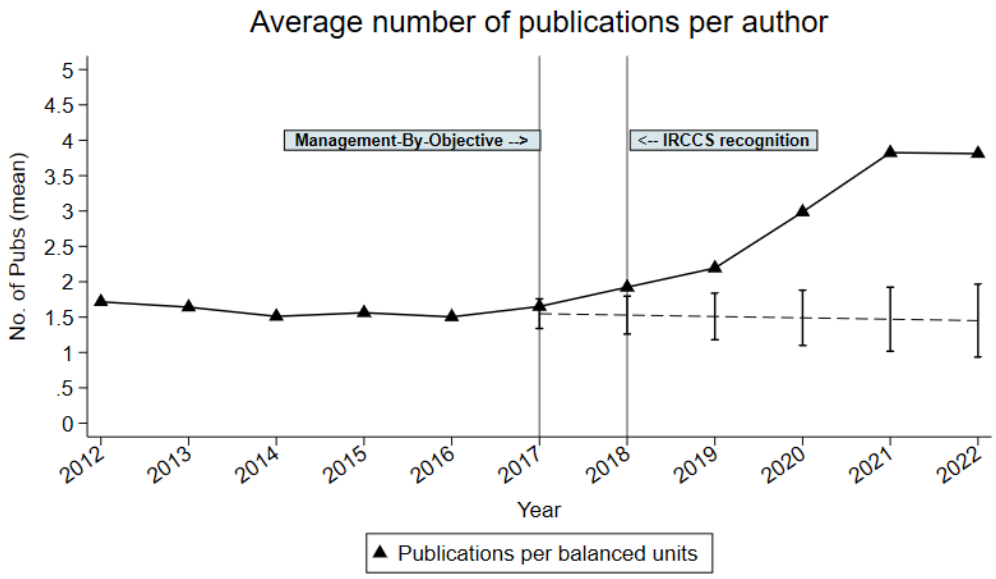


Figure 1: Average publications per researcher per year (2012–2022), balanced panel. Vertical lines indicate the introduction of the MBO policy in 2017 and the IRCCS recognition in 2018.

4 Dynamic Estimates

4.1 Empirical Strategy

Although the policies represent two separate interventions, they occur in close temporal sequence and share the same scope. Thus, we analyze them together first. To assess which policy proves to be more effective, we estimate dynamic event-studies specifications based on a standard two-way fixed effects difference-in-differences (TWFE DiD) framework. The first policy is the performance-based monetary incentive (MBO) implemented in 2017. The second one is the hospital's recognition as an IRCCS institute in early 2018, which allowed the hospital to access dedicated public research funding. The latter mainly affected academically affiliated physicians, who became able to receive additional resources and operate within a more prestigious research environment.

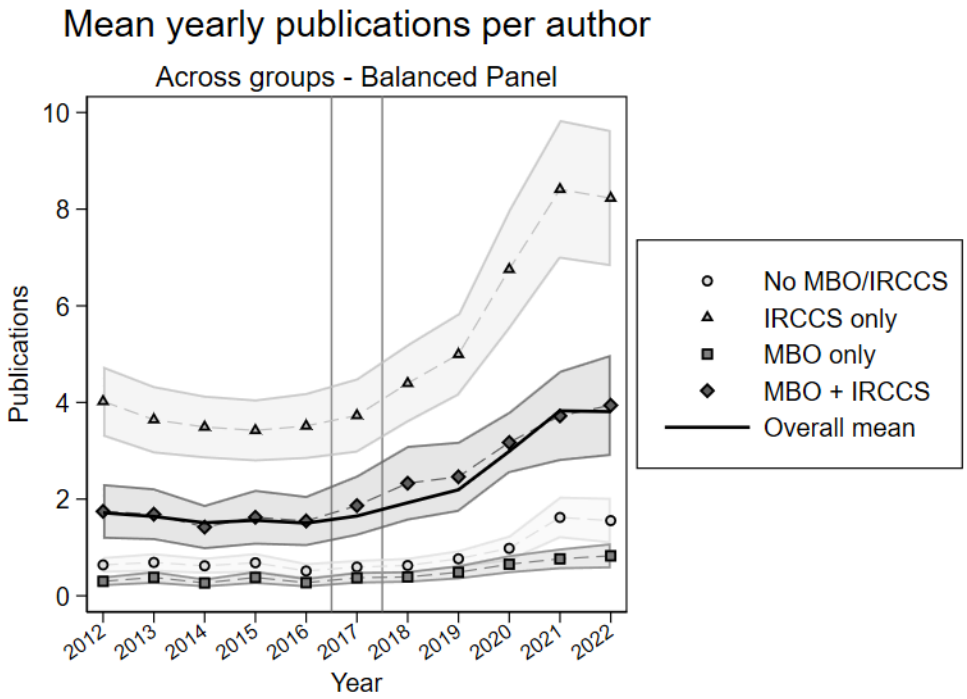


Figure 2: Average yearly publications across different groups of healthcare professionals, defined according their treatment status after 2017 and 2018.

We identify comparison groups using the classification presented in Table 1, which we combine into a unified treatment framework. The resulting groups are defined as follows: *Group 0: "Full control"*; it includes academic physicians, sanitary directors, and other research staff affiliated with the hospital or the university who are not part of the IRCCS research staff list. None of them is eligible for the MBO scheme. *Group 1: "MBO-only treated"*; these are non-academic physicians who are not included in the IRCCS research staff list. They are eligible for the monetary incentive but do not benefit from the additional public research funding. *Group 2: "IRCCS-only treated"*, which includes academic physicians (faculty members performing clinical activity) who are part of the IRCCS research staff list. They have access to IRCCS research funding but are not eligible for the MBO incentives. *Group 3: "Double-treated"*, who are the non-academic physicians tjhat are eligible for the MBO incentive and are also included in the IRCCS research staff list. They therefore benefit from both the monetary incentive and access to IRCCS research resources.

We estimate a dynamic TWFE difference-in-differences specification comparing changes in publication outcomes across these groups of physicians. The *full* control group (Group 0) serves as the reference category. Since our analysis focuses on dynamic treatment ef-

fects, we do not rely on a single post-treatment indicator as in a standard DiD framework. Instead, we estimate event-time coefficients capturing the evolution of research output before and after the introduction of the policies. We normalize the coefficients to the last pre-treatment year (2016), which serves as the omitted category.

We estimate the following equation:

$$Y_{it} = \alpha + \sum_{j \in \{1,2,3\}} \left\{ \sum_{h=2}^{H} \beta_{j,h} \, \mathbb{1}(G_i = j) \, \mathbb{1}(t = t_0 - h) + \sum_{g=1}^{K} \gamma_{j,g} \, \mathbb{1}(G_i = j) \, \mathbb{1}(t = t_0 + g) \right\} + \delta_i + \tau_t + \varepsilon_{it} \,. \quad (1)$$

where Y_{it} denotes the research output of individual i in year t. $G_i \in \{0, 1, 2, 3\}$ identifies the group to which physician i belongs, with $G_i = 0$ representing the control group. The indicators $\mathbb{1}(G_i = j)$ therefore capture appurtenance to the different treatment groups. The variables $\mathbb{1}(t = t_0 - h)$ and $\mathbb{1}(t = t_0 + g)$ are event-time indicators corresponding to pre-treatment and post-treatment periods relative to the policy implementation. Their interaction with the treatment-group indicators generates the event-study coefficients measuring differential trends for each group relative to the control group. To avoid perfect multicollinearity, we omit the year 2016, which is the last year before the introduction of the MBO policy and serves as the reference period. Individual fixed effects (δ_i) control for time-invariant differences across physicians, while year fixed effects (τ_t) absorb common shocks affecting all individuals in a given year. Standard errors are clustered at the individual level. The coefficients $\beta_{j,h}$ and $\gamma_{j,g}$ capture the differential evolution of research output for each treatment group relative to the control group over time. We present these coefficients graphically in the event-study figures in order to assess both the presence of pre-trends and the dynamic effects of the policies.

The key identification assumption is that, in the absence of the two policy interventions, the research output of treated and control groups would have followed parallel trends. The event-study specification allows us to visually inspect whether such parallel trends hold in the pre-treatment period.

4.2 Results

Figure 3 reports the dynamic event-study estimates obtained from Equation 1. The coefficients measure the evolution of publication outcomes across the different physician groups relative to the control group.

Three main patterns emerge. First, the IRCCS recognition generates a substantial and statistically significant increase in research productivity among academic physicians included in the IRCCS research staff. The effect becomes visible in 2019, when academic

physicians included in the IRCCS perimeter increase their publication output by about one paper, corresponding to an increase of approximately 28% relative to the pre-treatment mean.

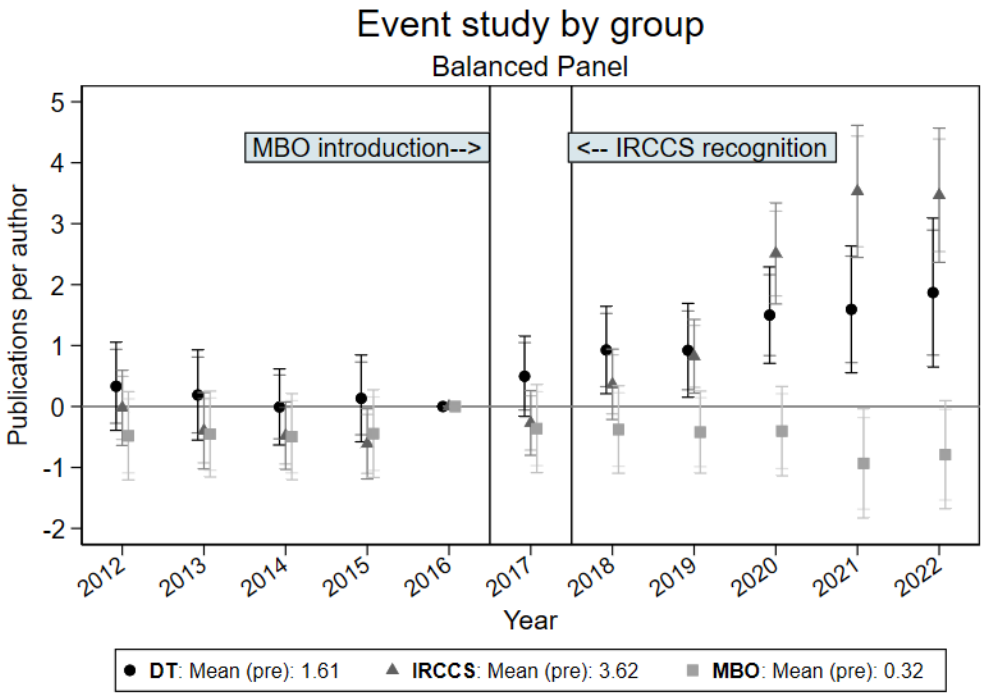


Figure 3: Event-study estimates of the effect of the MBO policy and IRCCS recognition on publications across physician groups, relative to the full control group and the pre-treatment mean of the outcome (balanced panel).

Second, the MBO incentive alone does not appear to stimulate research output. The estimates for the MBO-only group remain close to zero throughout the post-treatment period and are not statistically significant. Third, the monetary incentive appears to generate positive effects only when combined with IRCCS funding. Non-academic physicians who are both eligible for the MBO scheme and included in the IRCCS research staff (the "double-treated" group) exhibit a significant increase in publications after 2018. While baseline publication levels differ across groups, the event-study estimates show no evidence of differential trends in the pre-treatment period, supporting the parallel trends assumption.

To assess whether the results are brought about by sample selection due to the panel attrition emerged in the data collection process, Appendix Figure A3 reports the corresponding event-study estimates using the full unbalanced panel. The results remain qualitatively unchanged.

5 IRCCS Funding Analysis

Given the almost null effect of the MBO policy and the substantial increase in research output associated with IRCCS recognition, we now focus specifically on the impact of the IRCCS designation, granted in 2018.

5.1 Empirical Strategy

To estimate the causal impact of IRCCS recognition on research productivity, we employ a standard two-way fixed effects difference-in-differences specification estimated by OLS:

$$Y_{it} = \alpha + \theta(Post2018_t \times IRCCS_i) + \delta_i + \tau_t + \varepsilon_{it}, \tag{2}$$

where $Post2018_t$ equals one for years $t \geq 2018$ and zero otherwise, and $IRCCS_i$ indicates whether physician i belongs to the IRCCS research staff list. Individual fixed effects (δ_i) control for time-invariant differences across physicians, while year fixed effects (τ_t) capture common shocks affecting all individuals in a given year. In the baseline specification we adopt a conservative comparison group by including double-treated individuals, MBO-only physicians, and the pure control group among the controls. This choice provides a lower-bound estimate of the IRCCS funding effect.

Although the hospital updates the IRCCS research staff list annually, our baseline specification treats physicians who are eventually included in the IRCCS perimeter as treated from the beginning of the post-recognition period. Under this definition, the estimated coefficient should be interpreted as an intention-to-treat (ITT) effect rather than an average treatment effect on the treated (ATT). This assumption reflects the institutional organization of research activity within the hospital. After the IRCCS recognition, a relatively stable core of research-active physicians emerged and collaborated on projects financed through IRCCS funding. Some researchers were formally added to the IRCCS staff list with a delay, despite already participating in IRCCS-related research activities. We therefore model the treatment as absorbing. While some individuals are formally removed from the IRCCS list in later years, research projects and infrastructures financed through IRCCS funding typically extend over multiple years. Researchers involved in these projects are thus likely to continue benefiting from the associated resources even after their official status changes.

The coefficient θ captures the average ITT effect of IRCCS recognition for physicians who receive access to IRCCS funding after 2018, relative to MBO-eligible physicians and the full control group. The main identification assumption is that, in the absence of IRCCS

recognition, research productivity across groups would have followed parallel trends. While the inclusion of individual and year fixed effects accounts for time-invariant heterogeneity and common shocks, we assess this assumption through event-study specifications, as in the previous section.

Since inclusion in the IRCCS perimeter is determined by the institution and may partly depend on physicians' research performance, concerns about selection may arise. To address this issue, we complement our event-study estimates with robustness checks based on the methodology proposed by Rambachan and Roth, 2023. The *honestDiD* approach allows for mild violations of the parallel trends assumption by constructing confidence sets that remain valid even when pre-treatment trends differ slightly across groups.

In addition to the baseline specification, we estimate a series of alternative comparisons across groups in order to assess heterogeneous effects. In particular, we perform the following pairwise comparisons: (1) double-treated vs. MBO-only; (2) double-treated vs. full control; (3) IRCCS-only vs. MBO-only; (4) IRCCS-only vs. full control; (5) IRCCS-only vs. double-treated. These comparisons are obtained by re-estimating Equation 2 while changing the reference group in each specification.

We then estimate dynamic effects of the IRCCS intervention through event-study specifications for this set of models, interacting $IRCCS_i$ with year dummies around the 2018 recognition. The baseline comparison in Equation 3 adopts a conservative approach: the treatment indicator equals one only for physicians included in the IRCCS perimeter *exclusively*, that is, academic doctors listed in the IRCCS research staff and not eligible for the MBO scheme. The control group therefore combines groups 3, 1, and 0. This specification allows us to visualize potential diverging trends prior to 2018. The estimated equation (via OLS) is the following:

$$Y_{it} = \alpha + \sum_{h=2,\, h\neq 1}^{H} \beta_h \, \mathbb{1}(IRCCS_i) \, \mathbb{1}(t = 2018 - h) + \sum_{g=1}^{K} \gamma_g \, \mathbb{1}(IRCCS_i) \, \mathbb{1}(t = 2017 + g) + \delta_i + \tau_t + \varepsilon_{it}. \quad (3)$$

Although it is common practice to omit the period immediately preceding the treatment to avoid collinearity, we follow the same normalization adopted in the previous event-study and use 2016 as the reference period. This choice facilitates comparability across the different sets of estimates presented in the paper.

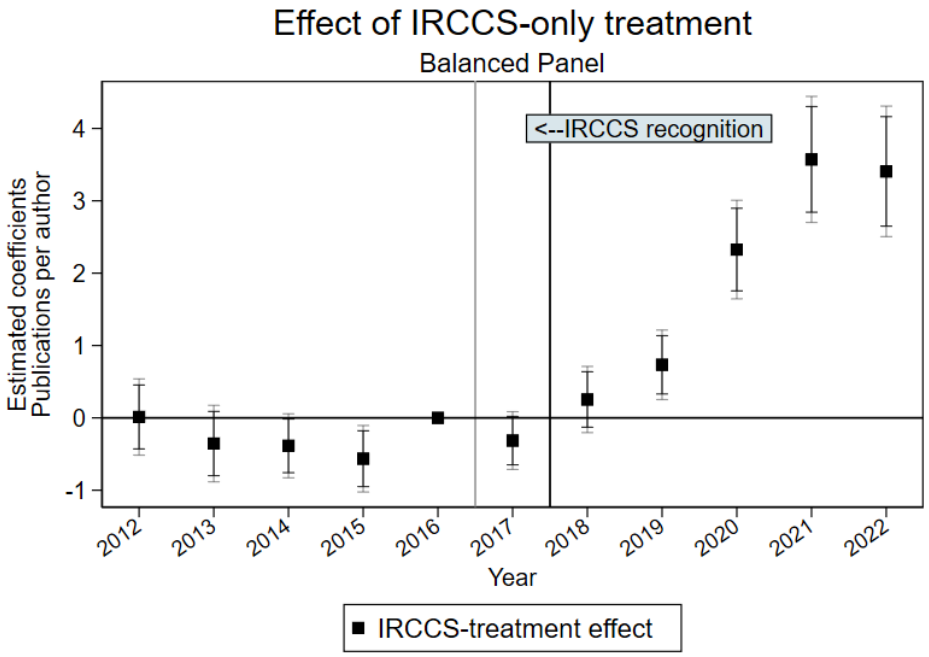
5.2 Results

Table 3 reports the difference-in-differences estimates from Equation 2. The results confirm the patterns previously observed in the event-study analysis and highlight the dominant role of IRCCS recognition relative to the performance-based incentive scheme. Column (1)

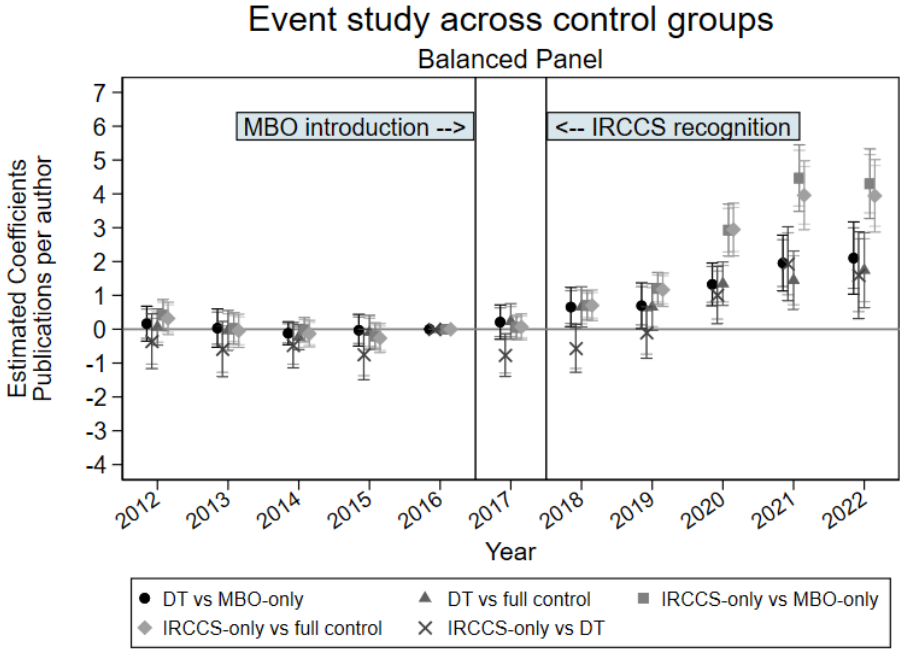
reports the baseline estimate of the interaction term $Post2018_t \times IRCCS_i$. The coefficient indicates that being listed as an IRCCS researcher is associated with an increase of about 2.34 publications per year (significant at the 1% level) relative to the other groups. This corresponds to an increase of roughly 64% relative to the mean publication level of the IRCCS group. Columns (2)–(6) report pairwise comparisons across groups. The results are consistent across specifications. Double-treated physicians publish about 1.24 more papers per year than MBO-only physicians (75%), and about 1.14 more papers than the pure control group (69%). However, IRCCS researchers publish around 1.11 more publications per year than the double treated (66%). The estimates indicate that IRCCS recognition substantially increased research productivity both for academic physicians and for the subset of non-academic doctors exposed to both policies. Nevertheless, the largest increase in publication output is observed among IRCCS-only researchers, suggesting that access to dedicated research funding plays a more important role than performance-based monetary incentives alone.

	(1) IRCCS vs. others	(2) DT vs MBO	(3) DT vs PC	(4) IRCCS vs MBO	(5) IRCCS vs PC	(6) IRCCS vs DT
Publications	2.3416***	1.2398***	1.1365***	2.6706***	2.5569***	1.1125***
(SE)	(0.2925)	(0.2285)	(0.2348)	(0.3098)	(0.3156)	(0.3347)
N	6424	2623	2430	3987	3794	2739
R^2	0.777	0.596	0.582	0.796	0.787	0.768
Time FE	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
Mean	3.64	1.65	1.65	1.65	3.64	3.64
Panel (bal.)	Full	DT, MBO	DT, control	IRCCS, MBO	IRCCS, control	IRCCS, DT
Time	2012–2022	2012–2022	2012–2022	2012–2022	2012–2022	2012–2022

Table 3: Impact of IRCCS recognition on Annual Publications (Difference-in-Differences).



a) Event-study of IRCCS recognition effect on publications for IRCCS-only treated individuals compared with all the other units.



b) Event-study of IRCCS policy effect on publications for comparisons across different groups.

Figure 4: Event-studies of IRCCS policy effect.

The baseline dynamic specification for the IRCCS effect is presented in panel (a) of Figure 4. The definition of treated units in this specification is deliberately conservative, as IRCCS-only academic physicians are compared with the rest of the hospital personnel, including individuals who receive the performance-based bonus and those who are exposed to both policies. Prior to 2018, the publication trends of IRCCS-only researchers are relatively flat and do not diverge significantly from those of the comparison groups. Although their average level of scientific output is substantially higher (as expected for professional researchers; see Table 2), the absence of differential pre-trends supports the validity of the identifying assumption.

Following the IRCCS recognition, the coefficients display a gradual increase. A small and marginally significant effect emerges in 2019, plausibly reflecting the lag inherent in the publication process. The estimated effect becomes much larger in 2020, exceeding 100% of the pre-treatment mean. This pattern is likely driven both by the maturation of research projects initiated after the policy change and by the surge in scientific activity associated with the COVID-19 pandemic. The effect continues to grow in 2021 and stabilizes in 2022.

While this baseline specification is conservative, it is not perfectly sharp in identifying the treatment effect. Some researchers included in the comparison group are non-academic physicians who receive the performance-based bonus and may also participate in IRCCS-funded activities. Although earlier results indicate that the MBO policy has only a limited effect, its presence could still introduce some bias in the baseline estimates, particularly for double-treated individuals. For this reason, we estimate additional dynamic specifications that compare IRCCS-only and double-treated individuals with each of the other groups. The corresponding event-study estimates are reported in panel (b) of Figure 4. These comparisons confirm that the performance-based incentive alone does not substantially affect publication output. The MBO scheme appears to shift upward the publication trajectory of double-treated physicians only slightly, whereas the access to IRCCS funding generates a much stronger and persistent increase in research productivity for academic researchers. Some pre-treatment differences emerge when comparing IRCCS-only researchers with double-treated physicians. In particular, IRCCS-only researchers exhibit a slightly negative trend prior to 2017 when compared to the double-treated, which likely reflects the selection of particularly productive non-academic doctors into the IRCCS research staff. We address this potential selection concern in the robustness analysis below. Overall, the evidence suggests that the IRCCS funding intervention produced a large, rapid, and persistent increase in research output. The effect appears to be amplified by the pandemic-related surge in scientific activity and only marginally reinforced by the performance-based incentive 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 explicitly rewarding productivity in the short term, if such reward is not paired with additional resources. It must also be noted that additional resources already impact positively those who are selected into the perimeter, that are chosen by the management due to their supposed better performance in research-related activities; hence, such individuals possibly already embed a different incentive structure, if compared to the other non-academic doctors. In any case, the presence of academics as well in the pure control groups confirms that the heterogeneity in pre-existing incentive behavioral structure may not only exist across groups, but within them as well.

To sum up, it looks like that academic researchers, already incentivized by intrinsic reasons 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 monetary rewards, 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 addresses the concern of potential bias driven by the violation of the parallel trend assumption.

5.2.1 Parallel Trends

Although the event-study estimates suggest no systematic violations of the parallel trends assumption, we further investigate the robustness of our results given the magnitude of the estimated effects and the potential endogeneity of the treatment assignment. Visual inspection of the baseline event-study (Figure 4) indicates that pre-treatment coefficients are generally small and statistically insignificant. However, we observe a mild downward trend for IRCCS researchers prior to 2016, with the coefficient in 2015 approaching statistical significance at the 95% level. While this pattern does not appear large enough to threaten identification, it motivates a more formal assessment of the sensitivity of our results to potential deviations from parallel trends.

To address this issue, we implement the *honest DiD* methodology proposed by Rambachan and Roth, 2023. This approach constructs robust confidence sets that allow for bounded violations of the parallel trends assumption. We consider two types of restrictions: the *bounded relative magnitude* (BM) approach, which allows post-treatment deviations proportional to the maximum observed pre-treatment difference, and the *smoothness restriction* (SR), which constrains the slope of potential deviations from a linear extrapolation of the pre-treatment trend. The resulting confidence sets indicate that our results remain robust to sizeable deviations from parallel trends. In particular, the estimated

effects remain statistically significant even under substantial violations of the identifying assumption. Detailed results and graphical representations of the Honest DiD sensitivity analysis are reported in Appendix Figures A5–A11.

5.2.2 Validity of the TWFE 2x2 DiD identification

Although the IRCCS recognition occurred at the institutional level in 2018, the hospital updates the list of IRCCS researchers annually. As a result, while a large core group of researchers enters the IRCCS perimeter immediately, additional physicians are formally added in subsequent years (see Figure A2 in the Appendix). This institutional setting therefore generates a partially staggered adoption of the treatment. Table A4 in the Appendix reports the composition of the IRCCS group, distinguishing between the initial cohort of researchers included in 2018 (the "core"), and those added in subsequent years. Furthermore, Table A5 (still in the Appendix), shows how most individual characteristics of the members of the two groups are statistically significant, thus suggesting their heterogenous nature before the implementation of the policies.

In the baseline analysis, however, we define treatment according to an *ever-treated* criterion: individuals who are listed in the IRCCS registry at any point after 2018 are considered treated from 2018 onward. Under this definition, our parameter of interest corresponds to an *intention-to-treat* (ITT) effect rather than an average treatment effect on the treated (ATT). Formally, the ITT estimand can be written as

$$\text{ITT} = \mathbb{E}\left[Y_{it}(1) - Y_{it}(0) \mid \exists\, k \geq 2018 : D_{ik} = 1\right], \tag{4}$$

where D_{ik} denotes the treatment status of individual i in period k. This specification reflects the institutional organization of the research activity within the hospital, where the core IRCCS research group emerges immediately after recognition and subsequent additions often correspond to the formal inclusion of physicians already involved in IRCCS-related projects.

To corroborate the validity of this identification strategy, we perform several robustness checks. First, we verify that the estimated effects are not driven by late entrants into the IRCCS research staff by re-estimating the baseline specification excluding these individuals. The corresponding event-study estimates, reported in Appendix Figures A15 and A16, show patterns similar to the baseline specification, suggesting that the main results are not driven by delayed inclusion into the IRCCS perimeter. If anything, the estimated effects become larger once late adopters are excluded, indicating that the ITT estimates represent a lower bound of the underlying treatment effect. In addition, we

also perform the same analyses by excluding the 2018 adopters, thus accounting for the late adopters as treated from 2018 onward via a 2x2 DiD strategy, notwithstanding their staggered inclusion, to assess whether there are serious common trend divergences which are able to undermine our specification. Figures A17 and A18 suggest that this is not the case: the negative pre-trend shown in 2018 and the late take up of the effect in 2020 are clearly due to the exclusion of the core researchers from the estimation sample.

Second we relax our 2x2 basic assumption and implement a Staggered Difference-in-Differences methodologies on our main outcome, assigning the varying IRCCS-treatment to individuals according to their inclusion in the research perimeter. Then, following Goodman-Bacon, 2021, we decompose the TWFE estimator into its underlying 2×2 comparisons across treatment cohorts. The decomposition shows that the majority (84%) of the identifying variation stems from comparisons between treated and never-treated units rather than from comparisons across cohorts with different treatment timing. This suggests that the estimated effects are not driven by problematic comparisons across treatment cohorts. The corresponding decomposition results are reported in Appendix Figure A14.

Finally, we assess the implications of treatment-timing variation by complementing the baseline TWFE staggered DiD with the imputation-based estimator proposed by Borusyak et al., 2024. Appendix Figure A19 reports a comparison between TWFE OLS and the Borusyak et al. estimator. Borusyak et al., 2024 propose an imputation-based estimator for staggered adoption designs, both for the full balanced panel and for the subset obtained by removing all the core researchers included in the IRCCS list in 2018. The estimator first recovers untreated potential outcomes using only untreated observations (never-treated and not-yet-treated units), typically through unit and time fixed effects, and then imputes counterfactual outcomes for treated observations. Treatment effects are obtained as the difference between observed and imputed outcomes, thereby avoiding the forbidden comparisons that may bias conventional TWFE event-study estimates. In our setting, the estimates display limited evidence of pre-treatment dynamics: as shown in Figure A19, only the first lead is statistically significant, while the remaining pre-treatment coefficients are small and not systematically different from zero. This pattern plausibly mirrors the institutional selection mechanism discussed above, whereby physicians included later in the IRCCS perimeter exhibit slightly stronger pre-existing research trajectories. For the sample obtained by removing the core researchers, the effect are smaller and less precisely estimated, and the pattern noisier, possibly due to the smaller numerosity of the treatment group. However, the similarity between TWFE and Borusyak estimates in the post-treatment period suggests that such deviations are quantitatively limited. To further assess the robustness of our findings, we implement the *honest DiD* procedure of Rambachan and Roth, 2023 on the latter staggered event study estimates, at least for the full

sample (Appendix Figures A20 and A21). The resulting confidence sets confirm that both the main TWFE estimates and the obtained by the imputation of Borusyak et al., 2024 remain stable and statistically significant under plausible bounds on pre-trend violations.

6 Additional Analyses

In this section we present additional analyses aimed at further understanding and interpreting the baseline results, and assessing the robustness of the empirical findings. First, we explore heterogeneity across fields and clinical areas, both in terms of volume output and of research composition across individuals. Second, we investigate the responses to the IRCCS recognition across researchers with different baseline productivity levels, and investigate whether the policy also affects the extensive margin of their research activity (i.e., the probability of starting to publish). We then perform a series of robustness checks addressing concerns related to outcome scaling, the COVID-19 shock, spillover effects across groups, changes in individual status over time, and the presence of highly prolific researchers who might disproportionately drive the estimated effects.

6.1 Field-level Heterogeneity

Volume Margin

A natural question is whether this effect is uniform across scientific fields or instead reflects a reallocation of research effort toward specific clinical areas. To address this, we exploit the WoS-based field classification of each publication and construct, for each author-year, the number of publications produced in a given clinical-scientific area, grouped in macro-areas given the high number of the WoS fields in the dataset. [3]. Results reported in Table 4 indicate that the increase in research output associated with IRCCS recognition is widespread across scientific fields, although its magnitude varies substantially across areas. Each row reports the coefficient from a separate difference-in-differences regression, where the dependent variable is the annual number of publications in a given specialization. The reported coefficient (column "Coef.") corresponds to the intention-to-treat (ITT) effect of IRCCS recognition on field-specific output. Standard errors are clustered at the author level.

Statistical inference is reported in three ways. First, conventional p-values are shown in column "p-value". Second, we report Romano–Wolf stepdown adjusted p-values (column

[3]The methodology through which such macro-categories are constructed is reported in Appendix Table A6

"RW p-val."), which correct for multiple hypothesis testing across the set of specializations and therefore provide more conservative significance thresholds (Romano and Wolf, 2005, 2016). Third, we report a joint test of pre-treatment coefficients ("Pre (joint)"), which assesses whether pre-treatment trends are jointly equal to zero; failure to reject supports the parallel trends assumption.

The last two columns provide contextual information. In particular, "Note" summarizes the strength and reliability of the evidence using a simple classification: "++" for positive and statistically robust effect (greater than or equal to 50% of pre-treatment mean), "+" for a positive effect that is statistically significant but weaker (less than 50% of pre-treatment mean), "0" denotes no statistically significant effect, and "?" flags cases where estimates are potentially contaminated by pre-trends or where inference is less reliable.

The largest effects are observed in core medical and research-intensive areas. For instance, Oncology and Hematology exhibit an increase of 0.367 publications per year, corresponding to roughly +155% relative to the pre-treatment mean (0.236). Similarly, Internal Medicine (both generic and specialties) shows increases between 0.305 and 0.353 publications, implying effects in the order of +120–180%.

Biology also displays a large effect (+0.258, about +155%), while Cardiology, Diagnostics and Imaging, and Infectious Diseases show sizeable increases ranging between +85% and +180% relative to their respective pre-treatment means. By contrast, we do not detect statistically significant effects in fields such as Neurosciences, Surgery, Rehabilitation, or Technology and Engineering. These patterns suggest that the expansion in research output is concentrated in areas where research activity and infrastructure were already relatively developed prior to the policy intervention.

The absence of systematic pre-trends in most specializations supports the validity of the identification strategy. While a few fields (e.g., Maternal and Child Health, Oncology) display some evidence of pre-treatment dynamics, these cases are limited and do not alter the overall pattern of results. Such fields still display sizeable and relevant expansions after 2018.

The dynamic event-study estimates reported in Appendix Figures A22-A24 further corroborate these findings. They show that the increase in publications emerges after the IRCCS recognition and follows a similar temporal pattern across most of the affected fields, with no evidence of anticipatory effects. Overall, the evidence points to a broad-based expansion of research activity rather than a reallocation of effort toward specific disciplines.

Specialization	ITT			Pre (joint)		Mean	Note
	Coef.	p-value	RW p-val.	F-stat	p-value		
Biology	0.258***	0.000	0.012**	0.576	0.680	0.166	++
	(0.072)						
Cardiology/Cardiovasc. Diseases	0.156***	0.004	0.037**	1.174	0.321	0.179	++
	(0.054)						
Chemistry and Physics	0.095***	0.000	0.006***	0.235	0.918	0.022	++
	(0.025)						
Diagnostics and Imaging	0.138***	0.005	0.037**	1.146	0.334	0.143	++
	(0.049)						
Infectious Dis./ Immunology	0.150***	0.003	0.033**	0.727	0.574	0.084	++
	(0.049)						
Internal Medicine (Generic)	0.353***	0.000	0.001***	0.830	0.506	0.124	++
	(0.060)						
Internal Medicine (Specialties)	0.305***	0.000	0.005***	1.328	0.258	0.259	++
	(0.073)						
Maternal and Child Health	0.168**	0.019	0.042**	2.801	0.025**	0.139	?
	(0.071)						
Neurosciences and Psych.	0.082	0.175	0.311	0.224	0.925	0.245	0
	(0.060)						
Oncology and Hematology	0.367***	0.005	0.037**	3.110	0.015**	0.236	?
	(0.131)						
Pharmacology	0.091***	0.003	0.033**	0.980	0.418	0.068	++
	(0.030)						
Public Health and Health Systems	0.198***	0.000	0.002***	0.568	0.686	0.032	++
	(0.040)						
Rehabilitation and Nursing	0.018	0.370	0.558	0.528	0.716	0.021	0
	(0.019)						
Social Sciences and Humanities	0.006	0.591	0.668	2.246	0.063*	0.014	?
	(0.012)						
Surgery and Procedures	0.042	0.548	0.668	0.638	0.635	0.192	0
	(0.069)						
Technology and Engineering	0.047	0.210	0.341	1.291	0.272	0.074	0
	(0.038)						

Table 4: Each row reports results from a separate regression of the yearly number of publications in a given specialization on a post-treatment indicator for IRCCS-only researchers, controlling for unit and year fixed effects. Standard errors, reported in parentheses beneath the ITT coefficient, are clustered at the author level. The block Pre (joint) reports the F-test statistic and p-value from a joint test of all pre-treatment leads (2012–2015). Romano–Wolf stepdown adjusted p-values are reported to account for multiple hypothesis testing across the family of specialization outcomes. Depending on the variance estimator and finite-sample correction, the reported test statistic may be an F statistic or a chi-squared statistic. Mean pre-treatment level is computed over the pre-2017 period. Significance: * p<0.10, ** p<0.05, *** p<0.01.

Composition

Table A7 in the Appendix examines whether the increase in research output is accompanied by changes in the composition of research activity across fields. The results provide little evidence of systematic reallocation. Across virtually all specializations, the estimated effects on publication shares are small in magnitude and statistically insignificant. Even in cases where coefficients are marginally significant (e.g., Surgery), the economic magnitude remains limited. For example, the largest estimated effect corresponds to a change of about 3.2 percentage points, which is modest relative to baseline shares.

These findings indicate that the IRCCS recognition does not substantially alter the distribution of research effort across scientific areas. Instead, researchers tend to expand production along their existing lines of specialization. This interpretation is consistent with the volume results, which show strong increases in output within fields but no evidence of systematic shifts across them. The event-study patterns reported in the Appendix Figures A25-A27 confirm this interpretation. The dynamics of field shares remain largely flat both before and after the policy intervention, with no clear structural breaks around the IRCCS recognition.

6.2 Individual Heterogeneity

Baseline Productivity

We explore heterogeneous responses to the IRCCS recognition by splitting researchers according to their pre-treatment publication record. In particular, we construct measures of baseline productivity using the average number of publications in the pre-2017 period, and classify individuals into groups based on the quartiles of the distribution. Such analysis also corroborates our identification strategy, as by looking at these heterogeneity trajectories we can rule out potential concerns related to the structural differences between treated physicians (research-oriented physicians) and control ones (clinic-oriented doctors). To enforce this exercise even more, we also exclude from the sample those who do not publish a paper either before 2017 (whom we call *never-publishers*), to assess whether the main results are driven by a distorted comparison between prolific individuals and doctors who never committed into scientific production. Appendix Figure 5 reports the corresponding event-study estimates.

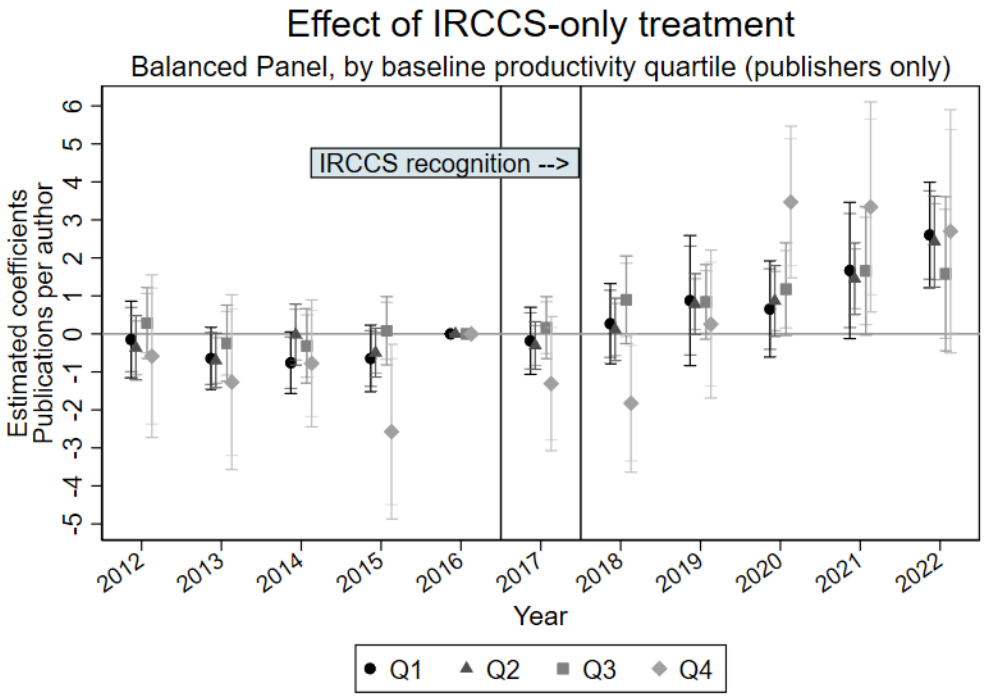


Figure 5: Event-study of the IRCCS effect by baseline productivity quartiles (balanced panel, publishers only).

Two main patterns emerge. First, the response to IRCCS recognition is broadly positive across all productivity groups, with relatively similar post-treatment dynamics across quartiles. While some heterogeneity is present, it is not strongly monotonic. In particular, researchers in the first and second quartiles display a more stable and consistent increase in publication output over time. By contrast, the response among top-quartile researchers (Q4) appears more volatile, with an initial decline around the time of the policy implementation followed by a subsequent increase in later years. This suggests that the adjustment process may be more heterogeneous among the most productive individuals. Overall, the evidence points to a relatively diffuse effect of the IRCCS recognition across the productivity distribution, rather than one concentrated exclusively among top-performing researchers.

Second, and importantly, we observe a positive and significant response among lower-productivity researchers. This suggests that the IRCCS recognition does not exclusively benefit already highly productive individuals, but also induces a non-negligible increase in output among less active researchers, among individuals with prior publication activity. When we include individuals with no pre-treatment publications, we add substantial variability in the estimated dynamics, as the treatment effect combines an intensive mar-

gin (increased output among active researchers) with an, albeit weak, extensive margin (entry into publishing). The latter is inherently more discrete and heterogeneous, leading to noisier estimates when both margins are jointly considered, as represented in the Appendix Figure A28. This highlights that the IRCCS recognition operates through distinct channels: while it certainly increases the productivity of already active researchers, it may also induce entry into research activity among previously inactive individuals, albeit in a more irregular and heterogeneous manner.

To further investigate the extensive margin of research activity, we focus on individuals with no publications in the pre-treatment period. Figure 6 shows that IRCCS recognition leads to a gradual increase in publication output even within this group, although the estimates are imprecise. This suggests that the policy may have contributed, albeit weakly, to activating previously non-publishing physicians. However, this pattern should be interpreted with caution, as the observed increase occurs mainly in the post-2020 period and may partly reflect the impact of the COVID-19 pandemic on research activity.

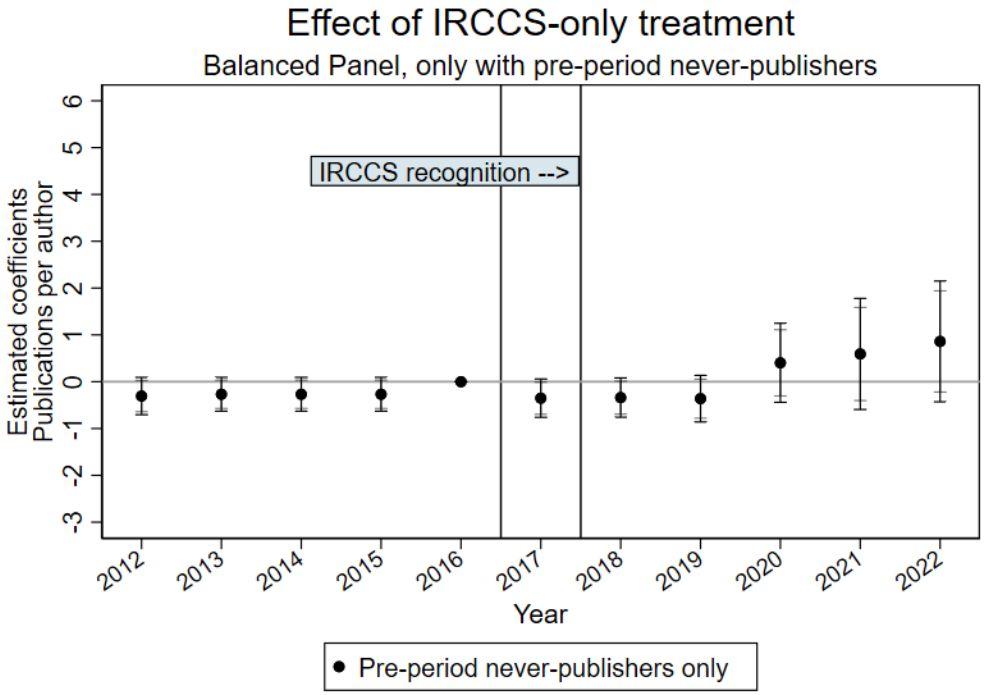


Figure 6: Event-study of the IRCCS effect for individuals with no pre-treatment publications (balanced panel).

At the same time, when restricting the sample to pre-period publishers only (Figure A29 in the Appendix), the magnitude of the effect remains large, confirming that the main results are not driven by a spurious comparison between active and inactive

researchers, and that entry into publishing matters weakly in the present context.

Other Characteristics

We further explore heterogeneous responses to the IRCCS recognition across individual characteristics, focusing on tenure at the hospital, as well as gender and age. Tenure and age are defined based on their distribution at the time of treatment, splitting individuals into below- and above-median groups. Appendix Figure A30 reports the corresponding event-study estimates by tenure. The estimated effects are broadly similar in shape and timing across tenure groups, with no evidence of strong differential responses. While high-tenure researchers exhibit slightly larger coefficients in the post-treatment period, the overall patterns remain qualitatively similar, indicating that the effects are not concentrated among more senior individuals. This is consistent with the evidence from baseline productivity, which already suggests that the policy does not disproportionately benefit a specific subset of researchers.

Overall, the findings point to a relatively diffuse increase in research activity across observable characteristics, reinforcing the interpretation of a broad-based productivity effect rather than one driven by specific subgroups. Additional evidence for heterogeneity by gender and age is reported in Appendix Figures A32 and A31, and confirms the absence of substantial differences across groups.

6.3 Robustness

We perform a battery of robustness checks to assess the sensitivity of our results to alternative outcome definitions, sample restrictions, and potential confounding factors; all detailed estimates are reported in the Appendix.

First, we verify that the results are not driven by the scale of the outcome variable by re-estimating the baseline specifications using log-transformed and inverse hyperbolic sine transformations of publications. The corresponding estimates are reported in Appendix Table A8. While the magnitude of the coefficients is mechanically reduced, all estimates remain consistent in sign and statistical significance.

Second, we address concerns related to the COVID-19 pandemic by excluding all publications explicitly related to COVID-19, identified through keywords in article titles[4]. The resulting estimates, reported in Figures A33 and A34, remain qualitatively unchanged, suggesting that the observed effects are not driven by pandemic-related publication dynamics.

[4]The keywords are: *Covid, Covid-19, Coronavirus, 2019 ncov, pandemic, lockdown, quarantine, ncov, SarS-COV-2, Omicron*, and *Delta*, written in various formats.

Third, we investigate the presence of spillover effects across groups by excluding co-authored publications between treated and control units, as well as double-treated individuals. The corresponding results are reported in Appendix Table A9 and Figure A35. The estimates indicate that cross-group spillovers are limited and do not materially affect the baseline results, although some within-group amplification effects emerge among IRCCS researchers.

Finally, we assess the role of compositional changes and highly prolific individuals. We first exclude all units who switch treatment status over time, namely physicians who transition between academic and non-academic positions after the policy implementation, thereby moving across treatment groups. These switches may arise, for instance, when non-academic doctors become faculty members or, conversely, when academics leave research-oriented positions. Excluding such individuals, we find virtually identical results (Appendix Figure A36). We then progressively trim the upper tail of the publication distribution by excluding individuals in the top percentiles of annual output; the corresponding estimates are reported in Appendix Figures A37 and A38. While the magnitude of the estimated effects decreases as more restrictive thresholds are applied, the direction and significance of the results remain stable, indicating that the findings are not solely driven by highly productive researchers.

7 Mechanisms

The evidence presented so far shows that the IRCCS recognition led to a substantial increase in research productivity. We now investigate the mechanisms underlying this effect. In particular, we examine whether the observed increase in publications is driven by improved access to research funding and/or by changes in the organization of research activity within the institution. To this end, we combine our main dataset with additional information on funded research activity from the Italian Ministry of Health's Research Workflow, which allows us to distinguish between *Current Research* (ordinary institutional funding) and *Targeted Research* (competitive project-based grants). We then study how the IRCCS recognition affected collaboration patterns, team expansion, and authorship structure, as well as the quality of research output as proxied by citations.

7.1 Access to IRCCS funds

Current Research

Current Research (*Ricerca Corrente*) represents the main source of institutional funding associated with the IRCCS recognition. These funds are allocated annually by the Ministry of Health to IRCCSs based on predefined research lines and amount, at the aggregate level, to approximately €150–175 million per year, with single-institution allocations ranging from a few hundred thousand euros up to about €10 million. Following the IRCCS recognition in 2018, the hospital under study received substantial additional resources—amounting to approximately €7 million per year from 2020 onwards—within the macro-areas of Personalized Medicine and Innovative Biotechnologies.

To assess whether the increase in research productivity can be linked to these resources, we exploit micro-data from the Ministry's Research Workflow, which reports all funded research lines and the publications associated with them. We merge these data with our publication-level dataset and identify all researchers who contribute to IRCCS-funded research lines [5]. This allows us to refine the treatment definition by distinguishing between IRCCS researchers who are actively involved in funded research lines and those who are not. We then estimate a 2×2 Difference-in-Differences model, comparing these groups to the full control group within our baseline empirical framework (Equation 1). Results are reported in Figure 7. The evidence shows that the increase in research productivity following IRCCS recognition is almost entirely driven by researchers involved in Current Research lines. In contrast, IRCCS-affiliated individuals who are not connected to any research line display no significant improvement in output. This pattern suggests that the productivity gains are closely tied to the access to institutional funding rather than to the IRCCS status (and prestige) per se. We display additional evidence by using a journal-level panel, where the unit of observation is the scientific outlet and journals are classified based on whether they ever host publications associated with IRCCS research lines; such findings confirm previous results, and are reported in Appendix Figure A40.

While Current Research funding cannot be linked to specific projects, its flexible and recurring nature likely supports a wide range of research-related activities, including per-

[5] We perform a two-step merging process: an exact merge based on the publication title, and then a fuzzy merge based on correspondence between the scientific outlets, reported in both sources of data, and on similarity of the publication titles (through the Blasnik, 2007 probabilistic matching procedure, by fixing a minimum score of 0.995). Note that the procedure is not able, in any case, to match all the publications provided by the Workflow, as many publications are authored by individuals who are not in our panel (external affiliates, PhD students, research fellows, collaborator and non-employees who still the affiliation anyway). For this reason, we keep our matching score extremely high to avoid false positives, as our outcome can at most be a lower bound nonetheless.

sonnel, data infrastructure, and smaller-scale inputs such as publication fees and research tools. Overall, these results point to a broad relaxation of resource constraints as a key mechanism behind the observed increase in scientific output.

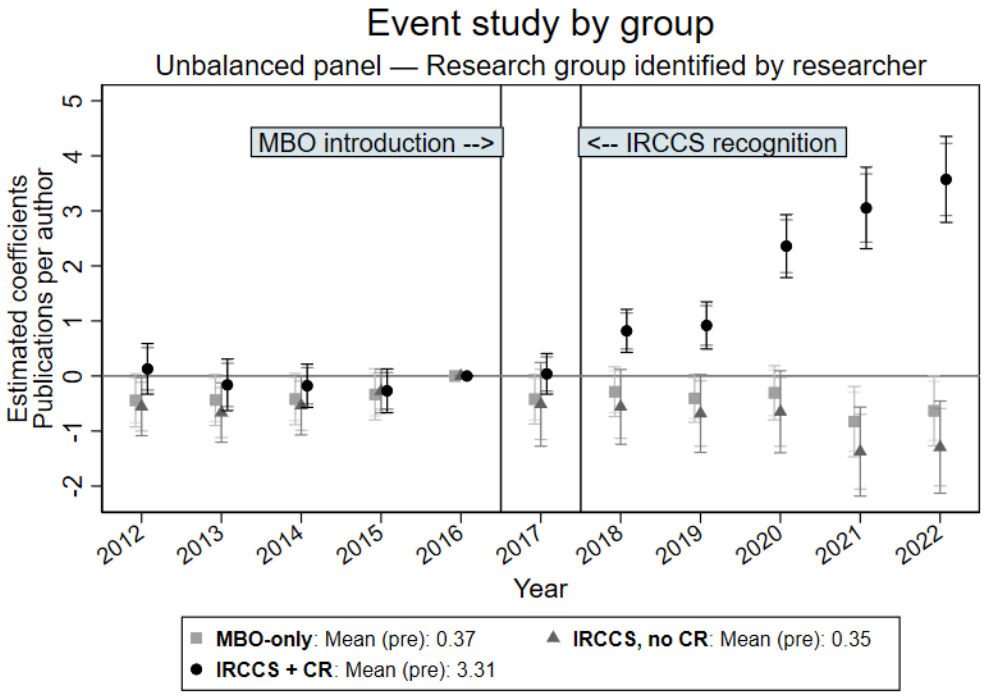


Figure 7: Event Study of the impact of IRCCS recognition + inclusion in Current Research lines on publications, across different comparison groups - by researchers.

Targeted Research

We next examine the role of Targeted Research (*Ricerca Finalizzata*)[6], namely competitive grants awarded by the Ministry of Health to finance specific research projects. Unlike Current Research funding, these resources are project-based and allocated through competitive calls, with projects led by Principal Investigators (PIs) within the institution. These grants typically finance projects of three-year duration, with awarded amounts generally ranging between €300,000 and €500,000 per project, and reaching up to €1 million in recent calls linked to post-Covid recovery programs.

[6]Although Targeted Research funds are allocated through competitive calls, IRCCSs benefit from a more direct and privileged access to these competitions. In particular, unlike most healthcare institutions, which can typically apply only through intermediary entities such as Regions or the National Institute of Health, IRCCSs are eligible to participate directly in the calls. This institutional feature increases both the likelihood of participation and the effective access to competitive funding for IRCCS-affiliated researchers.

To identify recipients of Targeted Research funds, we reconstruct the set of PIs and co-PIs associated with projects awarded to the hospital between 2018 and 2022, by combining information from the Ministry's Workflow with funding acknowledgments in publications and additional sources. We then flag individuals in our panel who are awarded a grant and exploit the staggered timing of these awards to estimate dynamic treatment effects within our baseline framework.

We implement two staggered Difference-in-Differences specifications. First, we compare grant recipients to all individuals who never receive Targeted Research funding. Second, we restrict the comparison group by excluding IRCCS-affiliated individuals who never obtain a grant, thus focusing on within-group variation. Results are reported in Appendix Figure A39.

Across specifications (OLS and Borusyak, to be consistent with our previous staggered event-study), we find a modest and imprecisely estimated increase in research output following the receipt of a grant. While point estimates are generally positive and peak in the first few post-treatment periods, they are rarely statistically significant and tend to fade after three years, consistent with the typical duration of funded projects. At the same time, the event-study coefficients display a clear upward pre-trend, suggesting that individuals who eventually receive Targeted Research funding were already on a rising productivity trajectory prior to treatment. This pattern is consistent with a selection mechanism, whereby more active or promising researchers are more likely to obtain competitive grants.

This evidence, combined with the results on Current Research funding, points to a more prominent role for the institutional and recurring component of IRCCS resources.

7.2 Networks and Research Organization

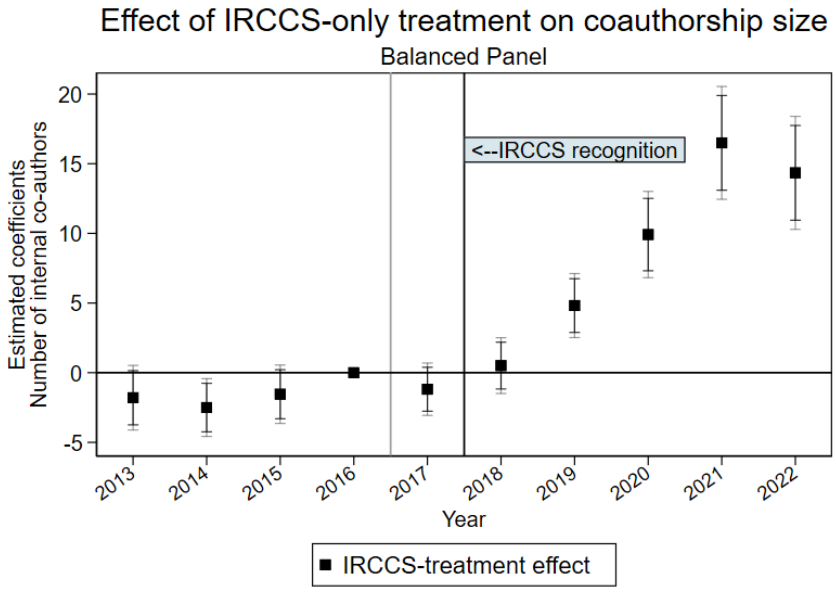
Team expansion

We begin by investigating whether the increase in research productivity is associated with changes in the size and composition of research teams. To this end, we consider three complementary outcomes: the number of internal co-authors per publication (i.e., affiliated to the institution under study), the number of new co-authors with respect to a given individual ("joiners"), and the share of joiners over total internal co-authors. For all outcomes, we exclude the first year of the sample (2012) from the event-study estimation. Since 2012 is the first observed year, collaborations initiated prior to the sample period may mechanically appear as new in that year, leading to a left-censoring issue that is particularly relevant for measures based on the entry of new collaborators. Dropping this initial year ensures a more consistent and comparable identification of pre- and post-

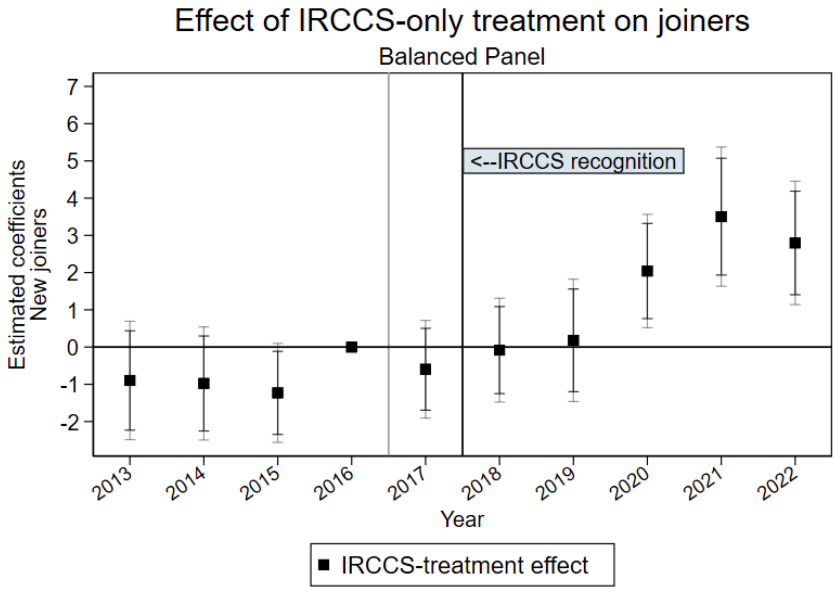
treatment dynamics across all specifications.

Figure 8 reports the dynamic effects of the IRCCS recognition on these outcomes. We find clear and consistent evidence of team expansion following the policy. The number of internal co-authors per paper increases sharply and persistently after 2018, indicating that research projects involve progressively larger teams. The effect is economically large and grows over time, pointing to a sustained scaling-up of collaborative activity. Turning to team composition, the number of new collaborators also increases after the IRCCS recognition, although the effect is more gradual and somewhat noisier.

Despite the increase in the number of joiners, which would suggest that part of the expansion in team size is driven by the entry of new researchers into existing collaboration networks, the share of new collaborators over total co-authors does not display a corresponding rise and, if anything, shows a mild decline over time (Appendix Figure A41). This implies that team expansion is not primarily driven by an extensive margin of entirely new collaborators to the treated physicians' research agenda, but rather by an intensive margin, whereby existing collaborators participate more frequently and in larger groups. In other words, IRCCS recognition appears to foster the consolidation and expansion of a core research network, rather than simply attracting a disproportionate number of new contributors.



a) Number of co-authors per publication.



b) Number of new collaborators (joiners).

Figure 8: Event-study estimates of the effect of IRCCS recognition on team size and composition.

The results indicate that the increase in research output is partly explained by a reorganization of the production process, with IRCCS researchers coordinating larger and more structured teams, consistent with a shift towards more resource-intensive and collaborative forms of research. Results are robust to alternative specifications comparing different treatment groups, reported in Appendix Figures A42– A44.

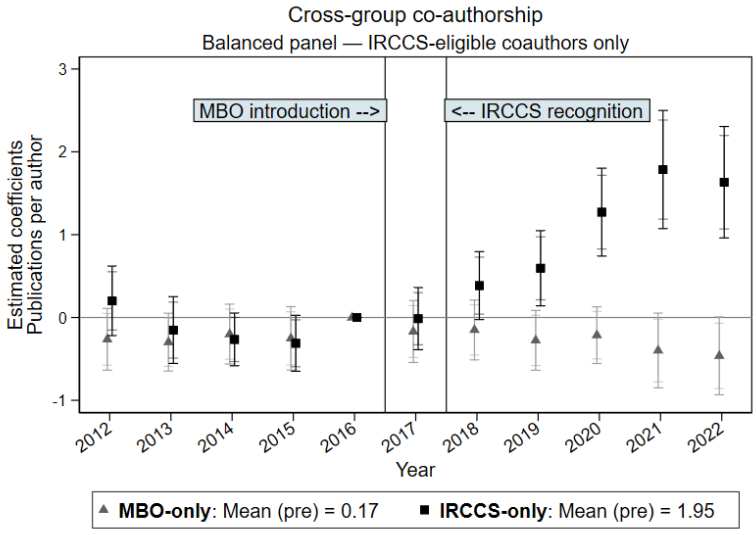
Cross- and within-group co-authorship

We next investigate whether the policies affected the structure of collaboration networks across different groups within the institution. In particular, we assess whether the increase in research output is associated with stronger interactions between IRCCS researchers and MBO-eligible physicians, or whether it mainly reflects an intensification of within-group collaborations.

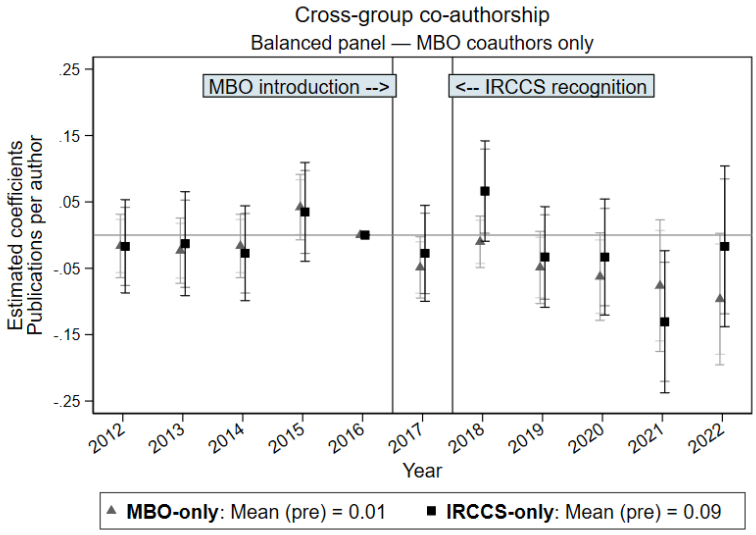
To this end, we define as "networked" all publications featuring at least two co-authors affiliated with the institution, and we classify them according to the type of collaboration involved. We distinguish between: (*i*) publications co-authored exclusively with IRCCS researchers (*Type A*); (*ii*) publications co-authored exclusively with MBO-eligible physicians (*Type B*). After excluding the overlapping category of double-treated individuals, we estimate dynamic effects following Equation 1, using the full control group as reference. Figure 9 reports the event-study estimates for within-group collaborations. Panel (a) shows the dynamics of publications co-authored with IRCCS researchers only, while panel (b) focuses on collaborations involving MBO-eligible physicians only. The corresponding estimates for the unbalanced panel are reported in Appendix Figure A45.

Two main patterns emerge. First, pre-treatment coefficients are close to zero and do not display systematic trends, suggesting that collaboration patterns across groups were stable prior to the policy interventions. Second, following the IRCCS recognition, collaboration increases are largely concentrated within groups. IRCCS-treated researchers experience a substantial and persistent rise in within-group co-authorship, whereas MBO-eligible physicians do not display a comparable increase in collaborative activity.

In contrast, we find little evidence of a significant expansion in cross-group collaborations. If anything, the dynamics of mixed co-authorship remain flat or slightly negative relative to the control group. This indicates that the additional resources associated with IRCCS recognition primarily reinforce existing research networks, rather than fostering integration between clinical (MBO) and academic (IRCCS) personnel.



a) Outcome: publications co-authored with IRCCS units only.



b) Outcome: publications co-authored with MBO-eligible units only.

Figure 9: Dynamic impact of the policy implementation on annual cross-group publications in the balanced panel.

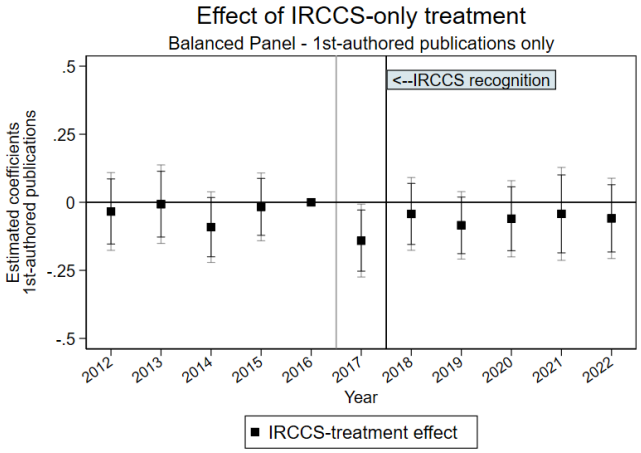
We can therefore conclude that the observed increase in research productivity is driven by the strengthening of within-group collaboration structures, with limited spillovers across groups. This pattern is consistent with our previous robustness check for the SUTVA, and with a model in which IRCCS recognition enhances the productivity of already active teams, without substantially altering the boundaries of collaboration networks within the institution. As a matter of fact, the positive effect we retrieved on joiners and the co-authorship size, ought to be credited to within-group dynamics.

Co-authorship structure

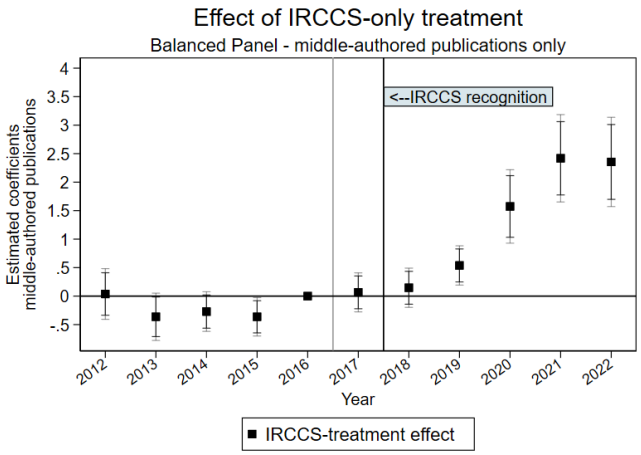
Having established that IRCCS recognition leads to an expansion of within-group research teams, we next analyse whether this is accompanied by changes in the internal organization of scientific production. In particular, we study how researchers' positions within the author list evolve following the policy, distinguishing between first-, middle-, and senior (last) authorship roles. Figure 10 reports the corresponding event-study estimates. We find that the increase in research output is highly heterogeneous across author positions. First-authored publications display no clear or systematic increase following the IRCCS recognition, with coefficients remaining close to zero throughout the post-treatment period. This suggests that the policy does not primarily operate through an increase in researchers' leading contributions within publications, but rather through a broader expansion of collaborative involvement.

In contrast, middle-authored publications exhibit a large and persistent increase after 2018. The magnitude of the effect grows over time and is substantially larger than that observed for other author positions. This pattern indicates a sharp rise in participation in collaborative projects, consistent with the expansion of research teams documented in the previous subsection.

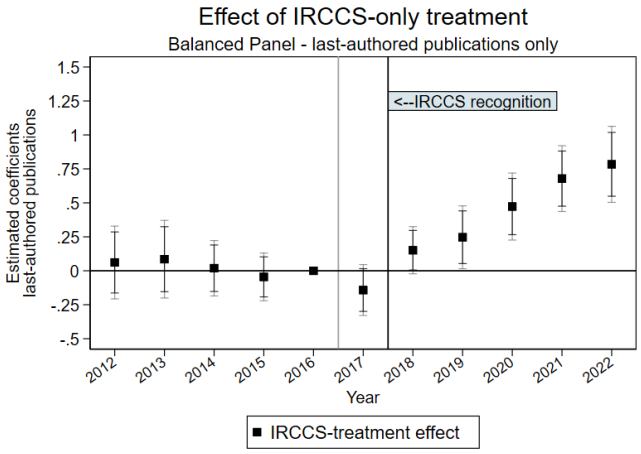
Evidence on senior roles, proxied by last-authored publications jointly, also points to a positive effect, although smaller and less precisely estimated compared to middle authorship. This suggests that IRCCS researchers increasingly take part in coordinating and supervising research activities, but that the dominant margin of adjustment operates through broader involvement across multiple projects rather than through a disproportionate increase in leadership positions. Pre-treatment coefficients are generally stable across all author positions and do not display systematic trends, supporting the validity of the identification strategy. Additional specifications comparing alternative treatment groups yield similar patterns and are reported in Appendix Figure A46.



a) Outcome: first-authored publications.



b) Outcome: middle-authored publications.



c) Outcome: last-authored publications.

Figure 10: Dynamic impact of the IRCCS recognition on authorship position within publications (balanced panel).

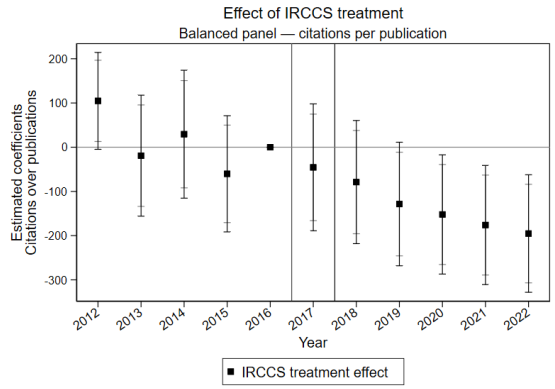
8 Research Impact

8.1 Research Quality

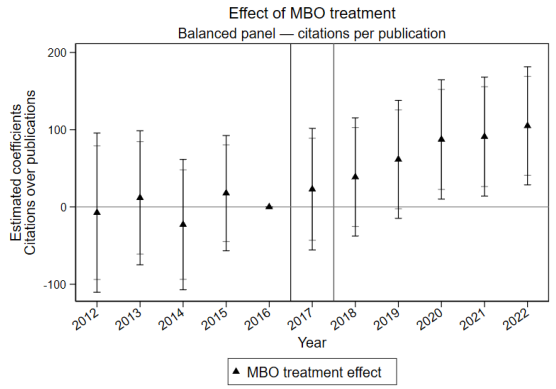
A further relevant question is whether the increase in publications induced by the two policies was accompanied by changes in the impact and quality of research output. To address this issue, we analyse citations as a proxy for scientific influence. In this subsection, differently from some of the previous exercises focused on the IRCCS effect only, we explicitly study the dynamic effects of both treatments: the IRCCS-only treatment and the MBO treatment (the latter including double-treated individuals, consistently with the baseline event-study specification).

Since citations are observed in Fall 2024 for papers published in different years, a direct comparison across cohorts of publications would be misleading, as older papers have had more time to accumulate citations. We therefore construct citation-based outcomes that account for publication age. First, we discount citations by the time elapsed between publication year t and 2024, using a yearly discount factor. We then build three complementary measures: (i) publications weighted by discounted citations, which capture whether the treatment affects output more strongly for highly cited research; (ii) the total number of discounted citations accumulated by the papers published in a given year; and (iii) discounted citations per publication, which we interpret as a measure of average research quality.

Figure 11 reports the dynamic effects on discounted citations per publication, which is the most direct measure of the quality margin. The contrast between the two policies is stark. For IRCCS-treated researchers, the average number of discounted citations per paper declines markedly after 2018, with coefficients becoming negative and increasingly large over time. This suggests that the expansion in scientific output associated with the IRCCS recognition came with a reduction in average impact per publication. By contrast, MBO-treated physicians display the opposite pattern: discounted citations per publication rise after the policy, indicating that, although the MBO scheme does not generate a sizable increase in the number of publications, it is associated with an improvement in the average impact of the papers produced.



a) IRCCS effect on discounted citations per publication.



b) MBO effect on discounted citations per publication.

Figure 11: Event-study estimates of the effect of IRCCS recognition and MBO treatment on discounted citations per publication (balanced panel).

These findings are consistent with the broader interpretation of our results. The IRCCS recognition appears to relax resource constraints and stimulate a substantial increase in research production, but this expansion is accompanied by a decline in average paper quality. The MBO scheme instead seems not to affect quantity, yet it is associated with a stronger quality orientation. One possible explanation is that the MBO reward, being tied not only to publication counts but also to journal quality through the Impact Factor, may have induced physicians to concentrate on fewer but more valuable outputs. Additional evidence based on discounted total citations and on citation-weighted publications is reported in Appendix Figure A47. These alternative outcomes confirm the same overall pattern: IRCCS increases aggregate citation counts mainly because it raises the number of publications, whereas MBO does not boost total output but is associated with higher citations per paper.

8.2 Strategic Behavior of MDs

While the previous evidence suggests that the MBO scheme does not increase the overall number of publications produced by treated physicians, two findings point to potentially non-trivial behavioral responses. First, we document a non-negligible transition of MDs into the IRCCS perimeter, together with anticipatory patterns in the staggered DiD estimates, consistent with a productivity-based selection process. Second, MBO-treated physicians exhibit an increase in citations per publication, in contrast with the decline observed for IRCCS researchers. These patterns raise the question of whether MDs respond to the incentive scheme by strategically adjusting their research activity, rather than by

genuinely increasing productivity. In particular, since the MBO bonus is tied to journal Impact Factors and adjusted for co-authorship, physicians may have incentives to target specific publication bundles that maximize payout without increasing total output.

To investigate this mechanism, we exploit administrative data on individual payouts, eligible publications, and Impact Factor–based scores, and compare MBO-treated physicians with untreated faculty members over the period 2017–2022. Given data limitations on pre-policy Impact Factors, we estimate a two-way fixed effects specification where treatment is defined by eligibility to the MBO scheme. A relevant feature of the institutional setting is that the set of publications used to compute the bonus is not mechanically constructed, but involves a validation process in which physicians can revise the list of eligible papers (e.g., by adding missing outputs or correcting misattributions). As a result, the number of "eligible" publications may differ from the full set of publications observed in bibliometric data.

Importantly, although only MBO-eligible physicians receive the monetary bonus, the same bibliometric indicators (including Impact Factor scores and implied payouts) are computed for all researchers in the institution, including faculty members who are not eligible for the scheme. This feature allows us to construct a meaningful comparison group and to interpret differences in these outcomes as reflecting behavioral responses to the incentive, rather than mechanical differences in measurement. We explicitly exploit this discrepancy by comparing total and eligible publications, which allows us to detect potential strategic reporting behavior. We consider four outcomes: (i) the monetary payout, (ii) the total Impact Factor score (adjusted for co-authorship), (iii) the number of eligible publications, and (iv) the gap between total and eligible publications, which captures potential manipulation in self-reported outputs.

The results, reported in Appendix Table A10, provide little support for the presence of systematic gaming behavior. MBO-treated physicians do not outperform their untreated counterparts in terms of payout, Impact Factor scores, or eligible publications. Moreover, we find no robust evidence of differential reporting behavior in the gap between actual and eligible publications. Therefore, the improvement in citation-based outcomes for MBO-treated physicians is unlikely to be driven by strategic manipulation of the incentive scheme. Rather, it is consistent with a genuine shift toward higher-quality research output, in contrast with the quantity-driven response observed among IRCCS researchers.

8.3 Additional evidence on hospital specialization

As a final descriptive exercise, we examine a selected set of hospital-level indicators from the Nationaal Outcome Program (*Programma Nazionale Esiti*, PNE), a public, national

monitoring system developed by AGENAS (the National Agency for Regional Health Services) to provide comparable measures of hospital activity, clinical outcomes, and care processes in Italy, under supervision of the Ministry of Health. These outcomes are scraped from the official web page of the system, and we chose them *ex ante* to capture potential patterns in clinical specialization following the 2018 IRCCS recognition, focusing on areas where a research-oriented institutional upgrade could plausibly lead to a greater concentration of high-technology, research-intensive, and complex services, particularly in oncology and cardiovascular care. The aim is not to establish causal effects, but to assess whether the hospital-level clinical trajectory aligns with the broader organizational transformation documented in the physician-level publication data.

We focus on two complementary sets of outcomes. First, we consider a broad set of standard PNE quality indicators, covering short- and medium-term mortality (e.g. 30-day and 1-year mortality for acute myocardial infarction, stroke, and major cancer surgeries) and readmission rates (e.g. 30-day readmissions for heart failure, pulmonary embolism, and elective procedures such as hip and knee replacement). These measures span the main clinical domains of hospital activity—cardiovascular conditions, oncology, and orthopedic care—and are routinely used to proxy the quality and appropriateness of care. While they allow us to detect potential changes in clinical performance following IRCCS recognition, they are not the primary margin along which research-oriented institutional upgrades are expected to operate.

Second, we examine a set of volume-based indicators normalized by hospital capacity (i.e. hospital beds, in thousands), aimed at capturing shifts in clinical specialization and technological intensity. Specifically, we consider procedure volumes in high-complexity and research-intensive domains, including pancreatic cancer resections (both total and resection-specific measures), lung cancer surgeries (distinguishing between overall and minimally invasive approaches), and colon cancer surgeries performed with minimally invasive techniques. We also include cardiovascular interventions, such as isolated coronary artery bypass grafts, total bypass procedures, and carotid revascularizations (both overall volumes and endarterectomy-specific measures). Finally, we consider procedure-specific indicators capturing the adoption of advanced treatments, such as acute myocardial infarction cases treated with thrombolysis. These outcomes provide a direct proxy for the extent to which clinical activity shifts toward more complex, technology-intensive, and research-oriented services.

The empirical exercise consists of estimating simple difference-in-differences specifications at the hospital-year level, comparing the treated hospital to a set of control hospitals within the same region (Lazio), with unit and year fixed effects. The control group includes public and private hospitals that are not granted IRCCS status over the sample period

and operate in comparable clinical domains, thereby capturing common regional shocks in demand, policy, and healthcare organization. The choice of the regional dimension for the selection of the control is motivated by the administration of Italian healthcare, which is run at the regional level. For each outcome, we report the average treatment effect (ATT), together with joint tests of pre-treatment coefficients to assess the credibility of parallel trends. Given the large number of outcomes, we adjust for multiple hypothesis testing using Benjamini–Hochberg corrections within each family (Benjamini & Hochberg, 1995).

The evidence, displayed in Appendix Tables **??** for quality outcomes, and **??** for specialization outcomes, points to a consistent pattern along the specialization margin. Quality indicators display mostly mixed, noisy, and unstable dynamics, with limited interpretability. In contrast, volume-based measures show more systematic increases in complex and technology-intensive procedures, particularly in oncology and in the adoption of minimally invasive techniques. A key qualification concerns however the presence of pre-existing trends in several of these outcomes. For such indicators, volumes were already evolving prior to 2018, in some cases along a clear upward trajectory. This weakens a rigorous difference-in-differences interpretation, as part of the observed post-treatment increase may reflect the continuation of an ongoing process rather than a discrete break induced by IRCCS recognition.

More precisely, a subset of indicators displays relatively clean pre-trends and a clear post-2018 break. This is the case for minimally invasive lung cancer surgeries and carotid endarterectomy, where pre-treatment dynamics are flat and non significant, and the post-2018 increase appears sharp, sustained and sizeable (almost 25% of the pre-2018 mean for the endarterectomies, and a more than 85% for the mini-invasive lung cancer surgeries). A similar pattern, albeit with more noise, emerges for standard pancreatic cancer surgery volumes, where the post-treatment expansion is large relative to the pre-period baseline, although the interpretation is flawed by pre-trends. By contrast, several other indicators are more ambiguous. Colon cancer minimally invasive surgery, overall carotid revascularization volumes, and coronary bypass activity exhibit visible pre-treatment dynamics, making it difficult to disentangle continuation from treatment effects. In these cases, post-2018 increases are present but broadly aligned with pre-existing trends, limiting causal interpretation. Finally, for some indicators (e.g. total lung cancer surgery volumes or aggregate bypass procedures), no post-treatment shift is detectable.

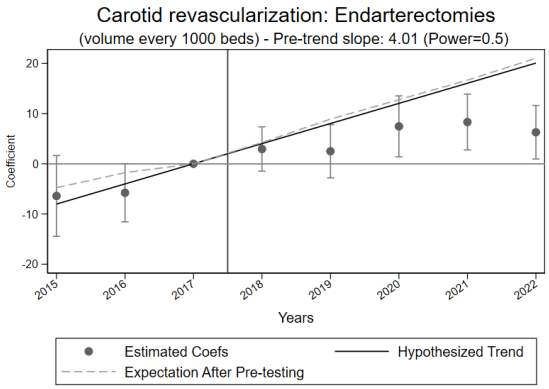
The event-studies in Figures 12 and 13 provide a more granular view of the dynamics underlying the baseline estimates, but also make the identification challenges transparent. A central limitation is the short pre-treatment window, which implies that standard joint tests of pre-trends are severely underpowered. Following Roth, 2022, failure to reject the null of no pre-trends cannot be interpreted as evidence in favor of the identifying

assumptions. Instead, the analysis explicitly considers the magnitude and direction of pre-treatment coefficients, and complements formal testing with visual inspection of the event-study profiles. To make this concern operational, the figures report both the estimated coefficients and an extrapolated counterfactual path based on the pre-treatment trend. This allows a direct comparison between post-2018 dynamics and what would be expected under a continuation of pre-existing trajectories, providing a more informative benchmark than conventional pre-trend tests alone.

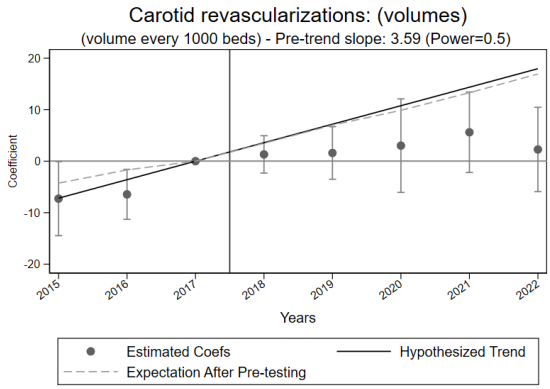
To such extent, the graphs reveal a heterogeneous set of patterns. A first set of outcomes displays clear pre-trends, which mechanically limits causal interpretation. This is particularly evident for carotid revascularizations (Panels a and b of Figure 12), where volumes are already increasing steadily over 2015–2017 and the post-2018 coefficients broadly align with the extrapolated pre-trend, if we except the last observations of the panel.

A second group shows relatively flat or modest pre-trends combined with a visible post-2018 acceleration. This pattern is most apparent for pancreatic cancer resection volumes and overall pancreatic cancer surgery volumes (Panel e of Figure 12 and Panel b of 13), and for mini-invasive lung cancer surgeries (Panel a of 13), where pre-treatment coefficients are close to zero and the post-2018 path displays a clear and sustained increase, which does not reflect completely the extrapolated pre-trend.

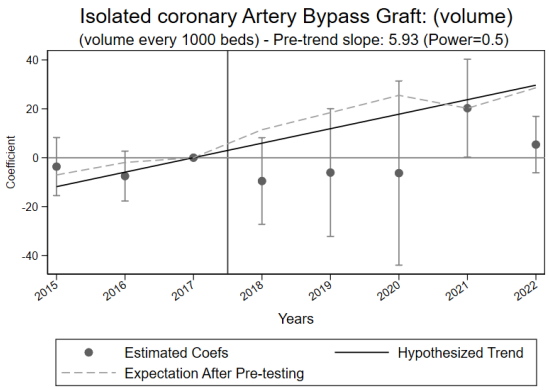
Finally, some outcomes—such as total lung cancer surgery and bypass procedures—are dominated by noise and wide confidence intervals. Point estimates fluctuate around zero in both the pre- and the post-period, and remain imprecise afterward, preventing any clear inference on either trend continuation or structural change.



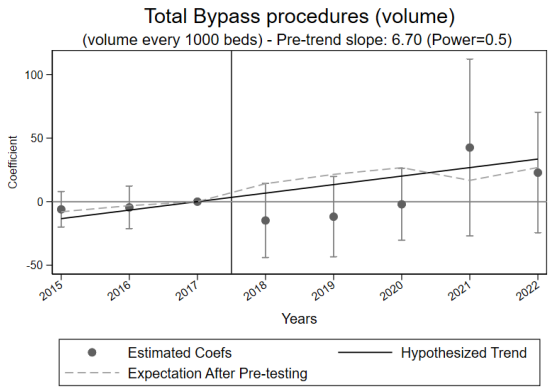
a) Carotid revascularization: endarterectomies



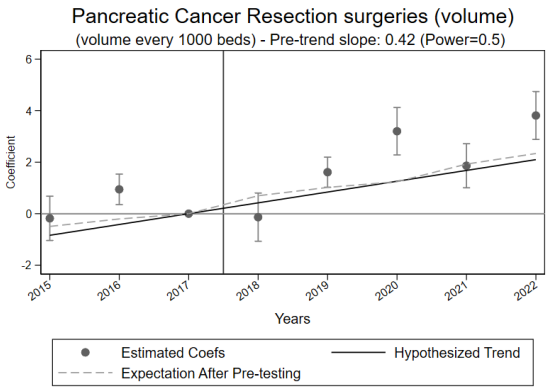
b) Carotid revascularizations (total).



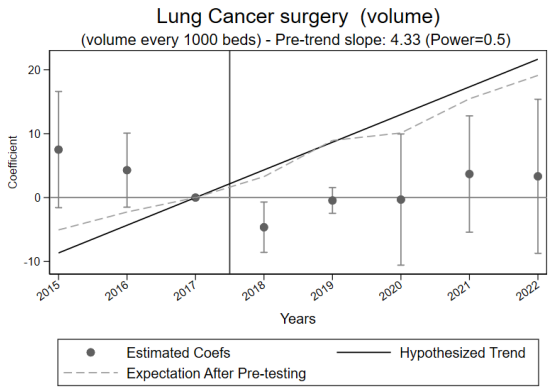
c) Isolated coronary artery bypass graft .



d) Total bypass procedures.

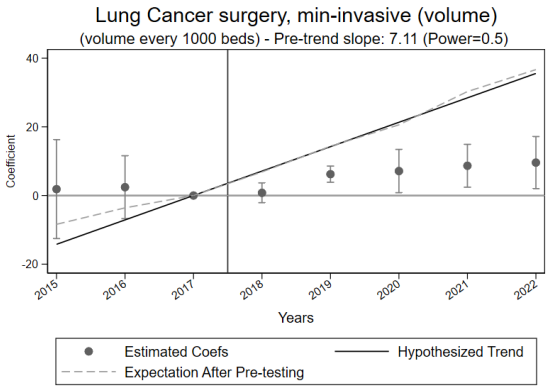


e) Pancreatic cancer resection surgeries.

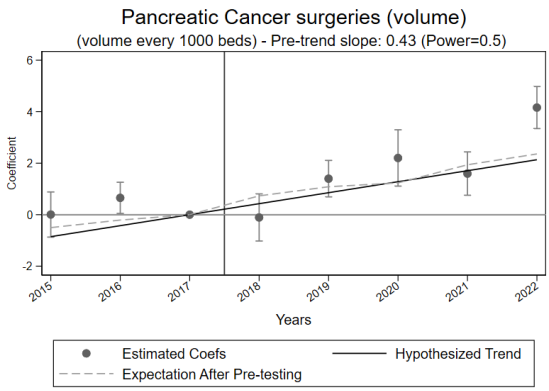


f) Lung cancery surgery procedures.

Figure 12: Event-study estimates of hospital-level PNE outcomes. Each panel reports dynamic treatment effects relative to the year prior to IRCCS recognition (normalized to zero), together with confidence intervals. The additional solid and dashed lines represent the extrapolated counterfactual paths based on the pre-treatment trend estimated through the procedure by (Roth, 2022). The reported "power = 0.5" corresponds to the magnitude of a linear pre-trend that the standard pre-trend test would detect with 50% probability, given the noise and sample size of the data.



a) Lung cancer surgery procedures with mini-invasive approach.



b) Pancreatic surgery procedures.

Figure 13: Event-study estimates of hospital-level PNE outcomes. Each panel reports dynamic treatment effects relative to the year prior to IRCCS recognition (normalized to zero), together with confidence intervals. The additional solid and dashed lines represent the extrapolated counterfactual paths based on the pre-treatment trend estimated through the procedure by (Roth, 2022). The reported "power = 0.5" corresponds to the magnitude of a linear pre-trend that the standard pre-trend test would detect with 50% probability, given the noise and sample size of the data.

9 Conclusions

This paper compares two alternative strategies to stimulate research activity in a health-care institution: performance-based monetary incentives targeted at individual physicians and increased institutional funding through public research support. Using individual-level data from a large Italian hospital, we show that the MBO incentive scheme does not generate a significant increase in research output among non-academic physicians. By contrast, IRCCS recognition and the associated access to public funding lead to a large and persistent increase in the productivity of academic researchers, with positive effects of monetary incentives emerging only when combined with institutional resources.

We provide evidence on the mechanisms underlying these effects. The increase in output is primarily driven by access to stable institutional funding, which relaxes resource constraints and enables sustained research activity. This expansion operates mainly through the intensive margin, with larger and more interconnected research teams, and is concentrated within existing collaboration networks, with limited cross-group spillovers. Consistently, we find that the increase in research output is broad-based across scientific fields, with little evidence of reallocation across areas, suggesting that the policy scales up existing lines of research rather than shifting effort across disciplines.

We also document heterogeneous effects on research quality. While IRCCS recognition increases total citations through higher publication volumes, it is associated with a decline in citations per paper. In contrast, MBO-treated physicians exhibit improvements in citation impact despite no increase in output. These patterns are consistent with a quantity–quality trade-off, rather than with strategic gaming of the incentive scheme.

Our findings have relevant policy implications. Performance-based incentives alone appear insufficient to stimulate research activity when resource constraints are binding, whereas institutional funding plays a central role in enabling research production. At the same time, such funding tends to reinforce existing research structures and does not automatically generate integration across professional groups. More broadly, effective research policies in healthcare settings require a combination of resources and incentives, where funding enables production and incentives shape its direction and quality.

Finally, our analysis focuses on a single research-intensive hospital, and external validity may depend on institutional context and baseline research capacity. Future work should assess the generalizability of these findings and their long-run implications for research organization and performance.

References

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.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Bertoli, P., & Grembi, V. (2019). Malpractice risk and medical treatment selection. *Journal of Public Economics*, *174*, 22–35.

Blasnik, M. (2007). Reclink: Stata module to probabilistically match records. *Statistical Software Components, Boston College Department of Economics*, (S456876).

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.

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.

Chandra, A., & Skinner, J. (2012). Technology Growth and Expenditure Growth in Health Care. *Journal of Economic Literature*, *50*(3), 645–680.

Checchi, D., De Fraja, G., & Verzillo, S. (2021). Incentives and Careers in Academia: Theory and Empirical Analysis. *The Review of Economics and Statistics*, 1–17.

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.

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.

Garisto, D. (2025). Trump moves to slash NSF: Why are the proposed budget cuts so big? *Nature*.

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).

Glenza, J. (2025). Private firms are trying to fill research gaps, but their 'puny' budgets are no match for federal funds. *The Guardian*.

Goodall, A. H. (2011). Physician-leaders and hospital performance: Is there an association? *Social Science & Medicine (1982)*, *73*(4), 535–539.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

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.

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.

Highman, L., Marginson, S., & Papatsiba, V. (2023). Higher education and research: Multiple negative effects and no new opportunities after Brexit. *Contemporary Social Science*, *18*(2), 216–234.

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.

Marchandot, B., & Morel, O. (2025). The Financialization of Healthcare in France: Trends and implications. *Public Health in Practice*, *9*, 100620.

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.

Peel, M., Bott, I., & Kuchler, H. (2025). The cost of Trump's attack on American science. *Financial Times*.

Rambachan, A., & Roth, J. (2023). A More Credible Approach to Parallel Trends. *Review of Economic Studies*, *90*(5), 2555–2591.

Romano, J. P., & Wolf, M. (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica*, *73*(4), 1237–1282.

Romano, J. P., & Wolf, M. (2016). Efficient computation of adjusted p<math><mi is="true">p</mi></math>-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, *113*, 38–40.

Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. *American Economic Review: Insights*, *4*(3), 305–322.

Rousseau, S., Catalano, G., & Daraio, C. (2021). Can we estimate a monetary value of scientific publications? *Research Policy*, *50*(1), 104116.

Shekhtman, L. M., Gates, A. J., & Barabási, A.-L. (2024). Mapping philanthropic support of science. *Scientific Reports*, *14*(1), 9397.

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.

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.

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.

10 Appendix

Target	MDs without faculty structured appointments
Object	Full papers published between Jan 1 and Dec 31 (each year)
Payment per IF point	€500 per IF point
Max reimbursable IF	20 IF points per year
Max annual gross	€10,000
Author role (institution affiliation)	**Bonus share**
First or only author from the institution	100%
First author with other institutional co-authors	60%
Co-author on a paper first-authored by an institutional member	40%
Co-author on a paper first-authored externally	50%

- **IF** = Impact Factor of the Journal.

Table A1: Description of the MBO-rewarding scheme.

Requirement	Description	Evaluator
Scientific excellence	Proven research productivity in a specific biomedical area	Scientific Committee (Ministry of Health)
Healthcare quality	High-level and specialized clinical care	Ministry of Health / AGENAS
Integration	Strong integration between research and healthcare activities	Ministry of Health / AGENAS
Qualified personnel	Presence of qualified scientific and healthcare staff	Ministry of Health
Infrastructure	Adequate facilities and technological resources	AGENAS
External evaluation	Periodic assessment of all requirements	Ministry of Health
Recognition validity	Five-year renewable accreditation	Ministry of Health

Table A2: Description of the IRCCS' requirement to achieve recognition (Source: DL 288/2003).

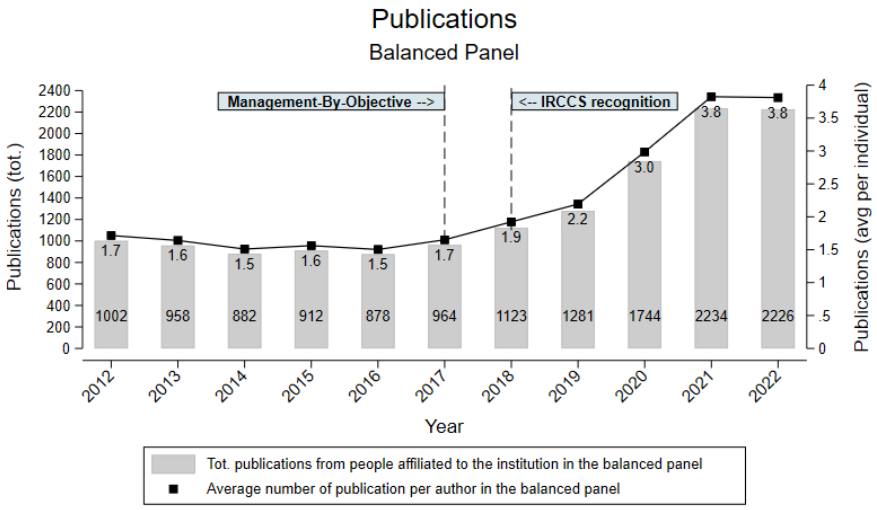


Figure A1: Total number of publications attributed to employees of the institution in the balanced sample (2012–2022).

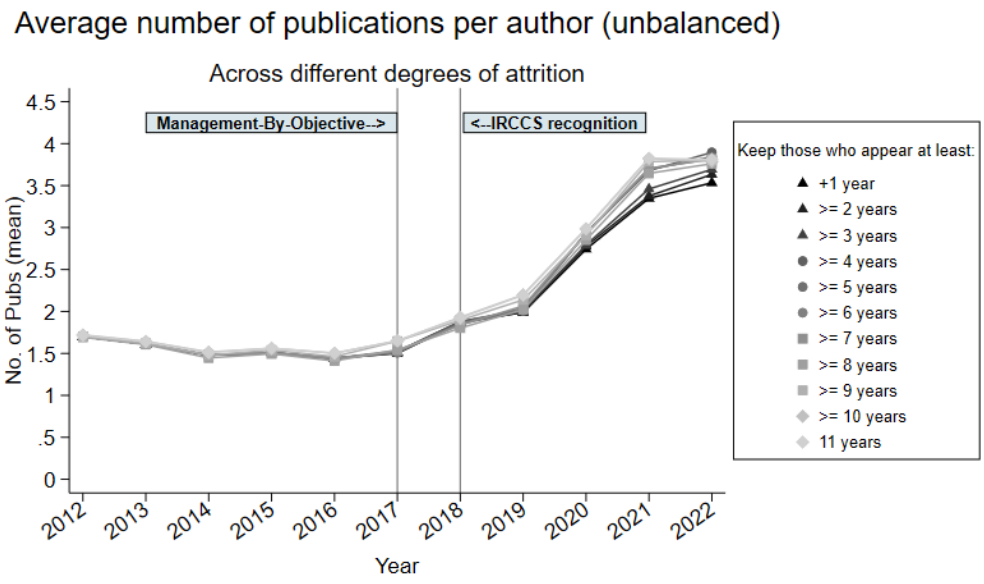


Figure A2: Average publications per researcher per year (2012–2022) across different degrees of panel attrition. Each line represents the mean trend obtained after progressively excluding individuals who are observed for fewer years in the dataset.

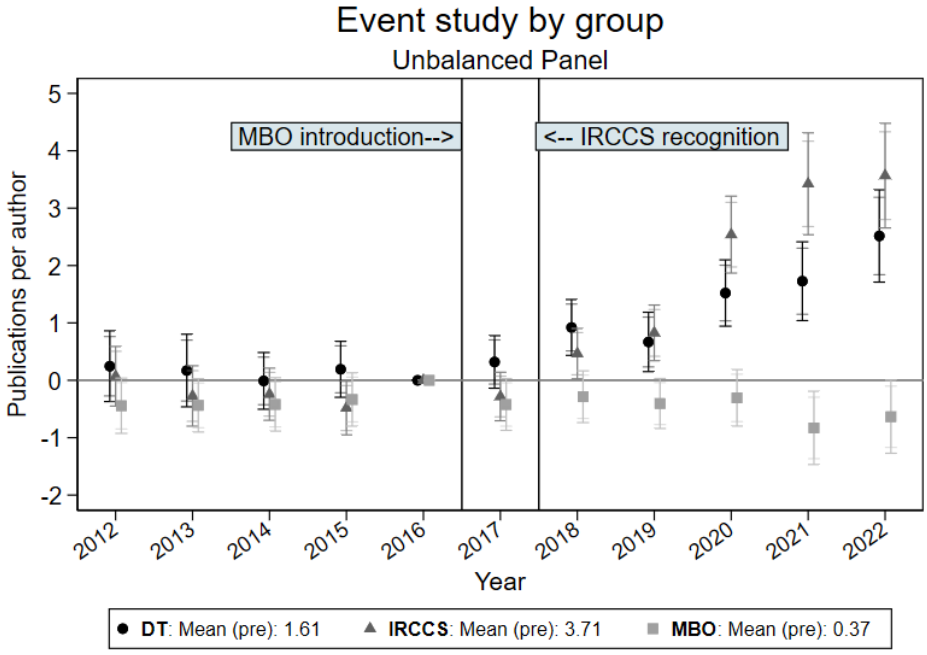
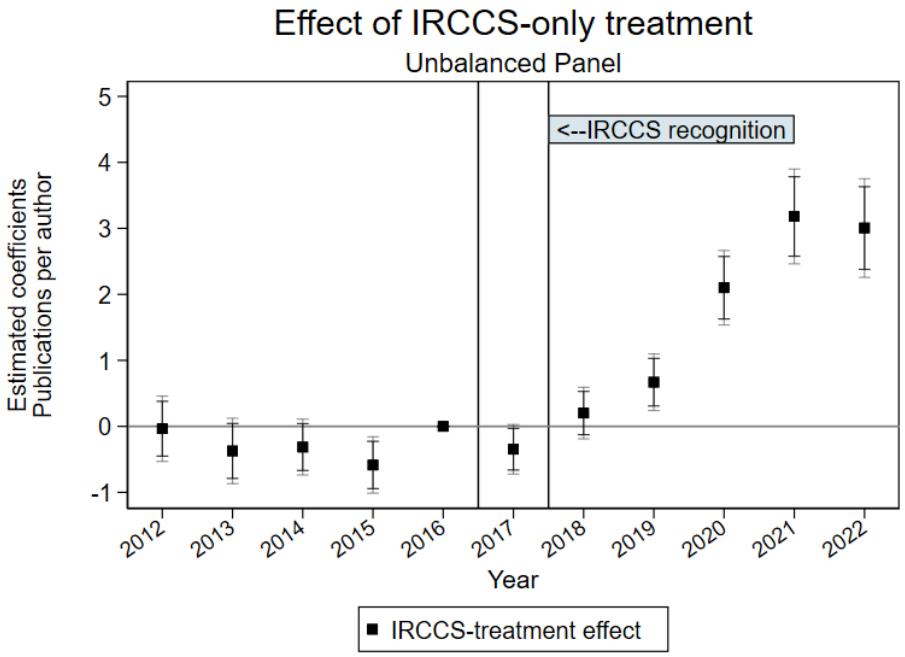


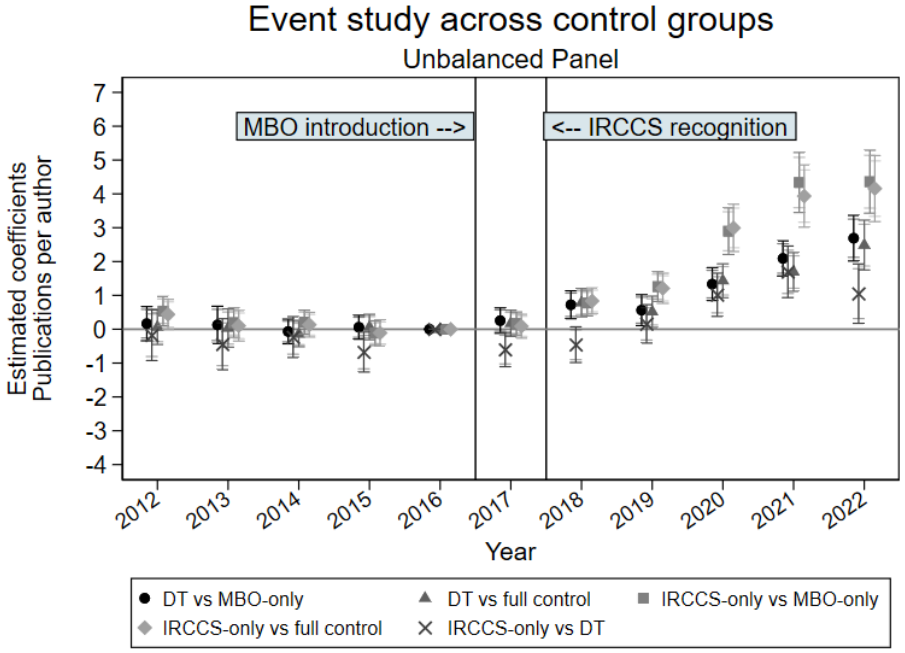
Figure A3: Event-study of the MBO joint with IRCCS recognition policy effect on publications across the different groups, and with respect to the full control group. - Full unbalanced panel.

	(1) IRCCS vs. others	(2) DT vs MBO	(3) DT vs PC	(4) IRCCS vs MBO	(5) IRCCS vs PC	(6) IRCCS vs DT
Publications	2.0587***	1.2984***	1.2281***	2.4704***	2.3267***	0.9739***
(SE)	(0.2407)	(0.1659)	(0.1719)	(0.2665)	(0.2640)	(0.2555)
N	10739	5487	4417	6296	5226	4445
R^2	0.772	0.660	0.647	0.802	0.796	0.761
Time FE	YES	YES	YES	YES	YES	YES
Individual FE	YES	YES	YES	YES	YES	YES
Mean	3.73	1.69	1.69	1.69	3.73	3.73
Panel (Unbalanced)	Full	DT and MBO	DT and control	IRCCS and MBO	IRCCS and control	IRCCS and DT
Time Range	2012–2022	2012–2022	2012–2022	2012–2022	2012–2022	2012–2022

Table A3: Impact of IRCCS recognition on Annual Publications in the unbalanced panel (Difference-in-Differences).

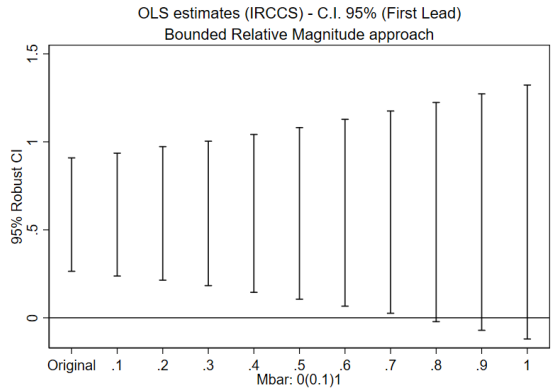


a) Event-study of IRCCS recognition effect on publications for IRCCS-only treated individuals compared with all the other units.

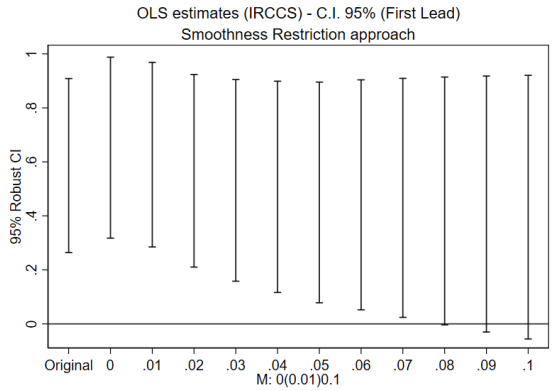


b) Event-study of IRCCS policy effect on publications for comparisons across different groups.

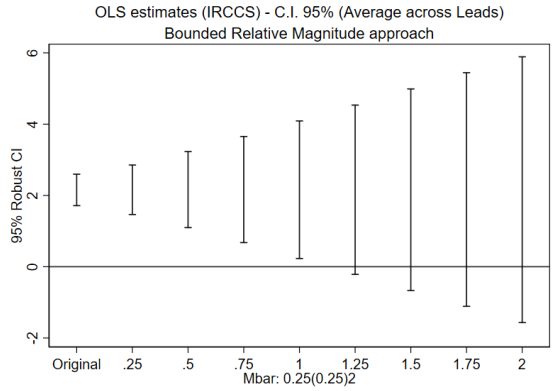
Figure A4: Event-studies of IRCCS policy effect for the unbalanced panel.



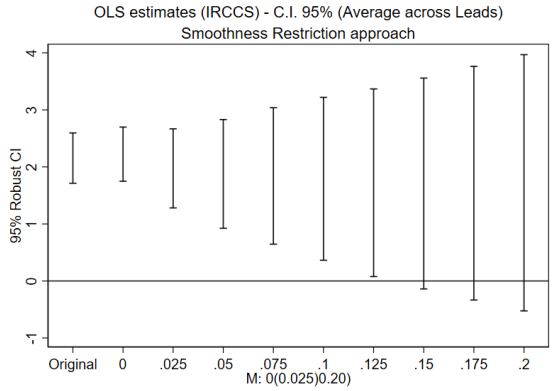
a) BM approach for the significance of the first lead.



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

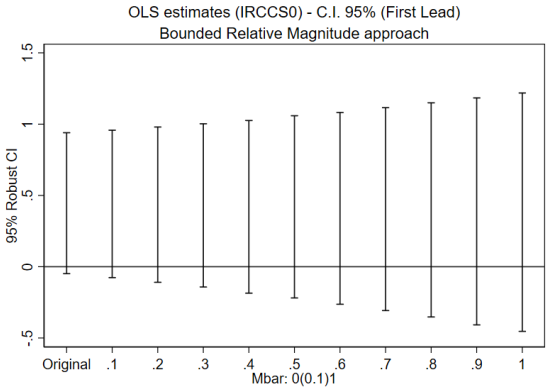


c) BM approach for the significance of the average across all leads.

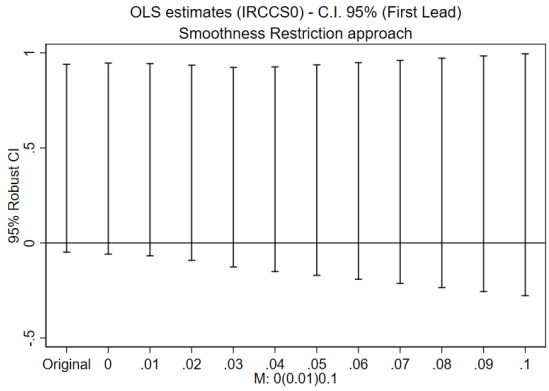


d) SR approach for the significance of the average across all leads.

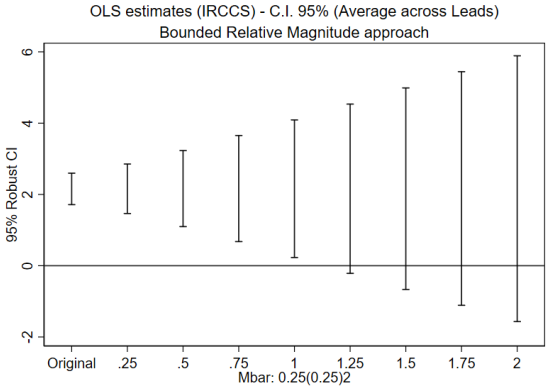
Figure A5: Honest DiD robust confidence sets for overall IRCCS effect estimated with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average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



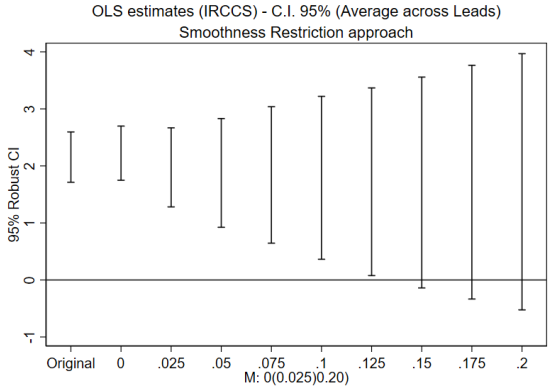
a) BM approach for the significance of the first lead.



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

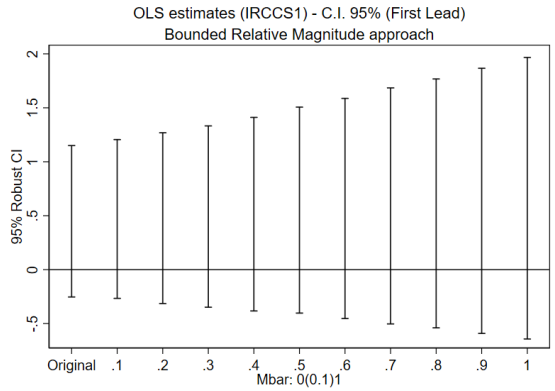


c) BM approach for the significance of the average across all leads.

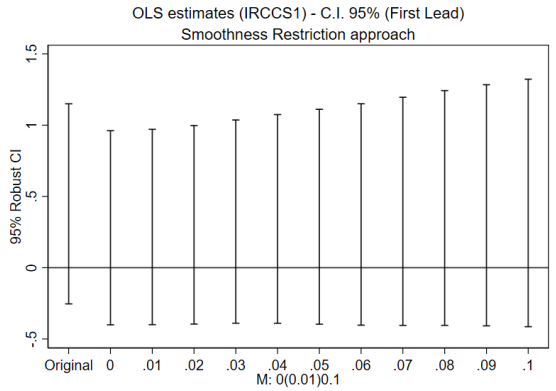


d) SR approach for the significance of the average across all leads.

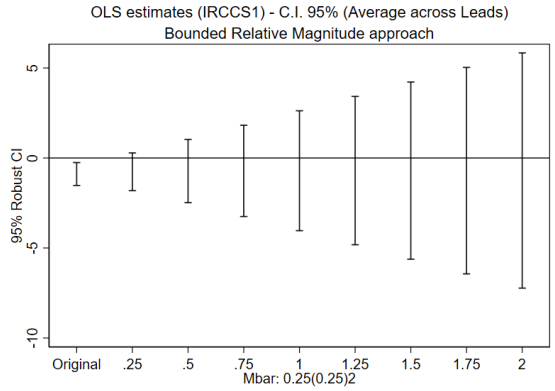
Figure A6: Honest DiD robust confidence sets in the comparison between double-treated and MBO-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of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



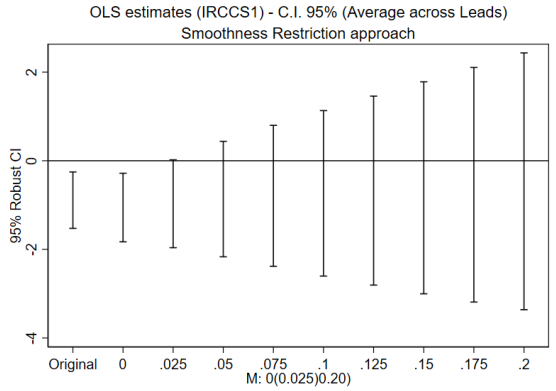
a) BM approach for the significance of the first lead.



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

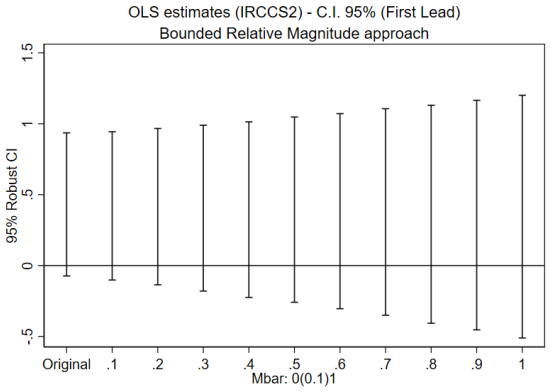


c) BM approach for the significance of the average across all leads.

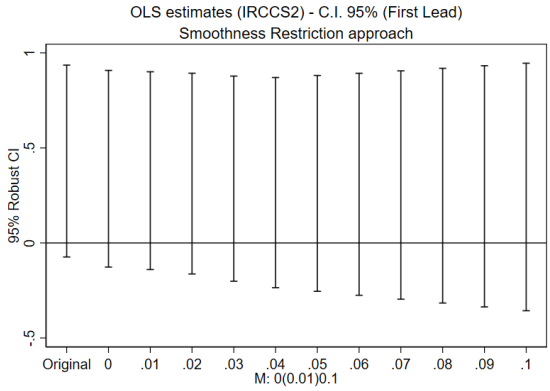


d) SR approach for the significance of the average across all leads.

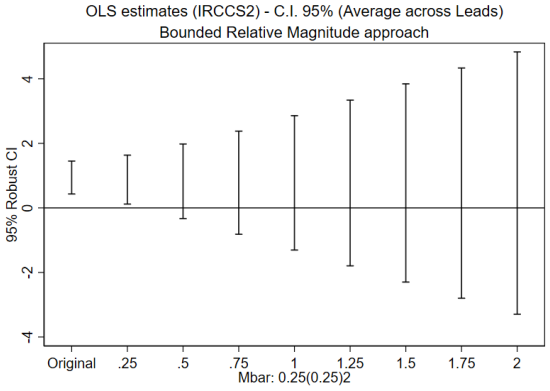
Figure A7: Honest DiD robust confidence sets in the comparison between double-treated and IRCCS-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



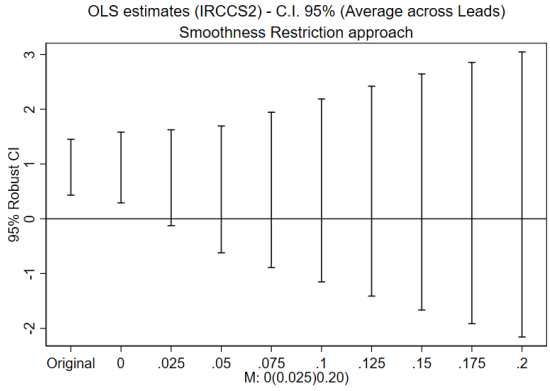
a) BM approach for the significance of the first lead.



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

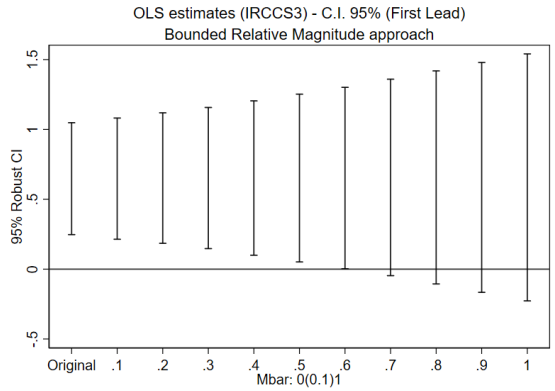


c) BM approach for the significance of the average across all leads.

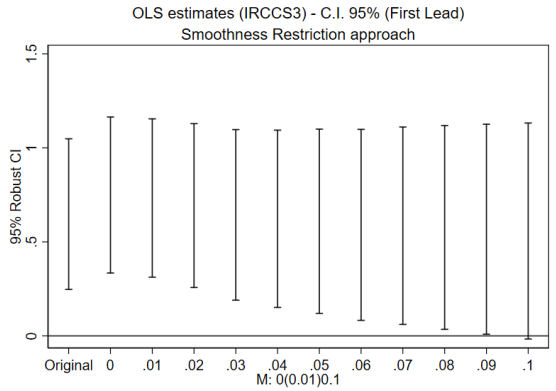


d) SR approach for the significance of the average across all leads.

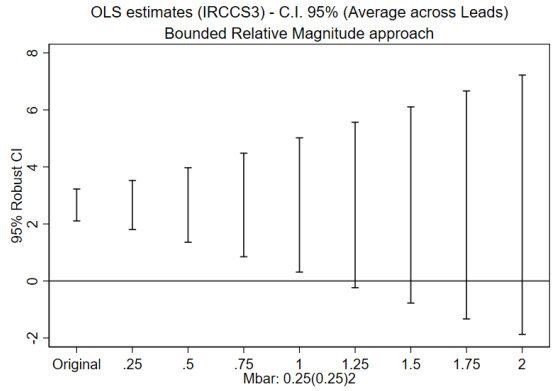
Figure A8: Honest DiD robust confidence sets in the comparison between double-treated and pure control units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



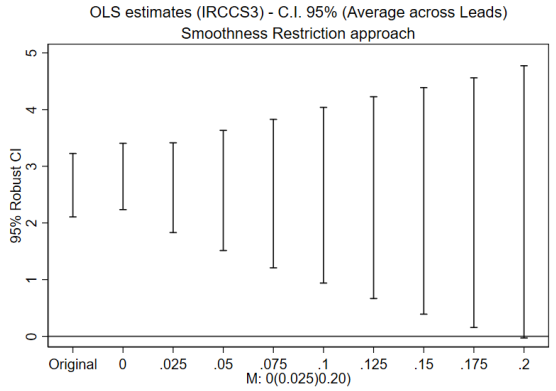
a) BM approach for the significance of the first lead.



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

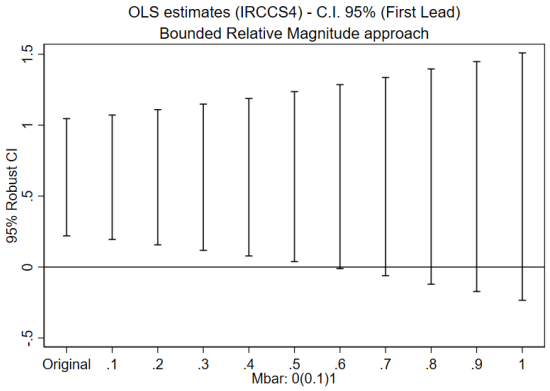


c) BM approach for the significance of the average across all leads.

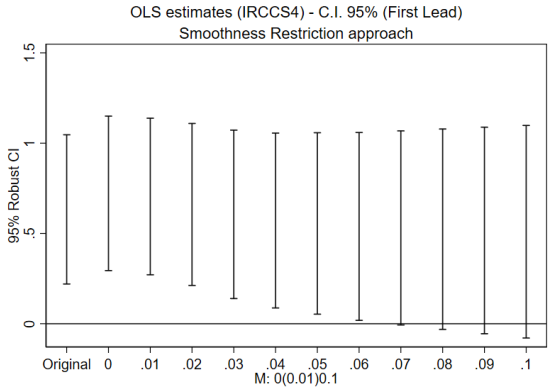


d) SR approach for the significance of the average across all leads.

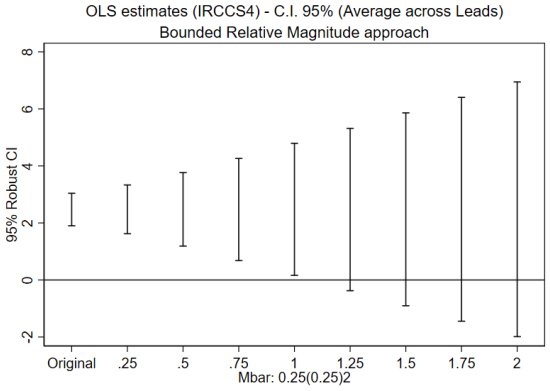
Figure A9: Honest DiD robust confidence sets in the comparison between IRCCS-only and MBO-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



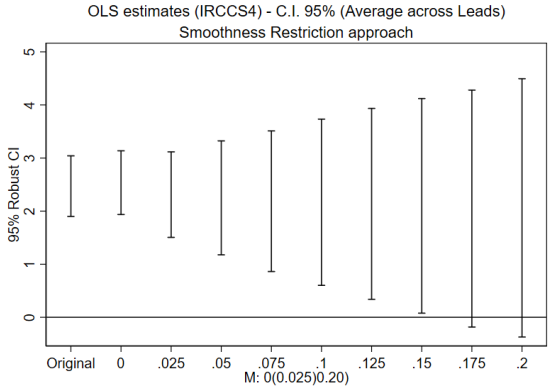
a) BM approach for the significance of the first lead.



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

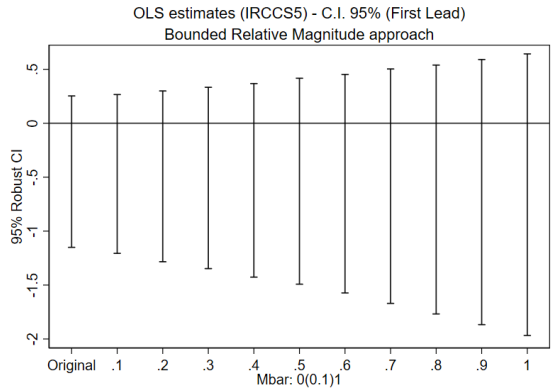


c) BM approach for the significance of the average across all leads.

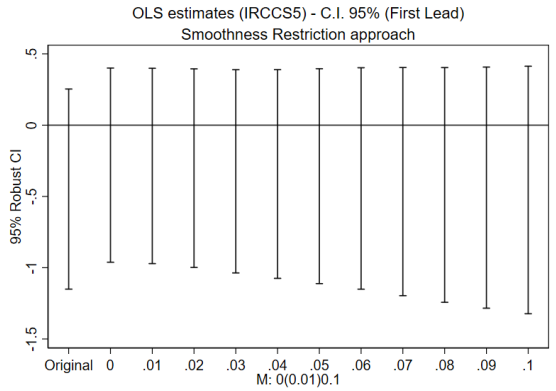


d) SR approach for the significance of the average across all leads.

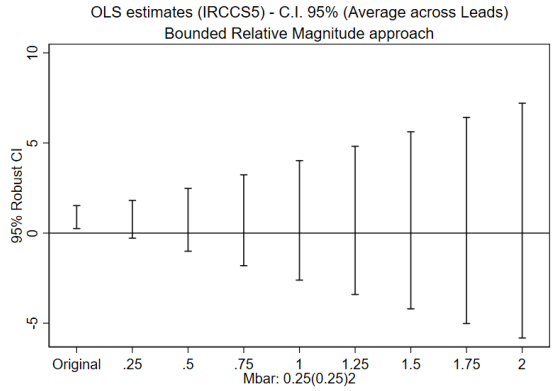
Figure A10: Honest DiD robust confidence sets in the comparison between IRCCS-only and pure control units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



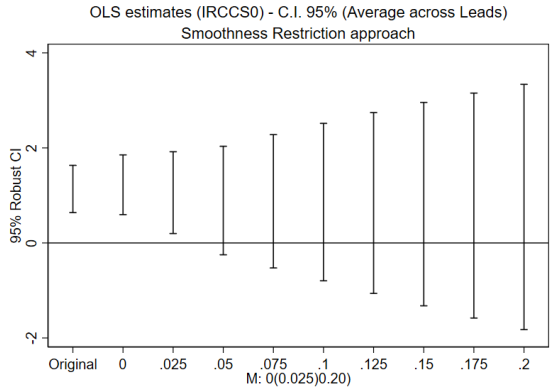
a) BM approach for the significance of the first lead.



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 A11: Honest DiD robust confidence sets in the comparison between IRCCS-only and double-treated 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)).

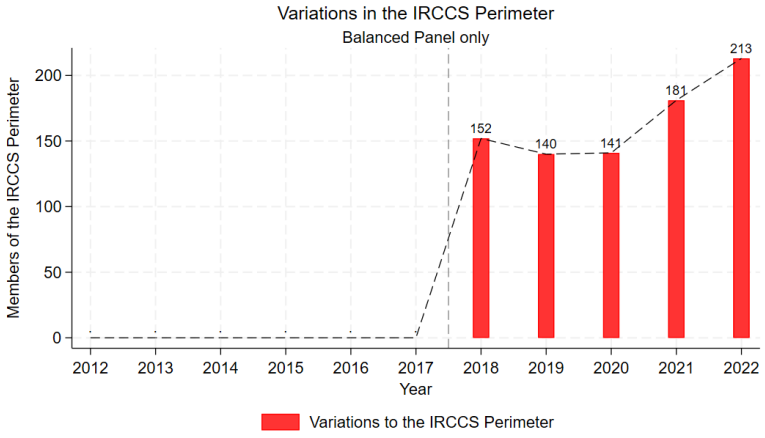


Figure A12: Evolution of the units "lately selected" into the IRCCS perimeter.

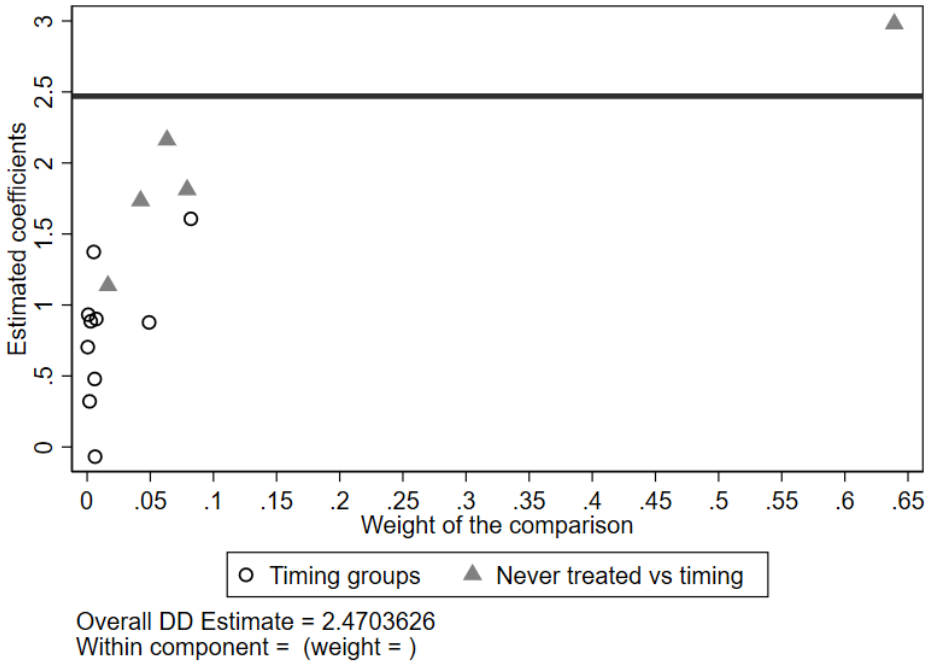


Figure A13: Bacon Decomposition of TWFE Estimate

	Beta	Weight
Timing groups	1.219	0.16
Never-treated vs timing	2.687	0.84
Overall effect	2.452	1

Figure A14: Goodman-Bacon decomposition: weights and estimates.

	N	Percent		N	Percent
IRCCS First - MBO	20	13.16	IRCCS Late - MBO	33	35.48
IRCCS First - No MBO	132	86.84	IRCCS Late - No MBO	60	64.52
Total	152	100	Total	93	100

Table A4: Composition of the IRCCS group, and distinction between "first-" and "late-" adopters.

	Core IRCCS	Late IRCCS	p-value
Age	49.4 (9.9)	46.1 (9.0)	0.001***
Publications	3.8 (4.7)	1.5 (1.9)	0.000***
Citations	150.0 (292.4)	42.2 (113.5)	0.000***
	Categorical variables (column %)		
Gender			
Female	28.2%	50.0%	
Male	63.5%	46.6%	
Status			
Medical Director	0.0%	0.0%	
Healthcare Professions Manager	37.6%	54.1%	
Sanitary Director	0.0%	0.0%	
Faculty Clinician	2.6%	3.4%	
Department			
Diagnostic Imaging	17.7%	10.1%	
Emergency & Anesthesiology	6.4%	6.1%	
Cardiovascular Sciences	9.4%	12.8%	
Women, Children & Public Health	16.5%	23.6%	
Lab Sci & Infectiology	7.1%	8.1%	
Gastro/Nephro/Endo	21.8%	11.5%	
Aging & Neurosciences	16.9%	27.0%	
Medical & Surgical Sciences	0.0%	0.0%	
Directorate General	0.0%	0.0%	
Clinical Governance	1.5%	0.0%	
Health Governance	2.6%	0.7%	

Table A5: Difference in characteristics (mean, SD in parentheses, and p-value of the t-test of thed iff. in means) and composition (%) of the group of researchers included in the IRCCS perimeter in 2018 (early) and those enrolled later.

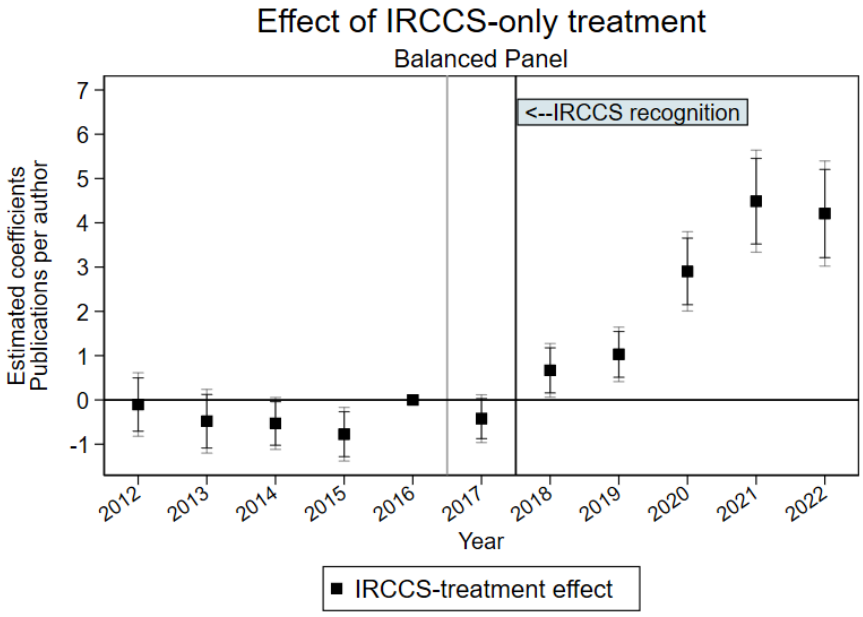


Figure A15: Event-study of IRCCS recognition effect on publications. Units included in the IRCCS perimeter in any year after 2018 are excluded from the estimates.

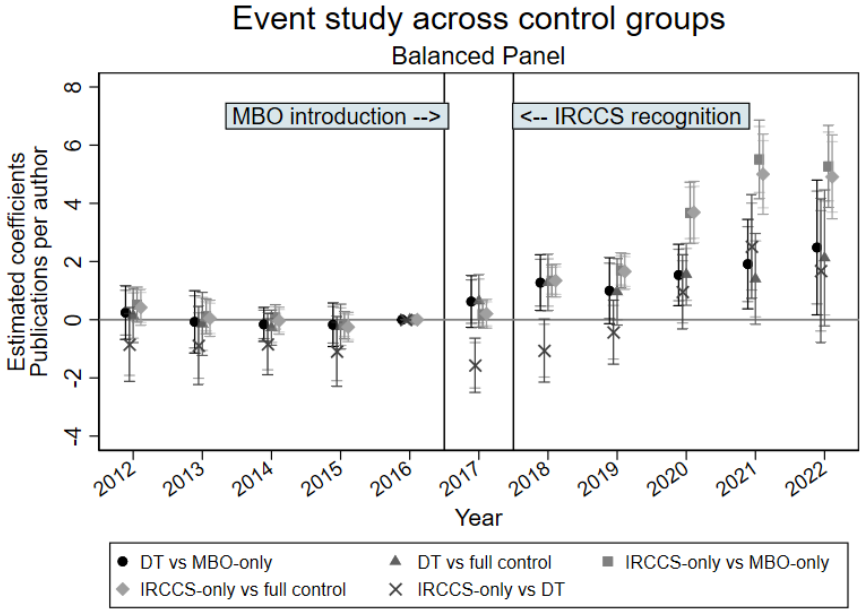


Figure A16: Event-study of IRCCS policy effect on publications for comparisons across different groups. Units included in the IRCCS perimeter in any year after 2018 are excluded from the estimates.

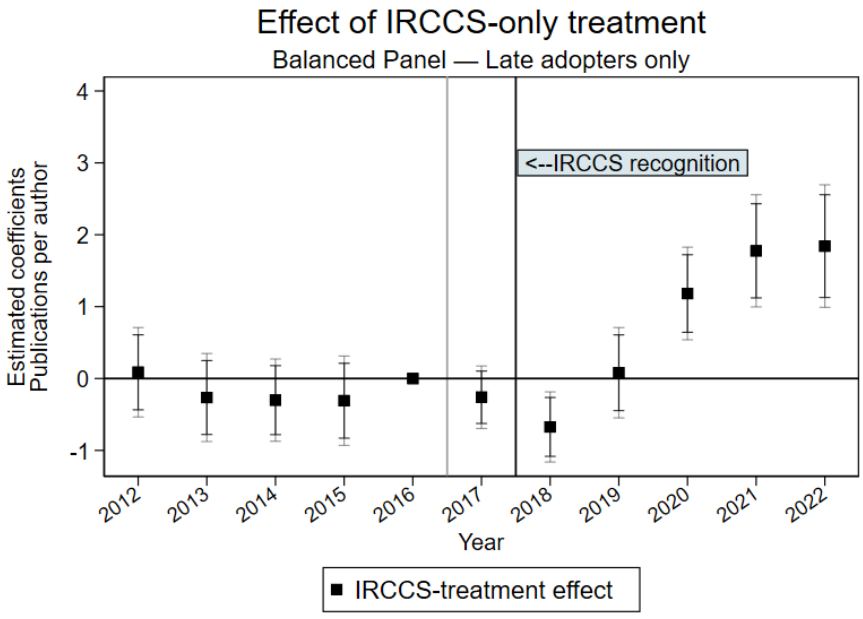


Figure A17: Event-study of IRCCS recognition effect on publications of late-enrolled. Units included in the IRCCS perimeter in 2018 are excluded from the estimates.

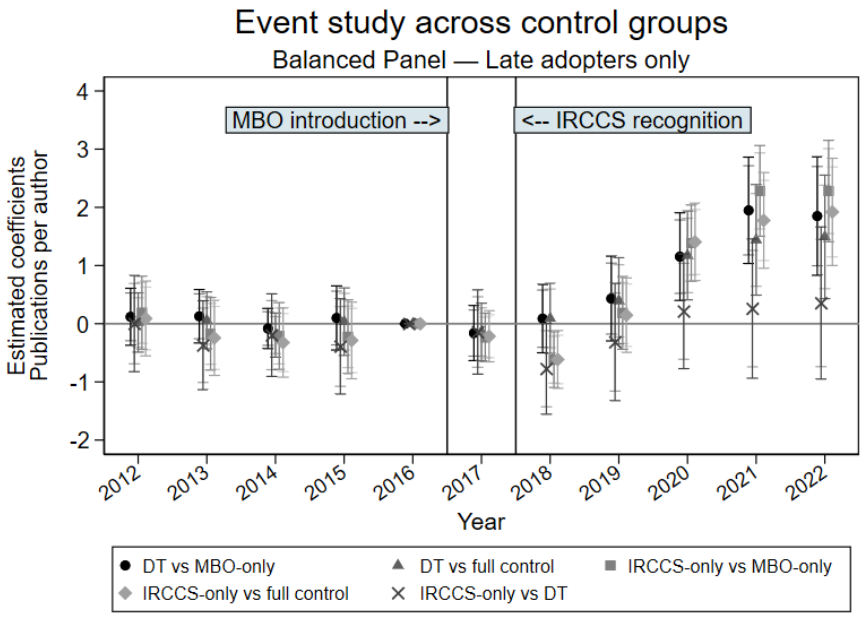
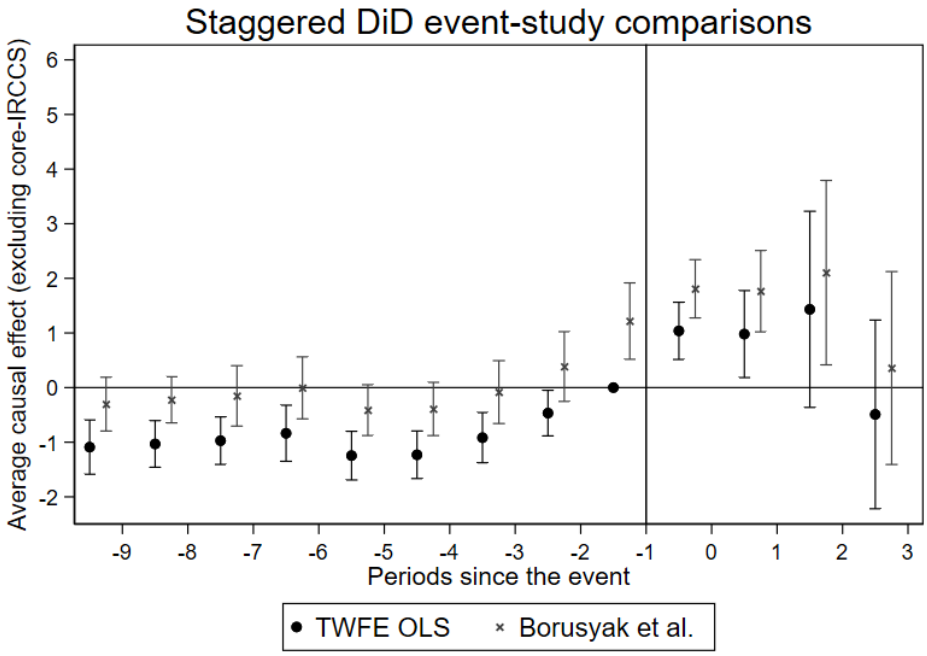
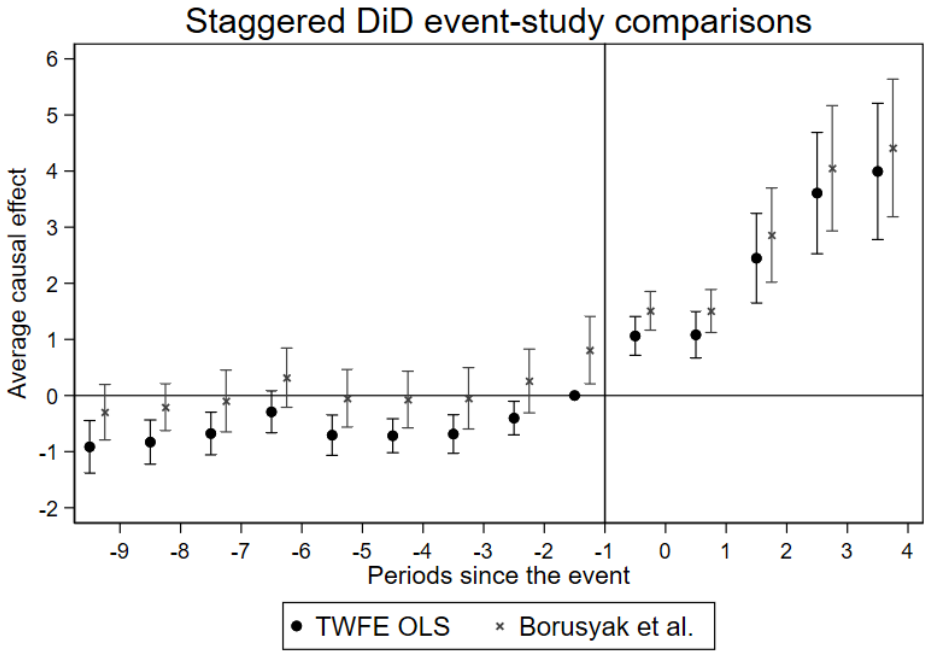


Figure A18: Event-study of IRCCS policy effect on publications late-enrolled for comparisons across different groups. Units included in 2018 are excluded from the estimates.

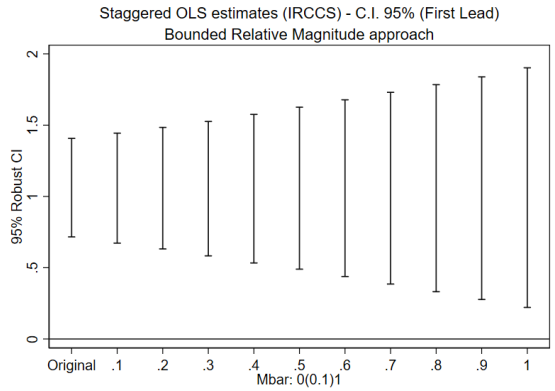


a) Event-study of the staggered inclusion in the IRCCS perimeter of late-enrolled only, excluding the "core" researchers treated in 2018.

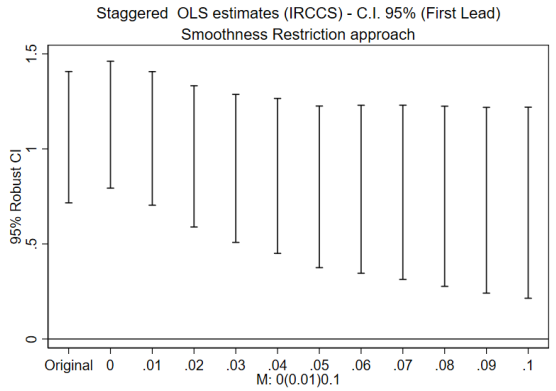


b) Event-study of the staggered inclusion in the IRCCS perimeter of all treated units, including the "core" researchers treated in 2018.

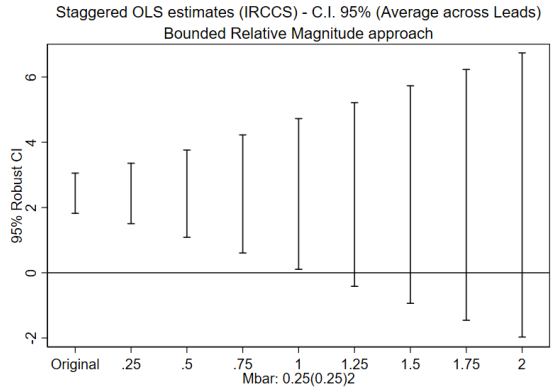
Figure A19: Event-study of IRCCS policy effect on publications for comparisons across different groups.



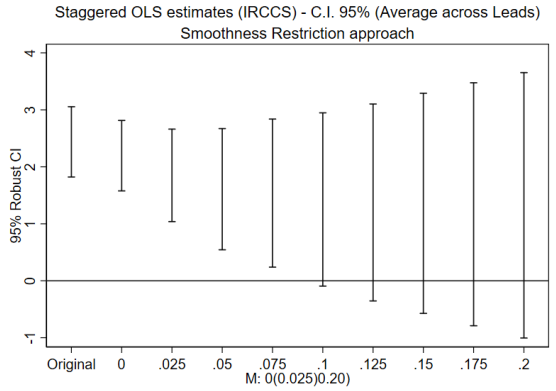
a) BM approach for the significance of the first lead.



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

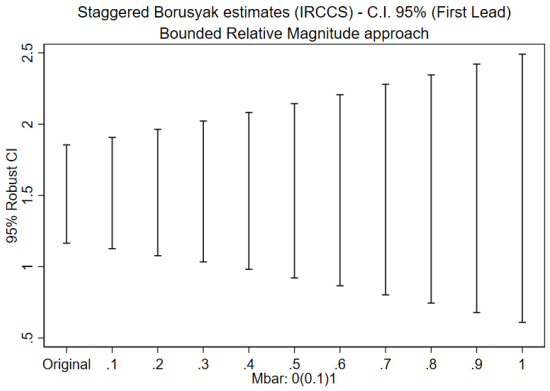


c) BM approach for the significance of the average across all leads.

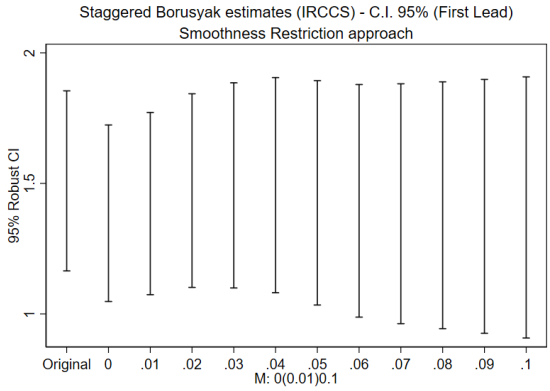


d) SR approach for the significance of the average across all leads.

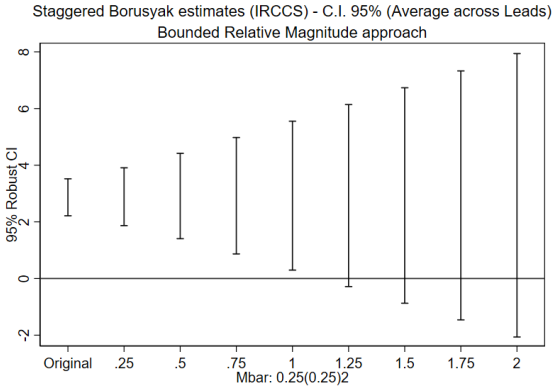
Figure A20: Honest DiD robust confidence sets in the OLS staggered event-study estimates on the IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



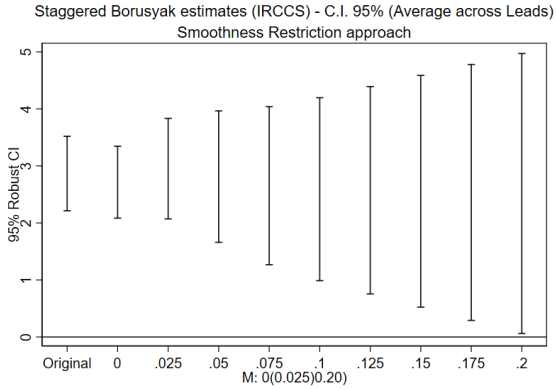
a) BM approach for the significance of the first lead.



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



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

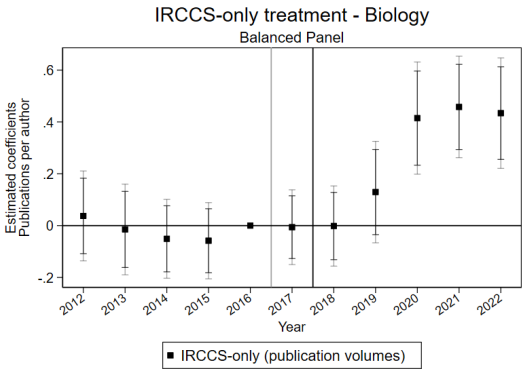


d) SR approach for the significance of the average across all leads.

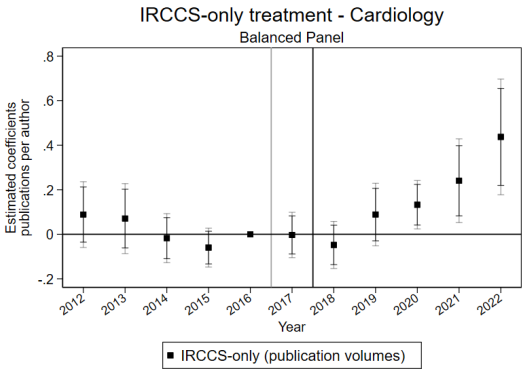
Figure A21: Honest DiD robust confidence sets in the staggered event-study estimates obtained via the procedure by Borusyak et al., 2024 on the IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).

Macro-area	WoS categories included
Biology	Biochemistry; Biotechnology; Cell Biology; Genetics; Life Sciences & Biomedicine; Microbiology; Biophysics; Developmental Biology; Evolutionary Biology; Physiology; Agriculture; Fisheries; Veterinary Sciences; Zoology
Technology/Engineering	Medical Laboratory Technology; Instruments and Instrumentation; Engineering; Computer Science; Mechanics; Metallurgy and Metallurgical Engineering; Microscopy; Nuclear Science and Technology; Robotics; Science and Technology - Other Topics; Food Science
Oncology/Hematology	Oncology; Hematology
Cardiology/Cardiovascular Diseases	Cardiology
Neuroscience/Psychiatry/Psychology	Neurosciences; Psychiatry; Psychology
Internal Medicine (Generic)	General and Internal Medicine; Emergency Medicine; Integrative and Complementary Medicine; Research and Experimental Medicine
Surgery/Procedural	Surgery; Anesthesiology; Orthopedics; Otorhinolaryngology; Urology; Transplantation; Ophthalmology; Dentistry
Diagnostics/Imaging	Radiology; Pathology; Medical Laboratory Technology
Mother/Child	Obstetrics; Pediatrics; Reproductive Biology
Infectious Diseases/Immuno-Allergology	Infectious Diseases; Virology; Parasitology; Tropical Medicine; Mycology; Immunology; Allergy
Chemistry/Physics	Chemistry; Electrochemistry; Crystallography; Materials Science; Physics; Polymer Science; Spectroscopy; Acoustics; Optics
Public Health/Health Systems	Health Care Sciences and Services; Public, Environmental and Occupational Health; Business and Economics; Public Administration; Biomedical Social Sciences; Environmental Sciences
Social Sciences and Humanities	Behavioral Sciences; Legal Medicine; Medical Ethics; Sociology; Social Sciences - Other Topics; Family Studies; Educational Research; Religion; Criminology; Anthropology; History and Philosophy of Science
Rehabilitation/Care	Rehabilitation; Nursing; Sport Sciences; Audiology
Pharmacology	Pharmacology; Toxicology; Substance Abuse
Internal Medicine (Specialties)	Gastroenterology; Endocrinology; Respiratory System; Rheumatology; Geriatrics; Dermatology; Nutrition and Dietetics

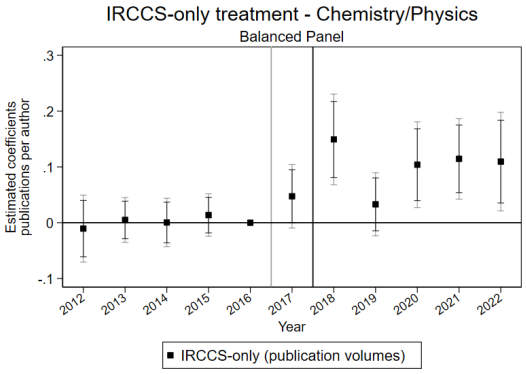
Table A6: Classification of research areas into macro-categories based on Web of Science (WoS) subject categories.



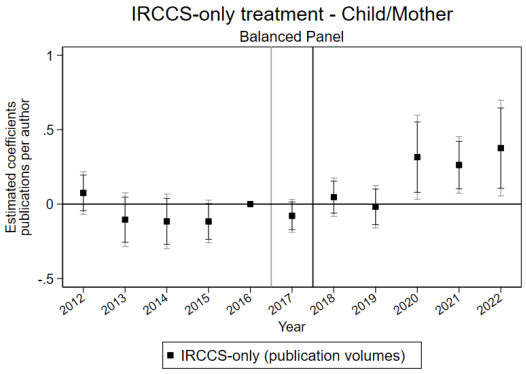
(a) Biology



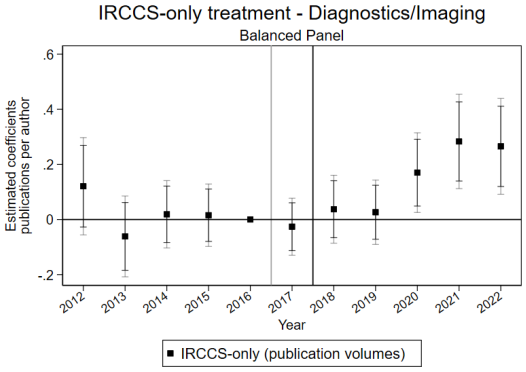
(b) Cardiology and Cardiovascular Systems



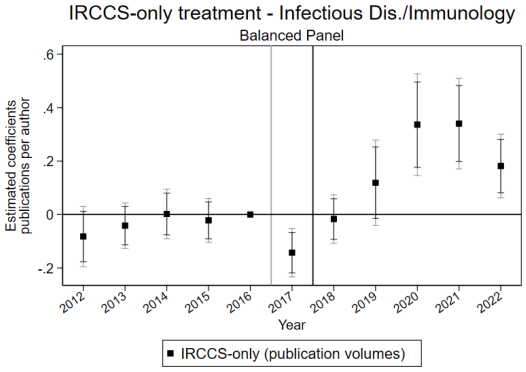
(c) Chemistry and Physics



(d) Child and Maternal Health

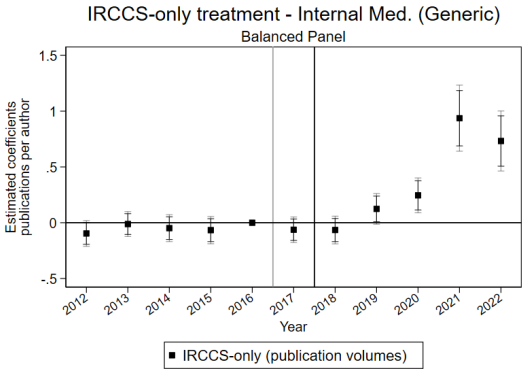


(e) Diagnostics and Imaging

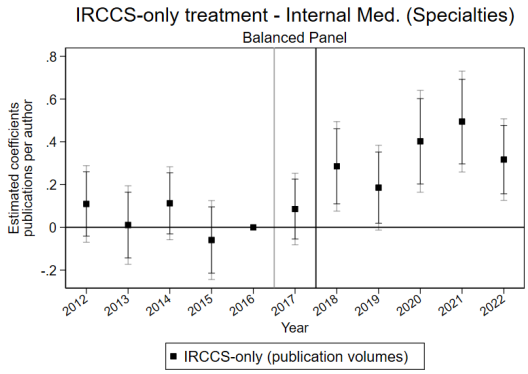


(f) Infectious Diseases, Immunology, Allergology

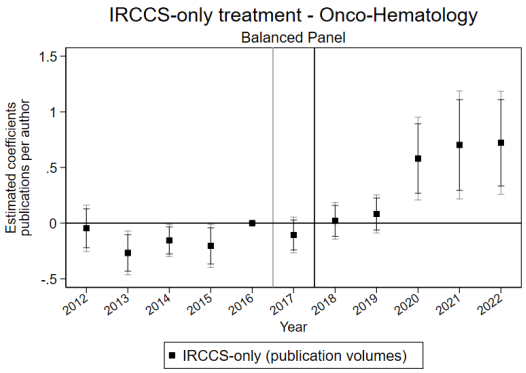
Figure A22: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the number individual publications in a specific area.



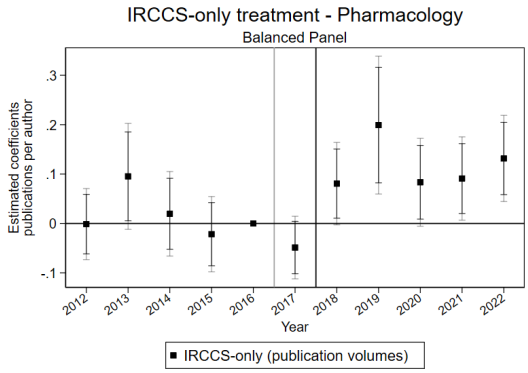
(a) Internal Medicine (General and Emergency Medicine)



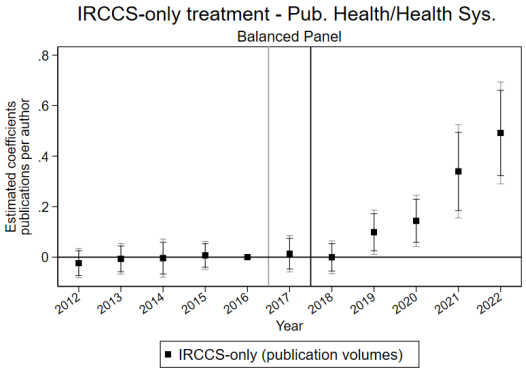
(b) Internal Medicine (Specialties)



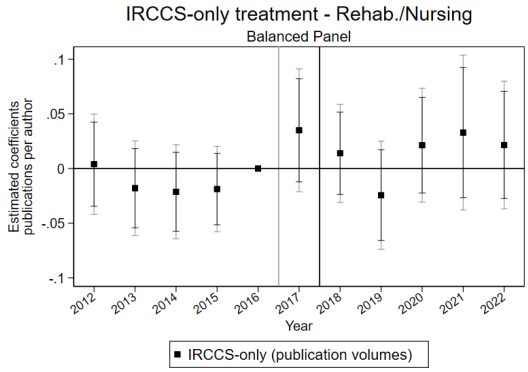
(c) Oncology and Hematology



(d) Pharmacology and Toxicology

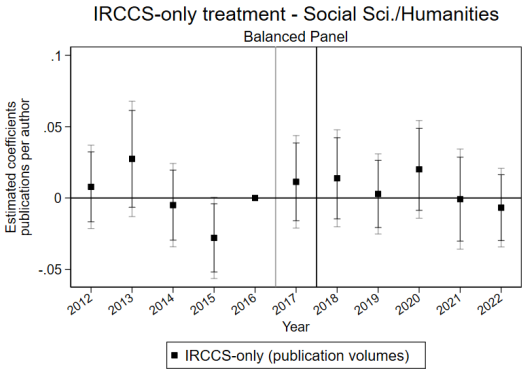


(e) Public Health and Health Systems

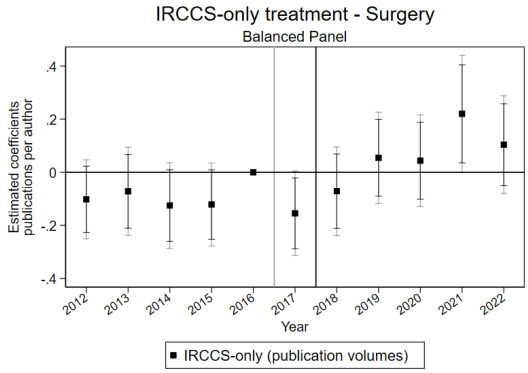


(f) Rehabilitation and Nursing

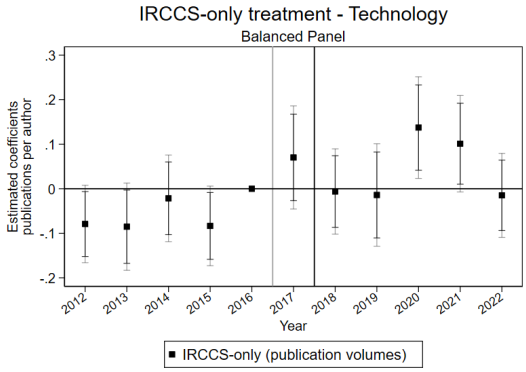
Figure A23: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the number individual publications in a specific area.



(a) Social Sciences and Humanities



(b) Surgery, Dentistry, Procedures

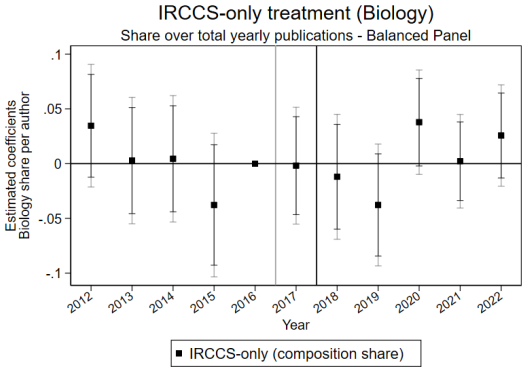


(c) Technology, Engineering, Computer Science

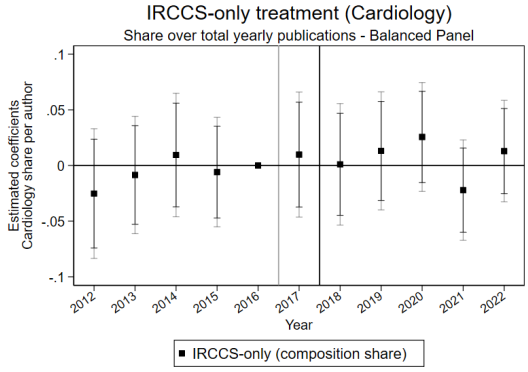
Figure A24: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the number individual publications in a specific area.

Specialization	ITT			Pre (joint)		Mean	Note
	Coef.	p-value	RW p-val.	F-stat	p-value		
Biology	0.003	0.790	0.998	0.324	0.862	0.047	0
	(0.013)						
Cardiology/Cardiovasc. Diseases	-0.001	0.934	0.998	1.209	0.306	0.041	0
	(0.013)						
Chemistry and Physics	0.008	0.207	0.713	0.302	0.876	0.006	0
	(0.007)						
Diagnostics and Imaging	-0.005	0.611	0.996	0.723	0.576	0.041	0
	(0.011)						
Infectious Dis./ Immunology	-0.004	0.702	0.998	2.837	0.024**	0.025	?
	(0.011)						
Internal Medicine (Generic)	0.013	0.426	0.983	0.425	0.791	0.046	0
	(0.016)						
Internal Medicine (Specialties)	0.002	0.891	0.998	1.602	0.172	0.083	0
	(0.016)						
Maternal and Child Health	0.004	0.682	0.998	2.226	0.065*	0.039	?
	(0.010)						
Neurosciences and Psych.	-0.023	0.140	0.533	2.264	0.061*	0.069	?
	(0.016)						
Oncology and Hematology	-0.004	0.778	0.998	0.230	0.922	0.061	0
	(0.014)						
Pharmacology	-0.007	0.500	0.984	0.989	0.413	0.021	0
	(0.010)						
Public Health and Health Systems	0.008	0.158	0.580	0.690	0.599	0.011	0
	(0.006)						
Rehabilitation and Nursing	-0.006	0.442	0.983	1.051	0.380	0.011	0
	(0.008)						
Social Sciences and Humanities	0.002	0.751	0.998	2.154	0.073*	0.006	?
	(0.006)						
Surgery and Procedures	-0.032*	0.061	0.245	0.251	0.909	0.073	0
	(0.017)						
Technology and Engineering	-0.009	0.438	0.983	0.605	0.659	0.020	0
	(0.011)						

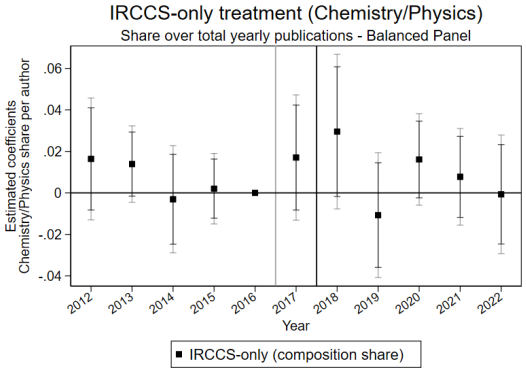
Table A7: Each row reports results from a separate regression of the yearly share of publications in a given specialization on a post-treatment indicator for IRCCS-only researchers, controlling for unit and year fixed effects. Standard errors, reported in parentheses beneath the ITT coefficient, are clustered at the author level. The block Pre (joint) reports the F-test statistic and p-value from a joint test of all pre-treatment leads (2012–2015). Romano–Wolf stepdown adjusted p-values are reported to account for multiple hypothesis testing across the family of specialization outcomes. Depending on the variance estimator and finite-sample correction, the reported test statistic may be an F statistic or a chi-squared statistic. Mean pre-treatment share is computed over the pre-2017 period. Significance: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.



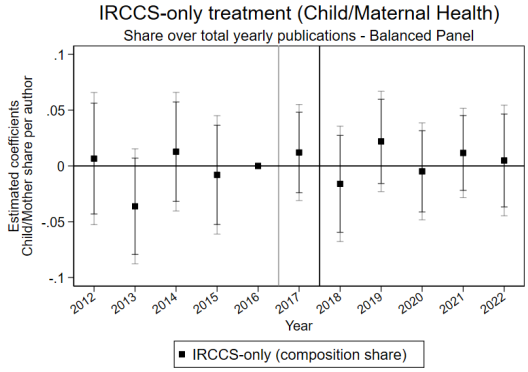
(a) Biology



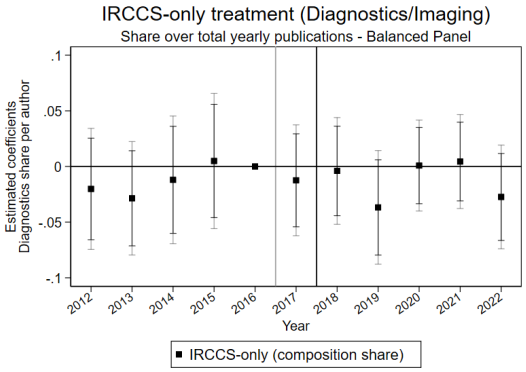
(b) Cardiology and Cardiovascular Systems



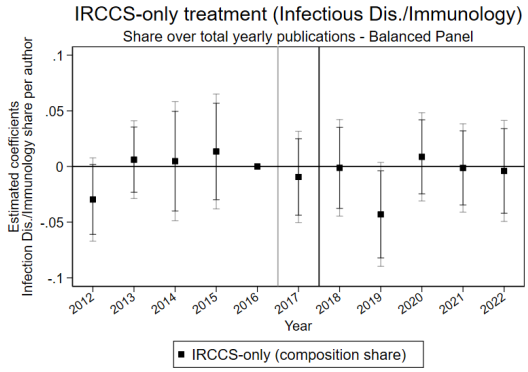
(c) Chemistry and Physics



(d) Child and Maternal Health

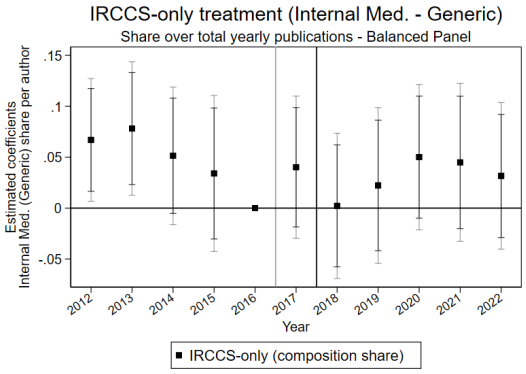


(e) Diagnostics and Imaging

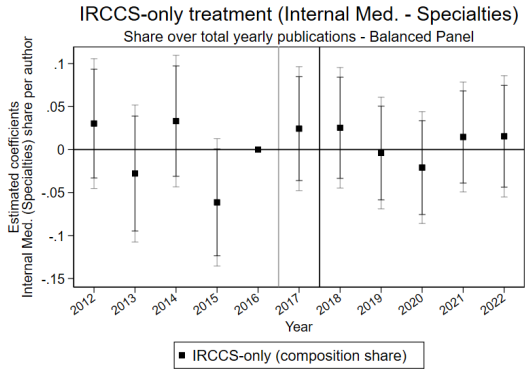


(f) Infectious Diseases, Immunology, Allergology

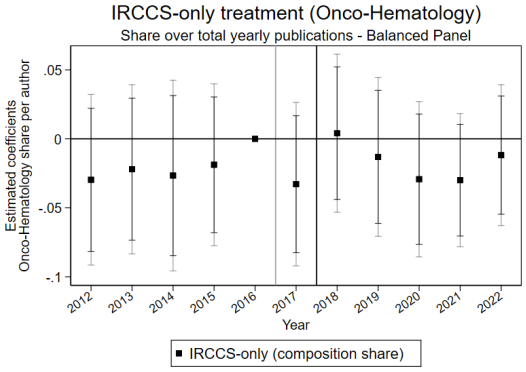
Figure A25: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the share of individual publications in a specific area over the total number of one's annual publications.



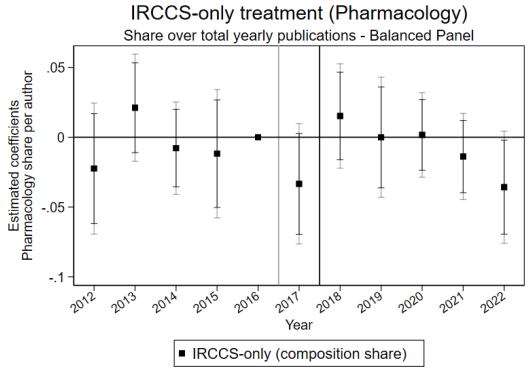
(a) Internal Medicine (General and Emergency Medicine)



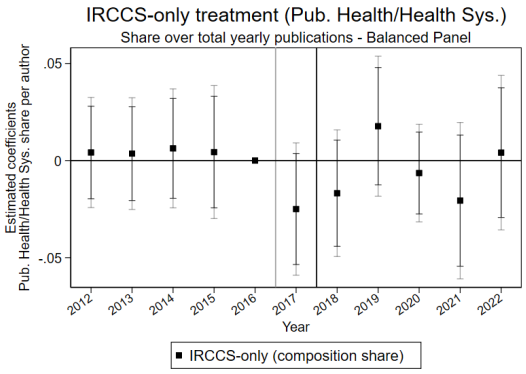
(b) Internal Medicine (Specialties)



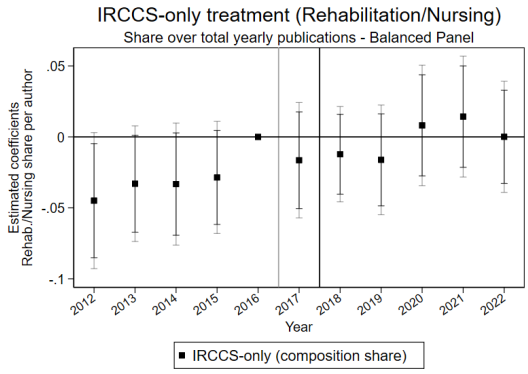
(c) Oncology and Hematology



(d) Pharmacology and Toxicology

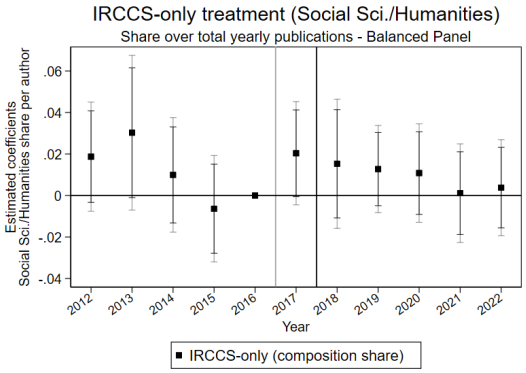


(e) Public Health and Health Systems

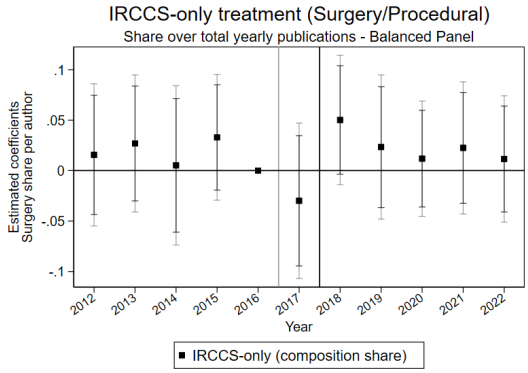


(f) Rehabilitation and Nursing

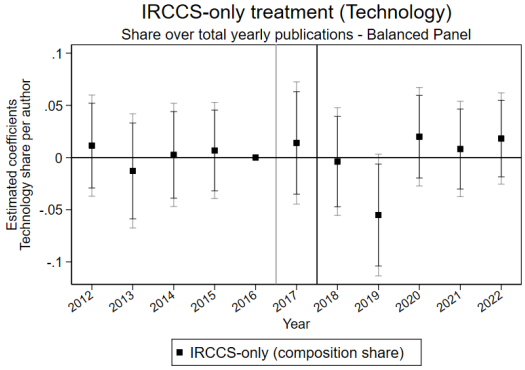
Figure A26: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the share of individual publications in a specific area over the total number of one's annual publications.



(a) Social Sciences and Humanities



(b) Surgery, Dentistry, Procedures



(c) Technology, Engineering, Computer Science

Figure A27: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the share of individual publications in a specific area over the total number of one's annual publications.

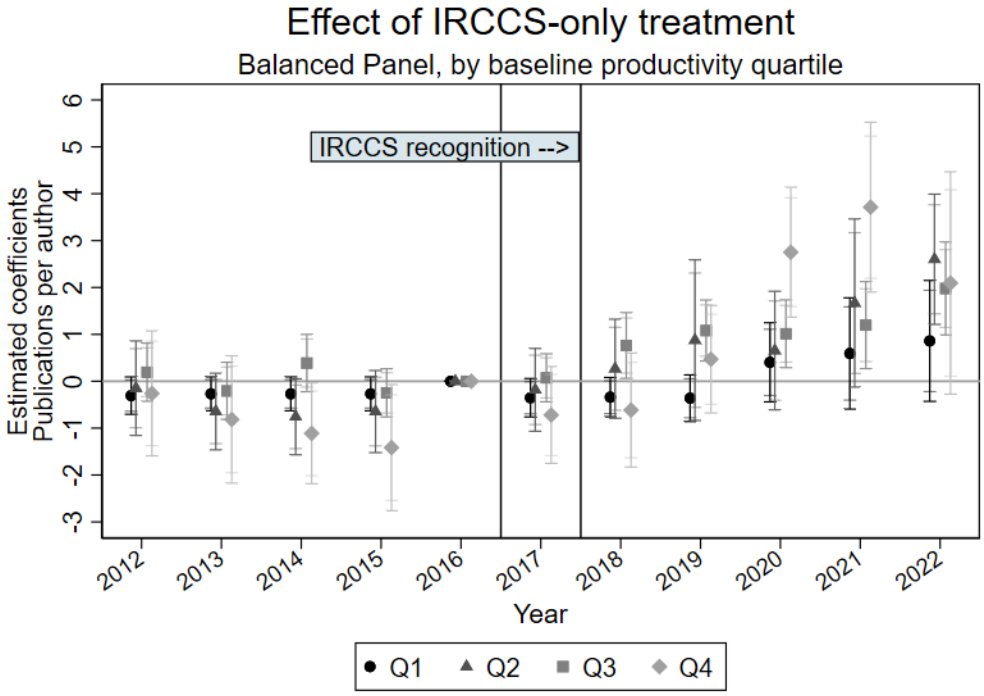


Figure A28: Event-study of the IRCCS effect by baseline productivity quartiles (balanced panel, including those who never published before 2017).

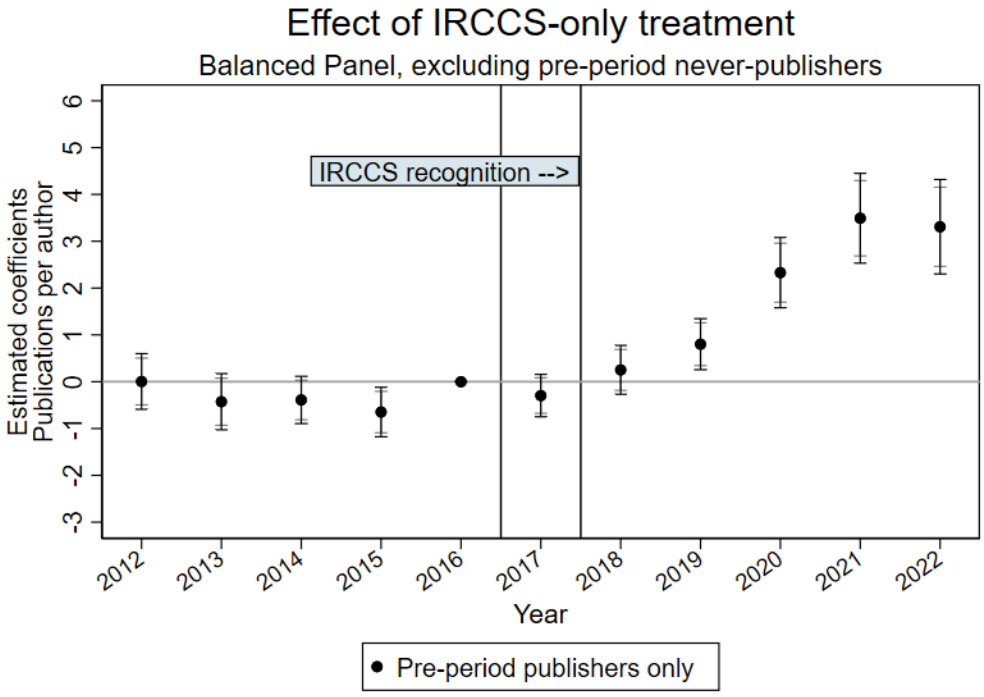


Figure A29: Event-study of the IRCCS effect excluding pre-treatment never-publishers (balanced panel).

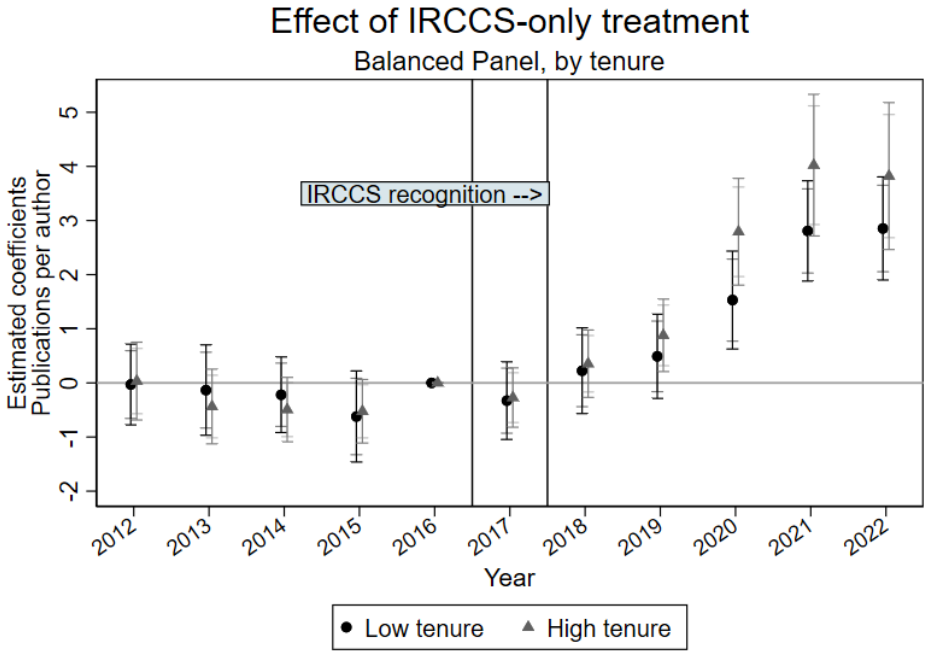


Figure A30: Event-study of the IRCCS effect by tenure (balanced panel). Tenure is defined based on the tenure distribution at the time of treatment, splitting individuals into below- and above-median groups.

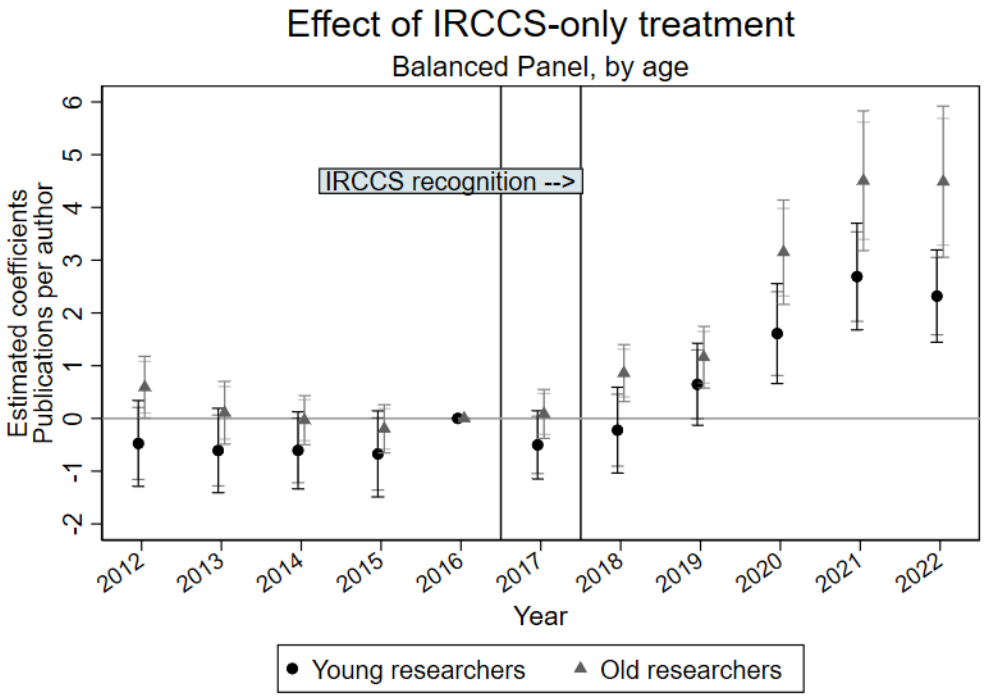


Figure A31: Event-study of the IRCCS effect by age (balanced panel). Age is defined based on the age distribution at the time of treatment, splitting individuals into below- and above-median groups.

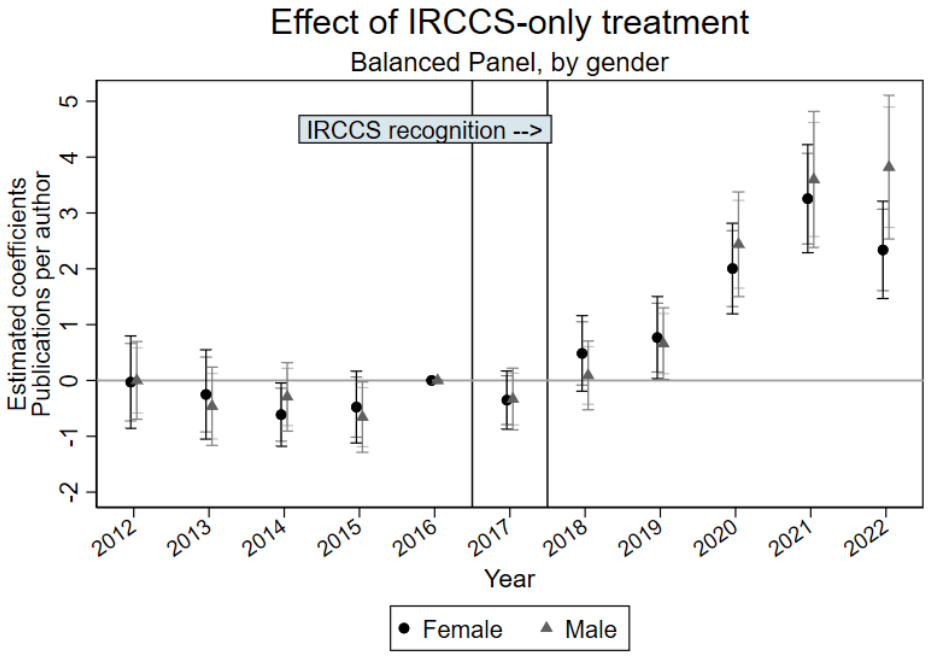


Figure A32: Event-study of the IRCCS effect by gender (balanced panel).

	(1) IRCCS vs. others	(2) DT vs MBO	(3) DT vs IRCCS	(4) DT vs PC	(5) IRCCS vs MBO	(6) IRCCS vs PC
Publications (log)	0.2691***	0.3027***	0.0011	0.2902***	0.3324***	0.3175***
(SE)	(0.0337)	(0.0587)	(0.0577)	(0.0594)	(0.0388)	(0.0400)
Publications (asinh)	0.3198***	0.3841***	0.0214	0.3691***	0.3997***	0.3814***
(SE)	(0.0425)	(0.0754)	(0.0735)	(0.0763)	(0.0492)	(0.0507)
N	6424	2623	2739	2430	3987	3794
R^2	0.752	0.605	0.726	0.588	0.802	0.774
Time FE	YES	YES	YES	YES	YES	YES
Individual FE	YES	YES	YES	YES	YES	YES
Panel	Full	DT and MBO	DT and IRCCS	DT and control	IRCCS and MBO	IRCCS and control
Time Range	2012–2022	2012–2022	2012–2022	2012–2022	2012–2022	2012–2022

Table A8: Impact of IRCCS recognition on Annual Publications, by changing the outcome to log-transformation and inverse hyperbolic sine (Difference-in-Differences).

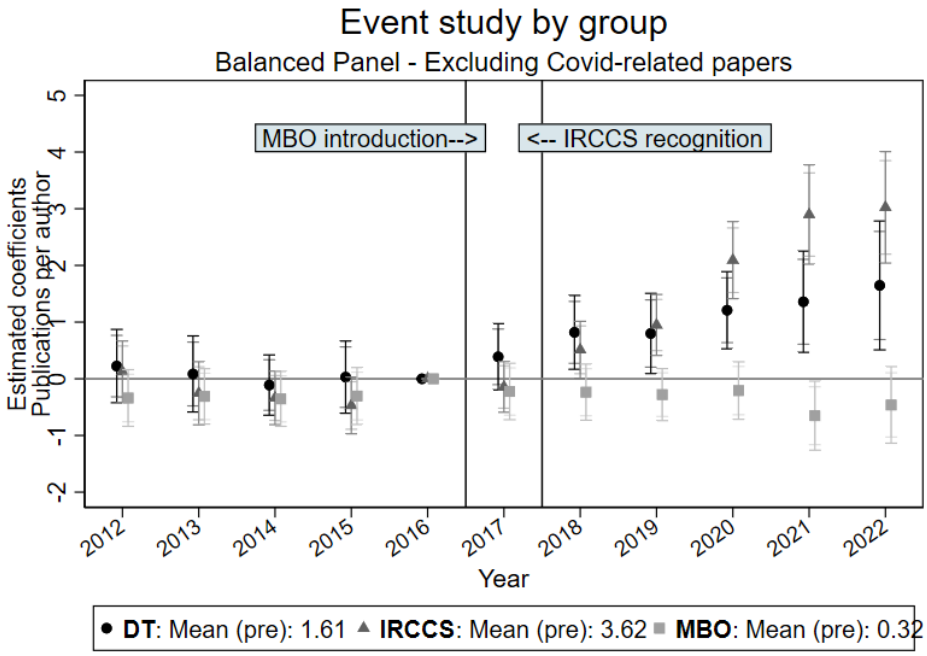


Figure A33: Event-study of the IRCCS after dropping COVID-related publications (balanced panel).

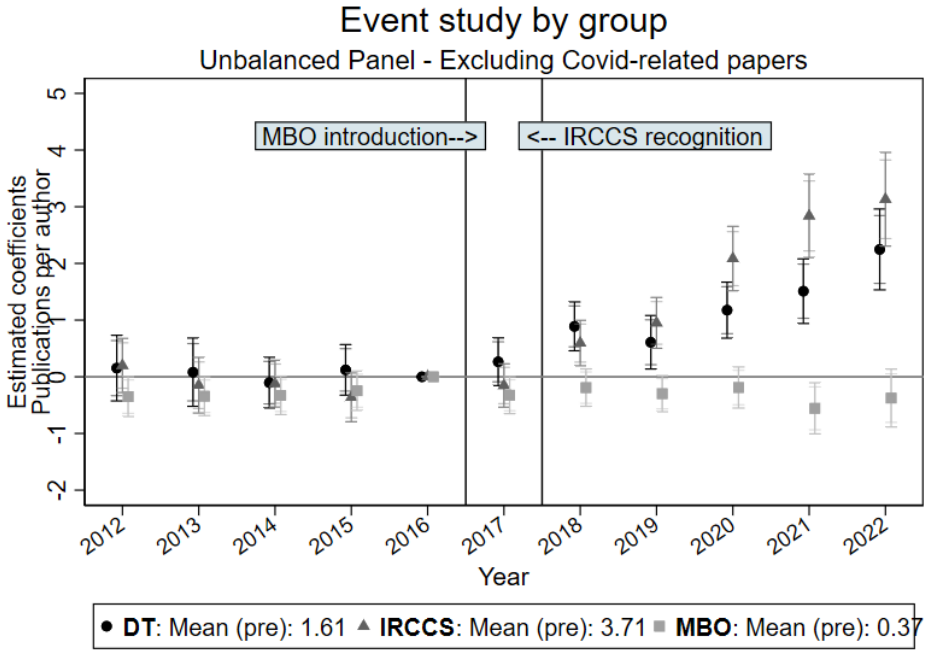


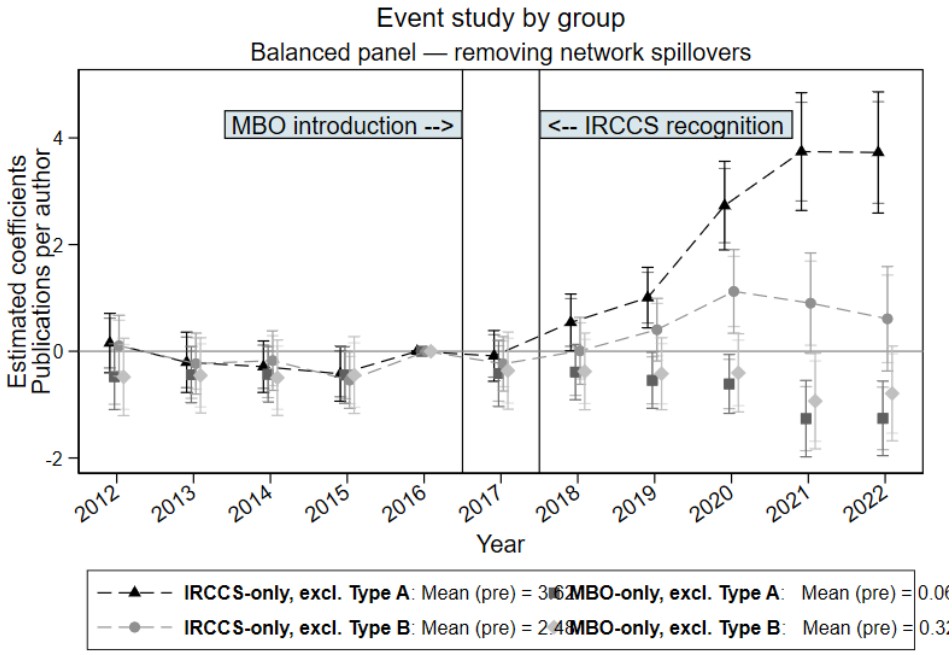
Figure A34: Event-study of the IRCCS after dropping COVID-related publications (unbalanced panel).

2012-2022; SEs clustered at individual level

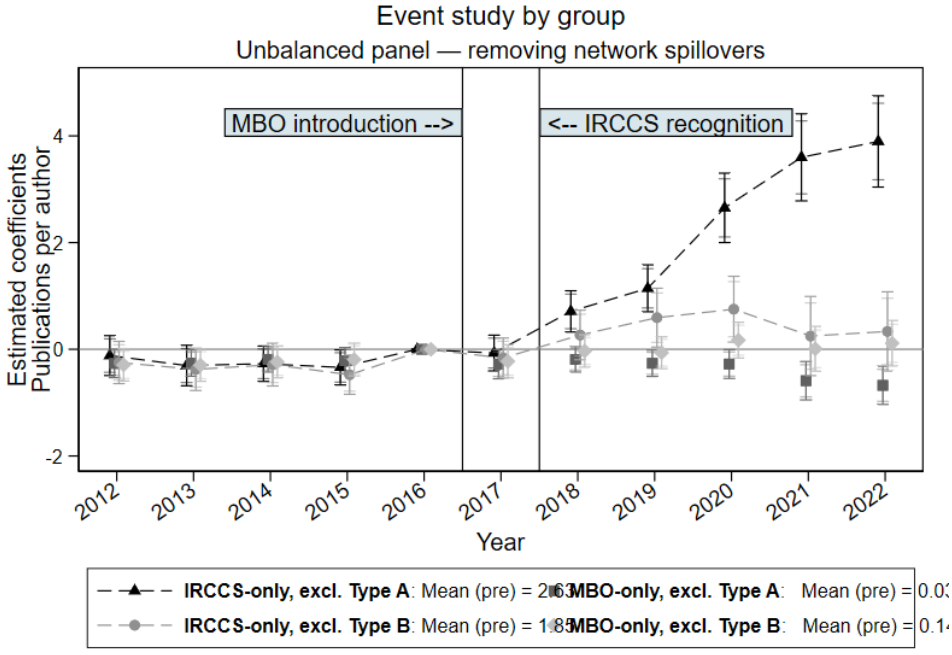
	(1) Pubs	(2) Pubs	(3) Pubs	(4) Pubs	(5) Pubs	(6) Pubs	(7) Pubs	(8) Pubs
Post 2017*MBO-only	-0.30943 ***	-0.18027			-0.13506 **	0.19353 ***		
	[0.10852]	[0.13296]			[0.06280]	[0.07404]		
Post 2017*MBO-only	2.11063 ***	0.49499 ***			2.19877 ***	0.59396 ***		
	[0.29400]	[0.17523]			[0.23569]	[0.15897]		
Post 2018*IRCCS-only			-0.38701 ***	-0.21920 *			-0.15469 **	0.24161 ***
			[0.10317]	[0.13252]			[0.06557]	[0.07850]
Post 2018*IRCCS-only			2.50004 ***	0.69136 ***			2.58720 ***	0.70018 ***
			[0.32107]	[0.19850]			[0.25718]	[0.17894]
Observations	5,311	4,871	5,311	4,871	12,202	12,803	12,202	12,803
R-squared	0.79049	0.66545	0.79539	0.66791	0.71808	0.52814	0.72473	0.52901
Individual FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	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 I	No Type A	No Type B	No Type A	No Type B	No Type A	No Type B	No Type A	No Type B
Panel II	Balanced	Balanced	Balanced	Balanced	Unbalanced	Unbalanced	Unbalanced	Unbalanced
Mean across treated	2.030	1.250	2.045	1.242	1.021	0.581	1.044	0.596

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

Table A9: Impact of IRCCS recognition on Annual Publications, by changing the outcome to log-transformation and inverse hyperbolic sine (Difference-in-Differences).

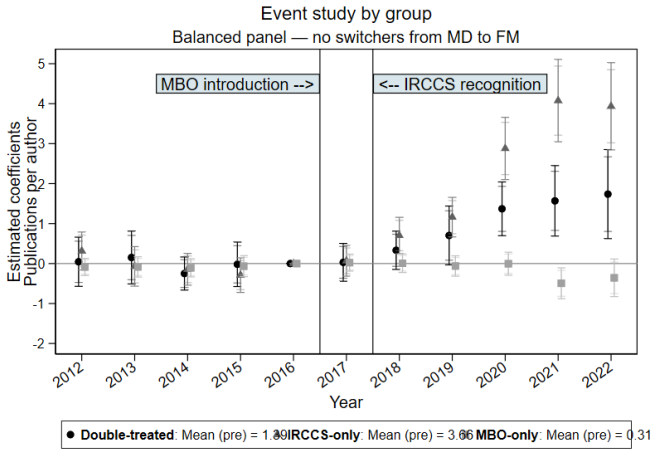


a) Balanced panel - Event-study of the two policies' effect on publications of the IRCCS-treated and MBO-treated only groups once excluded double-treated units and *Type A* researchers (MBO-treated units who *ever* co-authored with IRCCS researchers) in the first case, and double-treated units and *Type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

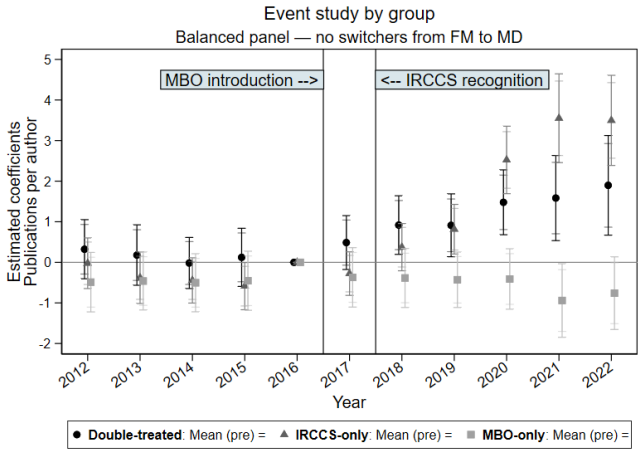


a) Unbalanced panel - Event-study of the two policies' effect on publications of the IRCCS-treated and MBO-treated only groups once excluded double-treated units and *Type A* researchers (MBO-treated units who *ever* co-authored with IRCCS researchers) in the first case, and double-treated units and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

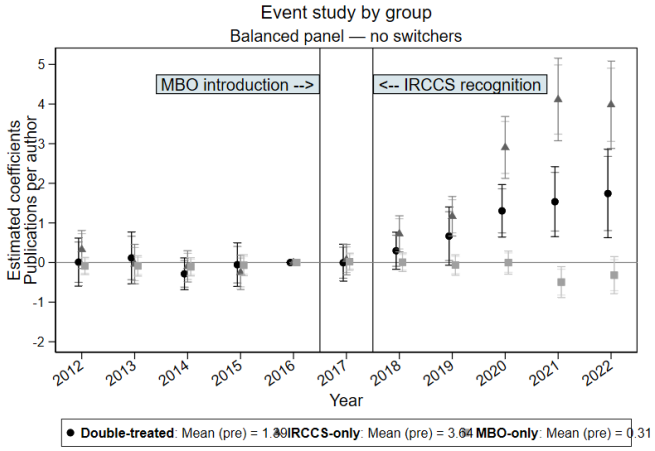
Figure A35: Event-study of for the dynamic effects of the MBO policy and the IRCCS recognition effect on publications in different samples subset based on collaboration dynamics. Double-treated units are always excluded. a) is the Balanced panel, b) the unbalanced one.



a) Removing switchers from Medical Direction to Faculty Membership.

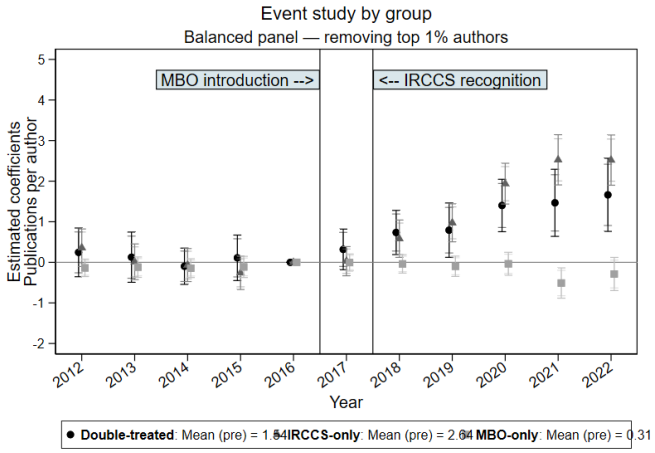


b) Removing switchers from Faculty Membership to Medical Direction.

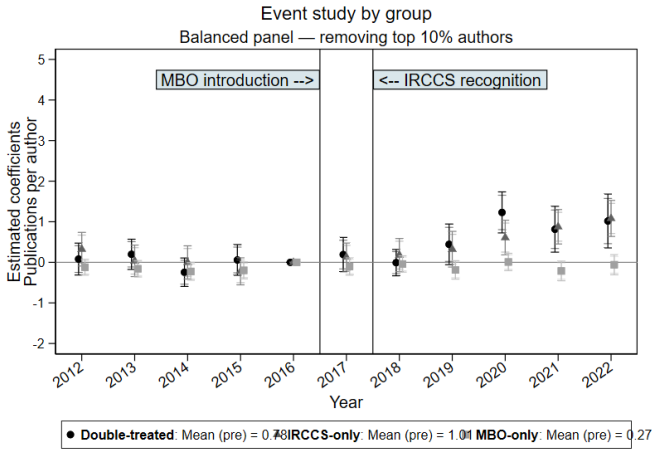


c) Removing all switcher.

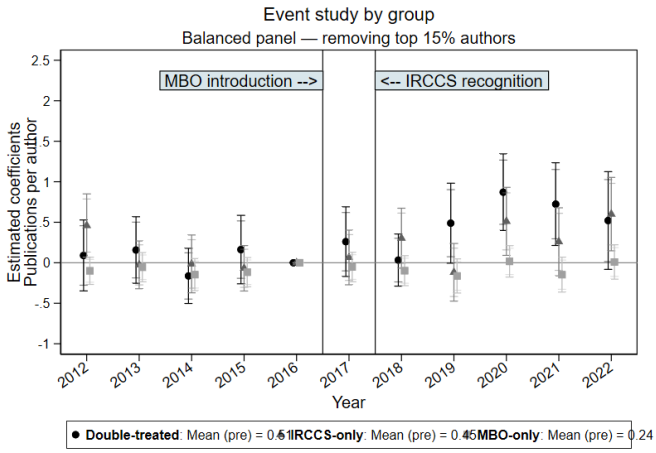
Figure A36: Dynamic mpact of MBO-implementation and IRCCS recognition on Annual Publications in the sub-samples obtained by excluding those who switch across groups.



a) Removing top 1% researchers.

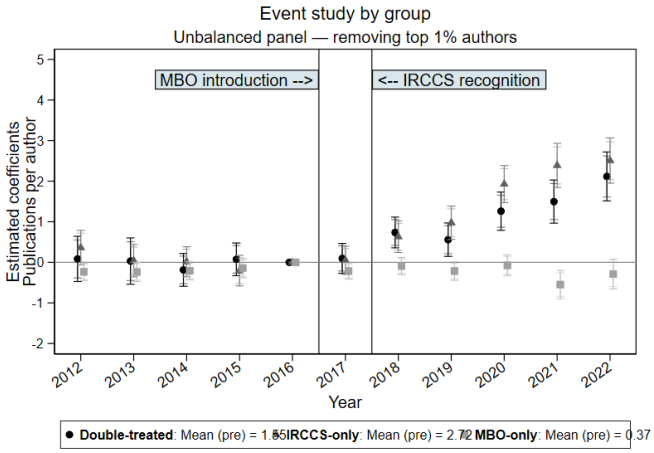


b) Removing top 10% researchers.

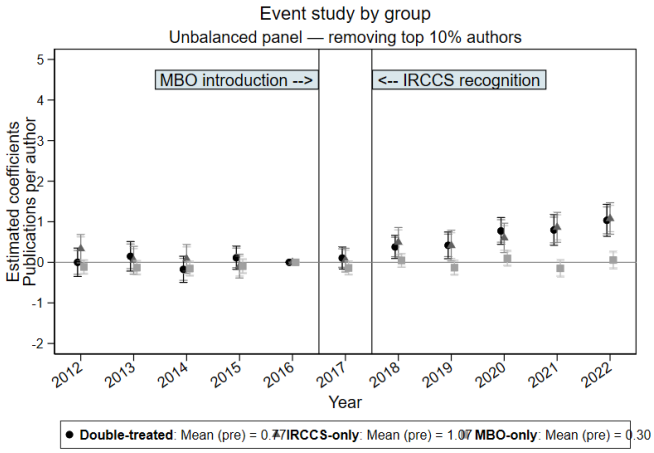


c) Removing top 15% researchers.

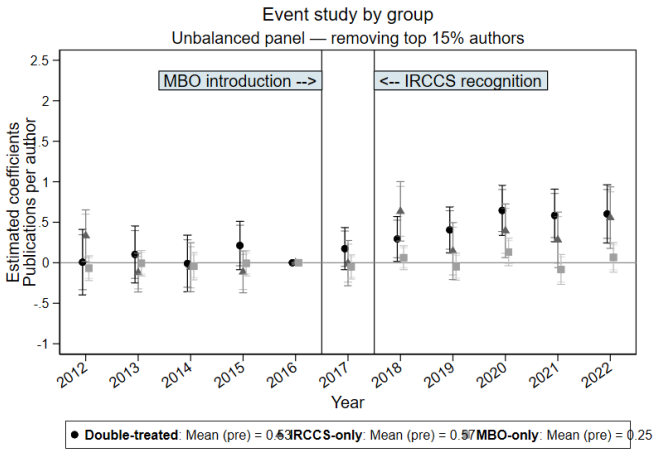
Figure A37: Balanced panel - Dynamic impact of MBO-implementation and IRCCS recognition on annual publications in the sub-samples obtained by excluding most prolific authors.



a) Removing top 1% researchers.

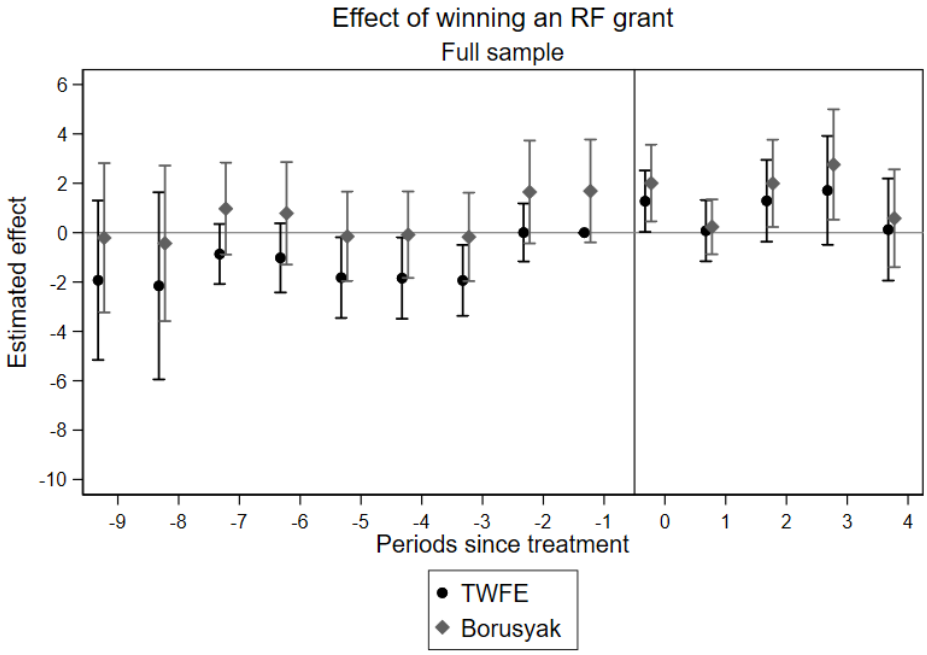


b) Removing top 10% researchers.

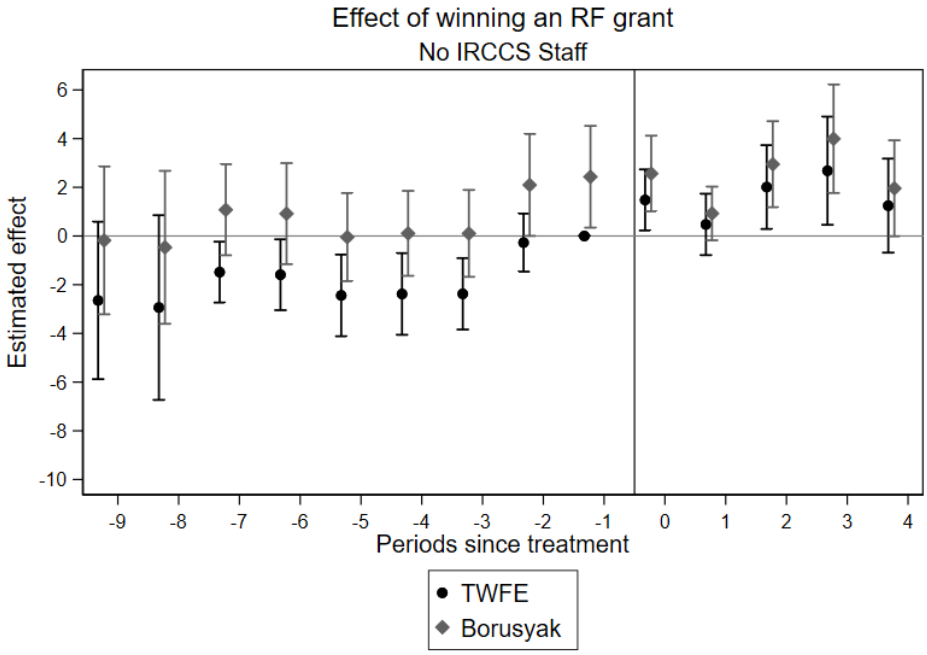


c) Removing top 15% researchers.

Figure A38: Unbalanced panel - Dynamic impact of MBO-implementation and IRCCS recognition on annual publications in the sub-samples obtained by excluding most prolific authors.



a) Event-study of the staggered awarding of Targeted Research funds to P.I.s of the IRCCS perimeters (full unbalanced sample).



b) Event-study of the staggered awarding of Targeted Research funds to P.I.s of the IRCCS perimeters (excluding IRCCS researchers without access to TR funds).

Figure A39: Event-studies of the staggered awarding of Targeted Research funds to P.I.s of the IRCCS perimeters.

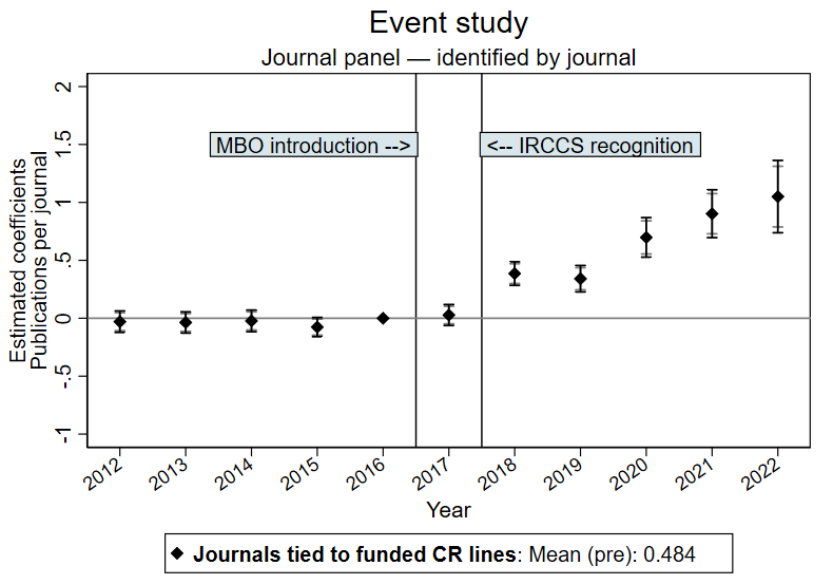


Figure A40: Event Study of the impact of IRCCS recognition + inclusion in Current Research lines on publications - by journals.

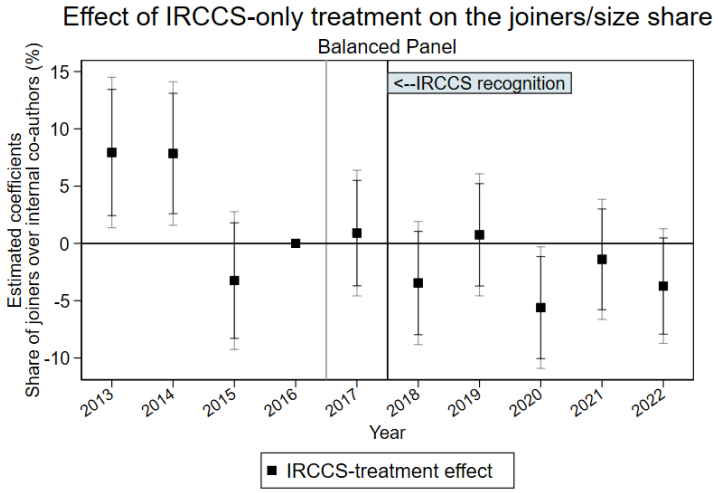


Figure A41: Event-study estimates for the share of joiners over total internal co-authors.

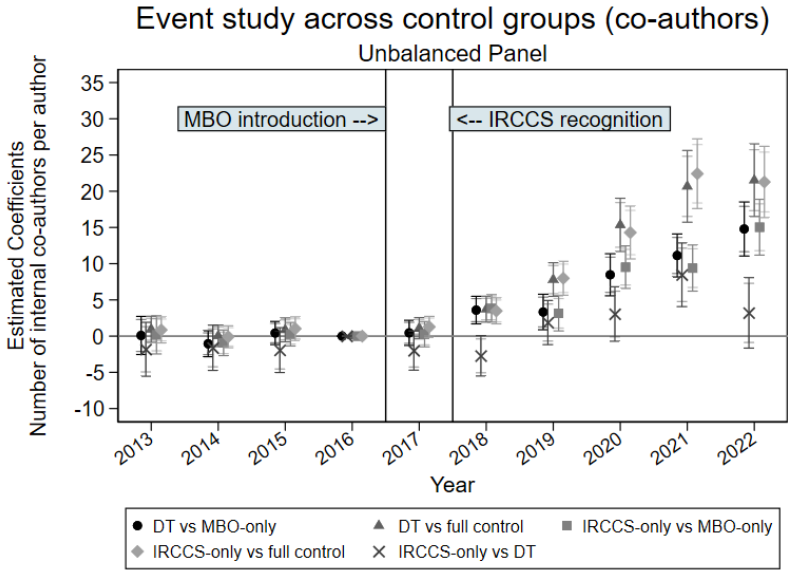


Figure A42: Event-study estimates for the total internal co-authors across alternative treatment group comparisons.

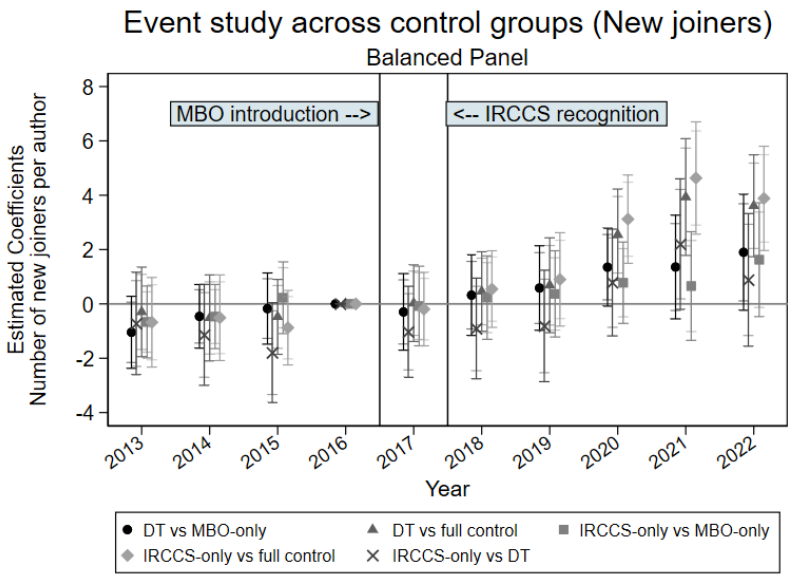


Figure A43: Event-study estimates for the number of joiners across alternative treatment group comparisons.

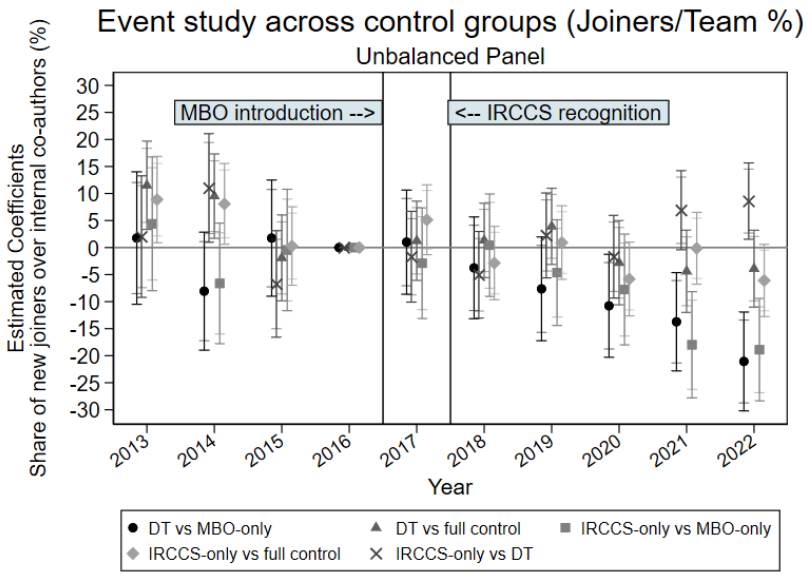
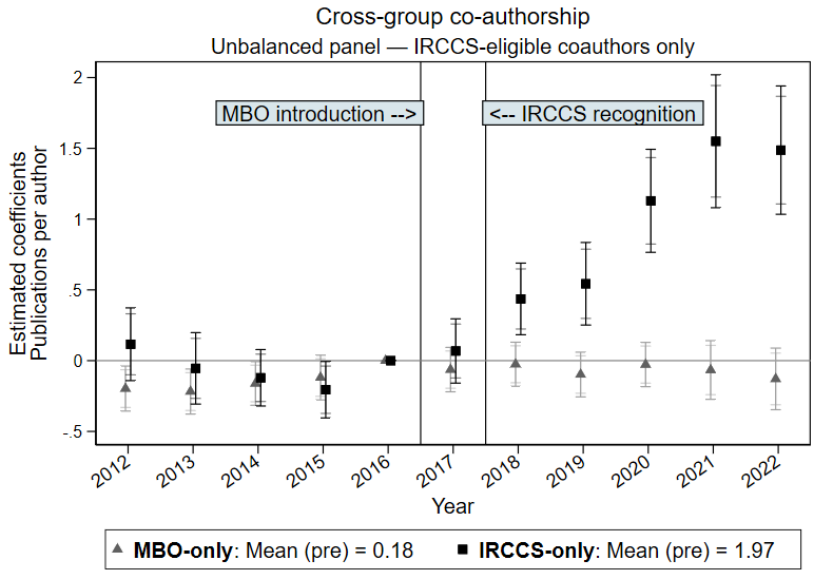
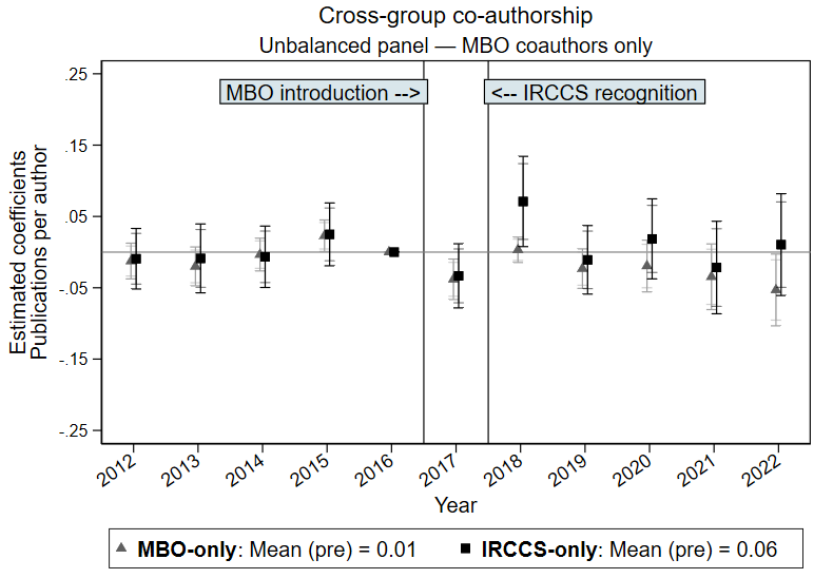


Figure A44: Event-study estimates for the nshare of joiners over total internal co-authors across alternative treatment group comparisons.

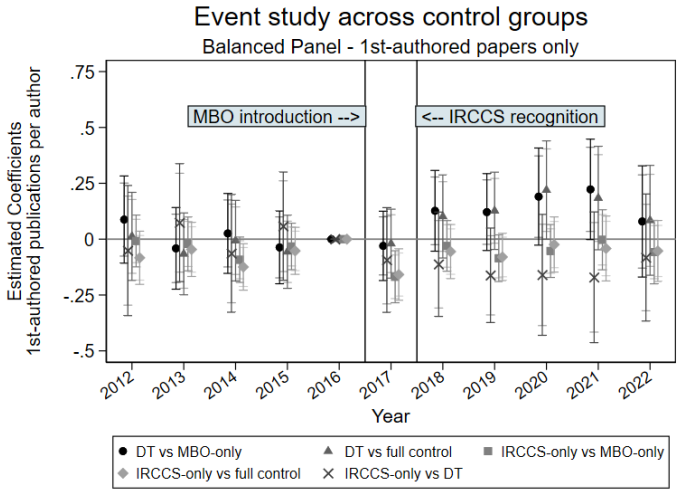


a) Outcome: publications co-authored with IRCCS units only.

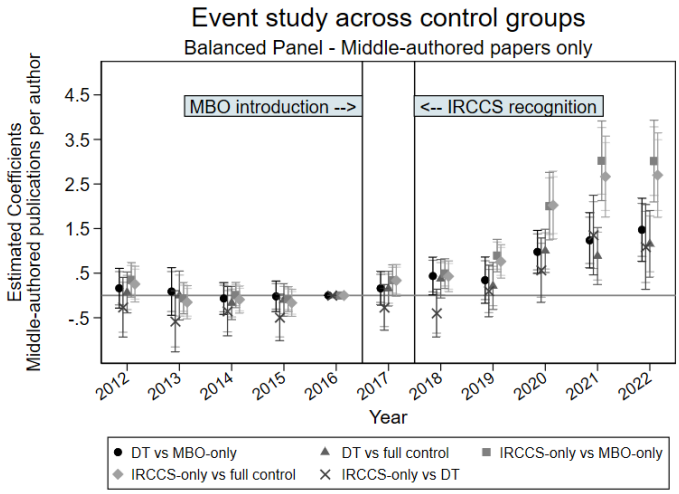


b) Outcome: publications co-authored with MBO-eligible units only.

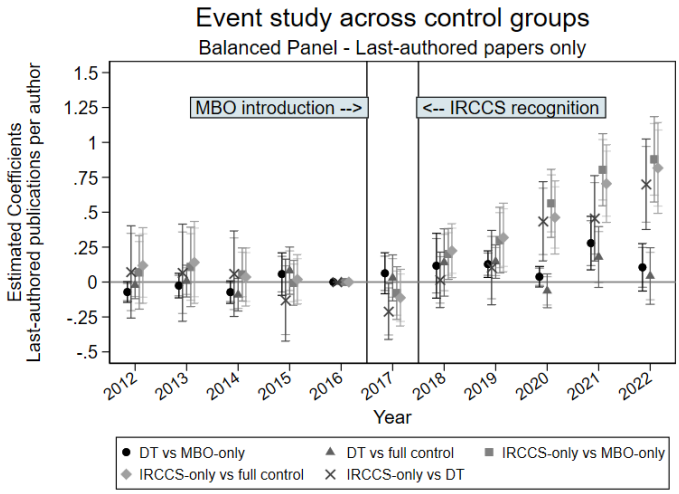
Figure A45: Dynamic impact of the policy implementation on annual cross-group publications in the balanced panel.



a) First-authored publications.

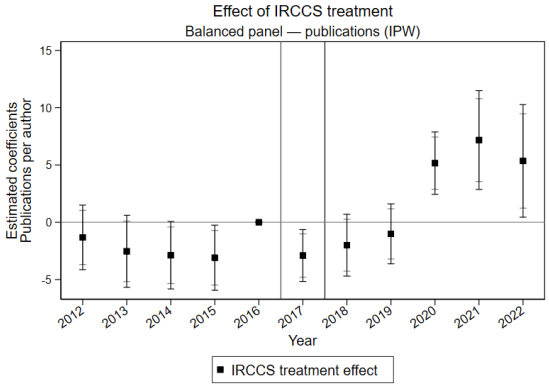


b) Middle-authored publications.

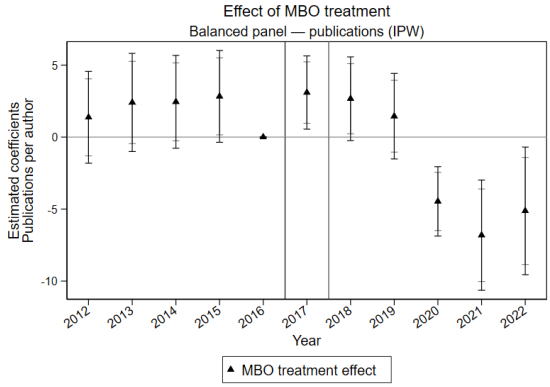


c) Last-authored publications.

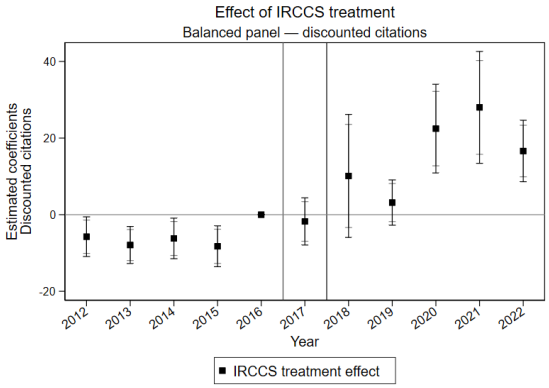
Figure A46: Event-study estimates across alternative treatment group comparisons for authorship positions



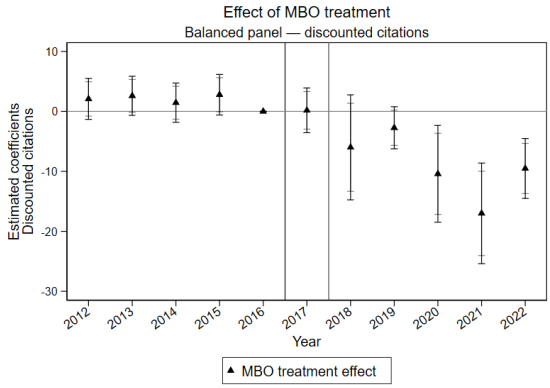
a) IRCCS effect on citation-weighted publications.



b) MBO effect on citation-weighted publications.



c) IRCCS effect on total discounted citations.



d) MBO effect on total discounted citations.

Figure A47: Additional event-study estimates on citation-based outcomes.

	(1) MBO vs. non-MBO	(2) DT vs IRCCS	(3) DT vs MBO	(4) DT vs control	(5) MBO vs IRCCS	(6) MBO vs control
Pay Out (€)	-1037.37939***	-1466.69556***	1780.15076***	1628.63623***	-3556.25537***	-269.15292**
(SE)	(389.74774)	(223.09830)	(188.85347)	(200.98660)	(200.08897)	(133.17281)
R^2	0.610	0.668	0.639	0.635	0.702	0.483
Mean	1196.52	2216.09	2216.09	2216.09	520.66	520.66
Sum of weighted IF	-0.05247	-11.46735***	3.66103***	3.92988***	-17.96440***	0.14813
(SE)	(2.09297)	(3.33887)	(0.93505)	(0.66406)	(4.41293)	(0.76541)
R^2	0.336	0.348	0.372	0.613	0.339	0.191
Mean	3.16	5.68	5.68	5.68	1.48	1.48
Eligible Papers	-1.25746	-4.68687***	4.03192***	3.58589***	-9.94389***	-0.69990***
(SE)	(0.98802)	(0.83610)	(0.34259)	(0.38195)	(0.94491)	(0.25696)
R^2	0.581	0.615	0.687	0.669	0.637	0.571
Mean	2.82	4.88	4.88	4.88	1.46	1.46
Delta (Pubs-Eligible)	0.17500	3.65820***	-2.78376***	-2.37526***	7.48813***	0.54775***
(SE)	(0.69514)	(0.66451)	(0.27954)	(0.30812)	(0.73891)	(0.18284)
R^2	0.406	0.477	0.586	0.553	0.471	0.520
Mean	-1.03	-1.88	-1.88	-1.88	-0.47	-0.47
N	8651	4186	4224	3612	5017	4465
Time FE	YES	YES	YES	YES	YES	YES
Individual FE	YES	YES	YES	YES	YES	YES
Panel	Full (no 0s)	DT and IRCCS (no 0s)	DT and MBO (no 0s)	DT and control (no 0s)	MBO and IRCCS (no 0s)	MBO and control (no 0s)
Time Range	2017–2022	2017–2022	2017–2022	2017–2022	2017–2022	2017–2022

Table A10: Regressions between enrollment into IRCCS perimeter and several factors.

PNE	ATT			Pre (joint)		Mean Pre	Note
	Coef.	p-value	BH q-val.	F-stat	p-value		
A.M.I.: 1y mort.	-2.370	0.270	0.332	17.840	0.000***	9.599	?
	(2.106)						
C.H.F.: 30d mort.	-1.491*	0.061	0.099*	9.523	0.000***	12.431	?
	(0.780)						
C.H.F.: 30d readm.	2.483***	0.000	0.000***	1.065	0.353	13.886	++
	(0.461)						
Colon CA surg.: 30d mort.	-1.380*	0.062	0.099*	2.160	0.136	3.896	0
	(0.705)						
Gastric CA surg.: 30d mort.	2.665	0.395	0.421	0.073	0.930	5.942	0
	(2.910)						
Hip Fracture: 30d mort.	1.778***	0.001	0.004***	0.932	0.403	6.459	++
	(0.508)						
Hip repl'ment: 30d readm.	-1.882***	0.000	0.001***	18.923	0.000***	4.323	?
	(0.477)						
Knee repl'ment: 30d readm.	-1.185**	0.013	0.029**	8.108	0.001***	1.315	?
	(0.450)						
Lung CA surg.: 30d mort.	-1.006	0.160	0.213	1.784	0.260	1.454	0
	(0.609)						
Nstemi: 30d mort.	0.447	0.816	0.816	0.270	0.765	9.383	0
	(1.910)						
P.E.: 30d readm. From adm.	-5.944***	0.000	0.000***	29.255	0.000***	10.717	?
	(0.872)						
Prostate CA surg.: 30d readm.	-4.491**	0.027	0.053*	16.070	0.002***	2.570	?
	(1.656)						
Stemi: 30d mort.	0.849	0.342	0.391	9.498	0.002***	9.654	?
	(0.868)						
isch. Stroke: 30d mort.	6.311***	0.000	0.000***	17.773	0.000***	11.458	?
	(0.992)						
isch. Stroke: 30d readm.	2.673***	0.000	0.001***	10.557	0.001***	7.823	?
	(0.615)						
valvulopl. Or Heart Valve repl'ment: 30d mort.	-1.703	0.122	0.177	0.817	0.485	3.243	0
	(0.945)						

Table A11: Each row reports results from a separate regression of yearly PNE outcomes on a post-treatment indicator, for the studied hospital controlling for unit and year fixed effects, compared to a set of control hospitals within the same region. Standard errors, reported in parentheses beneath the ATT coefficient, are clustered at the hospital level. The block Pre (joint) reports the F-test statistic and p-value from a joint test of all pre-treatment leads (2015–2017). Benjamini–Hochberg adjusted q-values are reported to account for multiple hypothesis testing within the reported family of PNE outcomes. Depending on the variance estimator and finite-sample correction, the reported test statistic may be an F statistic or a chi-squared statistic. Mean pre-treatment level is computed over the pre-2018 period. Significance: * p<0.10, ** p<0.05, *** p<0.01.

PNE	ATT			Pre (joint)		Mean Pre	Note
	Coef.	p-value	BH q-val.	F-stat	p-value		
A.M.I.: Vol. With thrombol. per 100 beds	0.359*	0.089	0.148	0.000	0.983	1.581	+
	(0.197)						
Carotid revasc'tion: Endarterectomy per 100 beds	5.482**	0.031	0.061*	2.538	0.100	19.638	++
	(2.387)						
Carotid revasc'tion: Vol. per 100 beds	2.759	0.361	0.452	5.375	0.010***	22.570	?
	(2.978)						
Isol. Surg. (colon CA) min-inv.: Vol. per 100 beds	4.130***	0.001	0.004***	3.835	0.030**	7.559	?
	(1.174)						
Isol. coron. Artery Bypass Graft: Vol. per 100 beds	0.645	0.928	0.928	6.165	0.024**	51.676	?
	(6.930)						
Lung CA surg. min-inv.: Vol. per 100 beds	6.527***	0.008	0.021**	0.385	0.694	7.472	++
	(1.795)						
Lung CA surg.: Vol. per 100 beds	0.322	0.925	0.928	1.427	0.275	18.989	0
	(3.340)						
Pancr. Ca Resection surg.: Vol. per 100 beds	2.072***	0.000	0.000***	23.879	0.000***	2.397	?
	(0.308)						
Pancreatic CA surg.: Vol. per 100 beds	1.852***	0.000	0.000***	7.985	0.002***	1.922	?
	(0.306)						
tot. Bypass Proc.s: Vol. per 100 beds	7.075	0.274	0.392	0.440	0.659	93.283	0
	(6.030)						

Table A12: Each row reports results from a separate regression of yearly PNE outcomes on a post-treatment indicator, normalized for 1000 beds, for the studied hospital controlling for unit and year fixed effects, compared to a set of control hospitals within the same region. Standard errors, reported in parentheses beneath the ATT coefficient, are clustered at the hospital level. The block Pre (joint) reports the F-test statistic and p-value from a joint test of all pre-treatment leads (2015–2017). Benjamini–Hochberg adjusted q-values are reported to account for multiple hypothesis testing within the reported family of PNE outcomes. Depending on the variance estimator and finite-sample correction, the reported test statistic may be an F statistic or a chi-squared statistic. Mean pre-treatment level is computed over the pre-2018 period. Significance: * p<0.10, ** p<0.05, *** p<0.01.