Research Scaling or Transformation? Incentive Structures and Research Capacity within a Healthcare Organization*

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

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

Abstract

This paper studies how different organizational approaches affect scientific productivity among high-skilled workers. We compare a managerial intervention that increases the returns to research through performance-based incentives with an institutional change that expands research capacity through public funding and organizational support. Exploiting two policies within a large research-intensive Italian university hospital and detailed physician-level panel data from 2012–2022, we estimate the effects of both interventions using a set of difference-in-differences strategies. We find that performance incentives alone do not significantly increase scientific output. By contrast, expanded research capacity generates a large and persistent increase in productivity among researchers exposed to the new research environment. The results reveal substantial heterogeneity across workers: individuals already engaged in research-intensive activities respond strongly to expanded research capacity, whereas physicians primarily devoted to clinical tasks show limited responses to publication-based incentives. The increase operates primarily through the intensive margin, reflecting the expansion of existing research teams and activities rather than the emergence of new collaboration patterns, scientific leadership or research specialization. While total citations increase, citations per publication decline, suggesting a quantity–quality trade-off. Consistent with a broader expansion of research-related activity, we also document increases in high-complexity clinical procedures, with no systematic changes in healthcare quality indicators. The findings suggest that policies designed to increase scientific output are not necessarily the same as policies capable of transforming the research process itself.

Keywords: High-skilled Workers, Knowledge Production, Management-By-Objective, Public Funding, Healthcare Research.

JEL Classification: I10, I23, J24, O31.

*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 Manfredi Aliberti, 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, the 4^{th} APHEC Conference, the Cattolica Workshop on IPR, Science and Innovation, the XXXVII SIEP Conference, the 5th Bank of Italy Human Capital Workshop, the XVI NERI Workshop, and the ESI Workshop at the BSE Summer Forum for their helpful comments and suggestions.

1 Introduction

Medical research plays a central role in modern healthcare systems. Although scientific activity is costly and its benefits are often uncertain in the short run (Grant & Buxton, 2018), advances in medical knowledge may generate substantial long-run improvements in treatments, population health, and healthcare efficiency (Chandra & Skinner, 2012; Murphy & Topel, 2006). The social returns to biomedical innovation can be considerable, provided that new knowledge is successfully translated into clinical practice and diffused across healthcare organizations (Barrenho et al., 2021, 2025; Chandra & Staiger, 2007). For this reason, hospitals are increasingly expected to deliver healthcare services joint with substantial contributions to the production of scientific knowledge. Yet scientific production within healthcare organizations requires substantial investments and organizational support, making it dependent not only on individual effort, but also on the institutional environment in which physicians operate. To this extent, healthcare organizations seeking to strengthen their research activity can broadly intervene along two margins; they may increase the returns to research by rewarding observable scientific output through performance-based incentives, extending schemes that are already common in healthcare settings, albeit typically targeted at clinical rather than research outcomes (Eijkenaar et al., 2013). Alternatively, they may relax the constraints that limit scientific production by facilitating access to research funding, infrastructure, and other organizational resources. While both approaches are commonly observed in practice, relatively little is known about their relative effectiveness and outcomes within healthcare organizations, where research activities compete with clinical responsibilities and where workers may differ substantially in their underlying motivations and incentive structures.

Physicians operate indeed under heterogeneous incentive structures. Academic physicians often face strong pre-existing incentives to conduct research, through the weight that the most ambitious ones might place to career advancement and reputational rewards (Checchi et al., 2021), or even due to altruistic motifs and scientific curiosity (Rousseau et al., 2021). At the same time, non-academic physicians must typically balance research activities with demanding clinical responsibilities, facing higher constraints in a multitasking setting à la Holmstrom and Milgrom, 1991. In such environments, incentives targeted at a single dimension of performance may induce effort reallocation across tasks, as documented in academic settings where stronger research incentives increase research activity at the expense of other duties (De Philippis, 2021). As a result, different forms of research support may generate heterogeneous responses across workers.

In addition to that, increases in scientific output do not necessarily imply a transformation of an organization's research activity. Research growth may simply reflect the expansion of existing teams and infrastructures, or instead emerge through broader changes in collabora-

tion patterns, field composition, leadership and authorship structures, and scientific impact. Understanding whether research policies generate research scaling (*more research*) or research transformation (*different research*), is therefore crucial for evaluating their effectiveness. In fact, policies that increase the quantity of research are not necessarily the same as policies that transform how research is conducted. Increases in output due to the expansion of the activity of existing teams, may have indeed very different implications from dynamics involving phenomena like collaboration structures and knowledge diffusion (Barrenho et al., 2025), the attraction of new researchers (Maynou et al., 2024), or the shift of the whole organizational profiles of healthcare institutions (Ghandour et al., 2022). Understanding which of these margins drives observed productivity gains is therefore important for evaluating the broader organizational consequences of research-support policies.

Accordingly, the research question of this paper is twofold. First, do different forms of research support generate heterogeneous responses across physicians with different underlying incentive structures? Second, when research output increases, does scientific activity merely scale up, or does it fundamentally change? We address these questions by exploiting two institutional reforms occurring within a major Italian private university hospital, which offers both high-quality care and medical education while undertaking impactful research. In 2017, the hospital introduced a performance-based bonus scheme (*Management-by-Objective*, or MBO), rewarding publications among non-academic physicians. One year later, in 2018, the institution obtained IRCCS status (*Scientific Institute for Research, Hospitalization and Healthcare*), gaining access to dedicated public research funding which substantially expanded its research capacity to be directly employed by a subset of its employees. The combination of these interventions allows us to compare research policies operating through individual incentives with policies operating through resources and capacity within the same organizational environment, and to assess their organizational effects. First, we contribute to the literature on incentives, multitasking, and management in public and imperfect markets, including education (Bloom, Lemos, et al., 2015; Muralidharan & Sundararaman, 2011), healthcare (Bloom, Propper, et al., 2015; Bloom et al., 2020; Goodall, 2011), and public agencies (Burgess et al., 2017; Dal Bó et al., 2013). This literature has shown that organizational incentives can substantially affect performance, and may even induce workers to reallocate effort across competing tasks (De Philippis, 2021). Recent evidence also suggests that access to research opportunities and infrastructure may generate heterogeneous responses in the long-term, depending on workers' pre-existing characteristics and career trajectory (Lepinteur & Nisticò, 2025). Despite this, much less is known about how different forms of research support influence scientific production within organizations where researchers may differ with respects to their tasks. We provide additional evidence on this from a healthcare setting. Second, we contribute to the literature on physicians' responses to financial incen-

tives (Gruber and Owings, 1996; Shurtz, 2013, 2014; Molitor, 2018; Bertoli and Grembi, 2019; Brosig-Koch et al., 2024). Existing studies focus almost exclusively on clinical decisions and healthcare provision. By contrast, we examine whether publication-based incentives affect scientific production and whether responses differ according to physicians' underlying group-level characteristics. Third, we contribute to the literature on research funding and scientific production. Most existing studies highlight the positive effects of grants, funding programs, or performance-based funding schemes on research output and innovation (Azoulay et al., 2011; Benavente et al., 2012; Ghirelli et al., 2023; Babina et al., 2023), while other works display mixed evidence (Banal-Estañol et al., 2023; Jacob & Lefgren, 2011). A related work also shows that funding can affect the mobility and allocation of researchers (Baruffaldi et al., 2020). However, our setting allows us to compare, within the same institution, the organizational effects of different research-support mechanisms. Beyond estimating effects on publications, we distinguish between research scaling and research transformation, examining whether the findings extend to how scientific activity is organized through collaboration networks, authorship structures, field composition, and scientific impact.

Previewing our results, we find that the performance-based MBO scheme does not significantly increase research productivity among non-academic physicians, whereas IRCCS recognition and the associated public funding increase publication output by roughly 72% among physicians operating in settings that are directly affected by the new funding stream. The two policies appear partly complementary for a subset of physicians: monetary incentives manage to generate positive responses only among physicians who also benefit from expanded research capacity. We further show that the increase in scientific production reflects research scaling rather than research transformation. Research activity expands across existing fields and teams, with little evidence of substantial changes in collaboration structures, scientific leadership, or field composition. Consistent with this interpretation, IRCCS funding increases total citations through higher publication volumes, but citations per publication decline. A major implication is that policies fostering research output shall not be deemed as equivalent to those aiming at transforming the organization and diffusion of scientific activity.

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 last 5 years, and more than 75 of its affiliated researchers were included in the top 2% ranking of scientists globally in 2024 (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 Management-by-Objective policy to encourage publications 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 based on output: 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, ranging to little below or little above according to the year. 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 authors who co-author with only 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 with 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 the Regions being awarded the funds themselves in public tenders): the research status allow IRCCS hospitals to compete directly with the Regions and the National Institute of Health for the targeted funds. The IRCCS label also carries obligations though: 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]. The bulk of the IRCCS staff list was made by academically affiliated physicians (university-employed doctors working at the hospital), but it also involved non-academic hospital physicians active in research. Importantly, all academic physicians, whether included in the IRCCS list or not, could not be part of the MBO bonus program as well, and were appointed by the hospital management; henceforth, assignment was determined by the institutional process underlying IRCCS accreditation, rather than by individual application or opt-in decisions. Meanwhile, non-academic physicians on the IRCCS list were still eligible for the MBO bonus, due to being hospital-employed clinicians and not structured academics. 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 directly benefitting from the IRCCS public funding) while also being eligible for the MBO

[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.

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.

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).

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 profes-

sional capacity (physicians, researchers, and other healthcare staff management) from 2017 onward. From that, we identify our population of interest: physicians and medical researchers continuously employed at the hospital between 2017 and 2022. For each person, the HR data provide demographic information (age, gender), employment details (job title or role, 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 available employer records begin in 2017, we supplemented them with hire dates and career information obtained from institutional records, CVs, professional registries, and other publicly available sources. This information allows us to retrospectively reconstruct physicians' affiliation histories and extend the panel back to 2012. Since workers who left the hospital before 2017 cannot be identified, the baseline analysis relies on a balanced panel of individuals continuously observed throughout the study period, avoiding attrition concentrated in the post-treatment years[2]. The resulting balanced panel covers 584 physicians and researchers observed annually from 2012 to 2022, yielding 6,424 person-year observations[3]. Treatment status is defined according to the categories reported in Table 1. Because some physicians entered the IRCCS perimeter after 2018, the baseline analysis adopts an ever-treated definition; alternative specifications accounting for staggered entry are discussed in Section 5.2.2.

	MBO		IRCCS-only		Double-treated		Control	
Sample composition								
Units	178		198		69		162	
Research output								
Publications	0.46	(1.05)	5.01	(6.93)	2.26	(2.77)	0.85	(1.68)
Citations	8.54	(33.92)	151.15	(341.13)	66.69	(144.07)	15.54	(37.68)
Worker characteristics (all)								
Age	51.13	(7.86)						
Female	0.42	(0.49)						
Tenure	16.49	(10.02)						
Co-authors per paper	9.98	(6.93)						
Pages per paper	8.29	(3.19)						
Observations	6424							

Table 2: Descriptives

[2]As the hire dates are missing for individuals who were not employed in 2020 and/or 2023, additional details for such missing units were collected from institutional websites, CVs, professional registries (FNOMCeO), LinkedIn profiles, and other publicly available sources. Details on the reconstruction procedure are available upon request.

[3]As a robustness check, we also construct an unbalanced panel that incorporates additional physicians identified through external sources. The unbalanced panel consists of 1,540 units.

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 both the balanced and unbalanced panels[4]. Each publication record provides the publication year, journal, number and name of co-authors, citations[5], and other bibliometric information. 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.

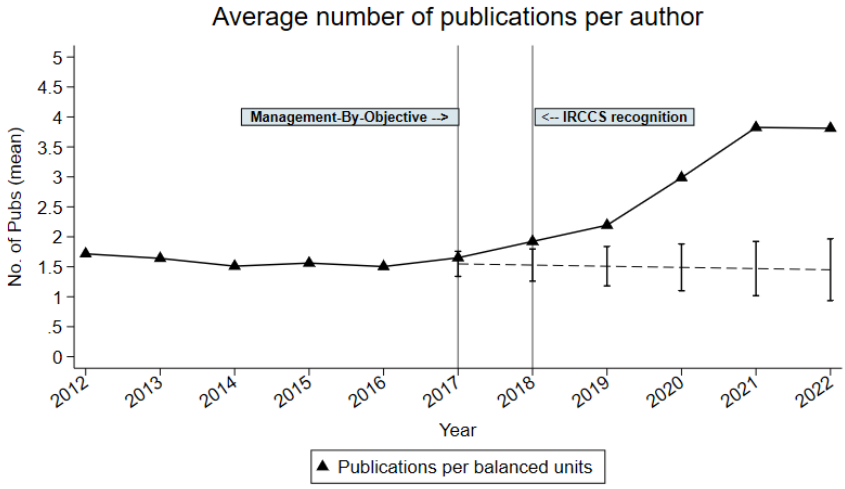


Figure 1: Average publications per researcher per year (2012–2022), balanced panel. Vertical lines indicate the introduction of the MBO policy in 2017 and the IRCCS recognition in 2018.

Table 2 reports summary statistics for the balanced sample[6]. Research productivity differs markedly across groups. Physicians ever included in the IRCCS perimeter exhibit substantially higher publication and citation levels than all other groups, while MBO-only physicians display comparatively limited research activity. These differences point to substantial heterogeneity in research orientation even before the introduction of the two institutional reforms.

[4]We ensured that the recovered authors were the ones affiliated to the hospital, by matching names and surnames, fields of research, and by double checking each of them, in order not to deal 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.

[5]As of the time of data collection, which is October 2024.

[6]Treatment groups are here defined on an ever-exposure basis; consequently, 23 physicians who change status during the observation period contribute to more than one category and are addressed separately in the robustness analysis

Figure 1 plots instead the average number of publications per physician per year between 2012 and 2022 for the balanced panel. Publication activity remains broadly stable until 2017 and increases markedly after 2018. The same pattern emerges across alternative samples with different degrees of attrition (Appendix Figures A1 and A2), suggesting a potential association between IRCCS recognition and the subsequent expansion of scientific output.

Figure 2 reports instead publication trends by treatment status and suggests that the post-2018 increase in scientific output is concentrated among physicians exposed to the IRCCS research environment.reated physicians display a marked increase in output.

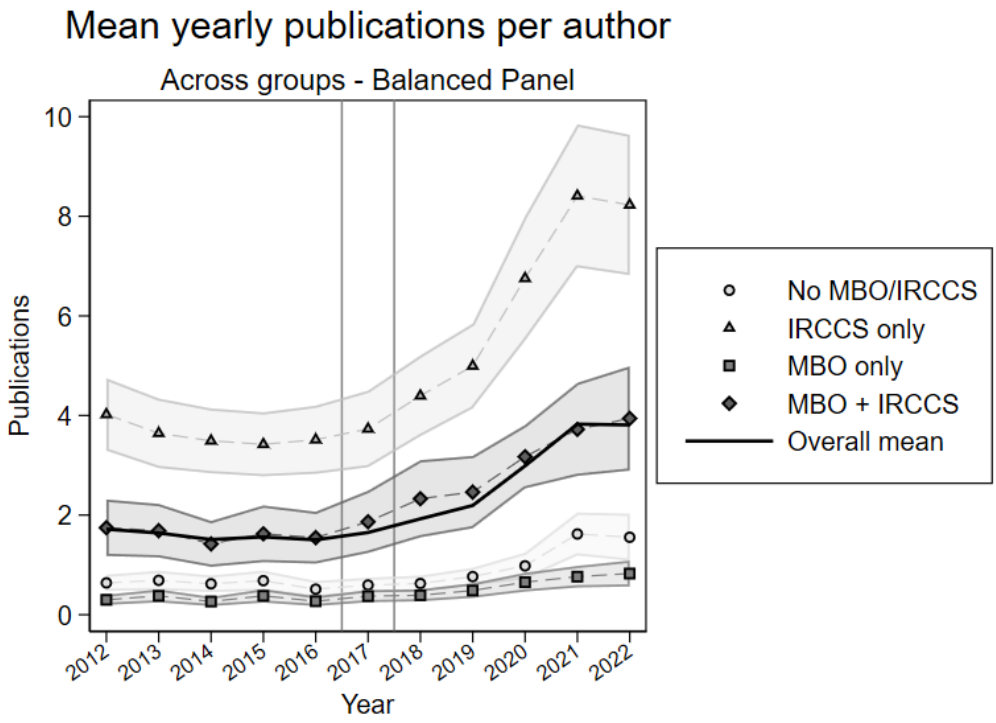


Figure 2: Average yearly publications across different groups of healthcare professionals, defined according their treatment status after 2017 and 2018.

4 Dynamic Specification

4.1 Empirical Strategy

The introduction of the MBO scheme in 2017 and the hospital's recognition as an IRCCS institute in 2018 represent two closely related institutional reforms aimed at strengthening research activity within the organization. While the MBO policy increased the monetary returns to scientific production for non-academic physicians, IRCCS recognition expanded

access to research funding, infrastructure, and organizational support. Because the two interventions occurred in close temporal sequence and affected partially overlapping groups of physicians, we begin by studying their joint impact through a unified dynamic event-study framework. We exploit the classification introduced in Table 1 to define four groups according to their exposure to the two policies. Group 0 serves as the control group and includes individuals exposed to neither intervention. Group 1 ('*MBO-only*") consists of physicians eligible for the monetary incentive but not included in the IRCCS research perimeter. Group 2 ('*IRCCS-only*") includes physicians exposed to IRCCS-related research funding but not eligible for the MBO scheme. Finally, Group 3 ("*double-treated*") comprises physicians simultaneously exposed to both interventions. Comparing the evolution of outcomes across these groups allows us to distinguish the role of monetary incentives from that of expanded research capacity.

We estimate a dynamic TWFE difference-in-differences specification comparing publication outcomes across the four groups, using Group 0 as the reference category. Rather than relying on a single post-treatment indicator, we estimate event-time coefficients that trace the evolution of outcomes before and after the introduction of the two policies. All coefficients are normalized relative to 2016, the last pre-treatment year for the first insstitutional change, the MBO, which serves as the omitted category. The estimated equation is the following:

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

where Y_{it} denotes the research output of individual i in year t. $G_i \in \{0, 1, 2, 3\}$ denotes the exposure group of physician i, with $G_i = 0$ representing the full control group. The summation over $j \in \{1, 2, 3\}$ therefore estimates dynamic effects for the MBO-only, IRCCS-only, and double-treated groups relative to Group 0. The indicators $\mathbb{1}(t = t_0 - h)$ and $\mathbb{1}(t = t_0 + g)$ denote event-time dummies for pre- and post-treatment years, respectively. Individual fixed effects, δ_i, absorb time-invariant differences across physicians, while year fixed effects, τ_t, absorb shocks common to all individuals in a given year. Standard errors are clustered at the individual level. The coefficients $\beta_{j,h}$ capture differential pre-treatment dynamics for group j relative to the control group, while the coefficients $\gamma_{j,g}$ capture the corresponding post-treatment differences. We present these coefficients graphically in event-study figures to evaluate pre-trends and the dynamic effects of the two policies.

The key identifying assumption is that, absent the two institutional changes, research productivity would have evolved similarly across the different physician groups. This comparison is facilitated by the fact that all individuals operate within the same hospital environment and are therefore exposed to the same organizational, institutional, and macroeconomic shocks. The event-study specification allows us to assess the plausibility of this assumption

by examining pre-treatment dynamics across groups.

4.2 Results

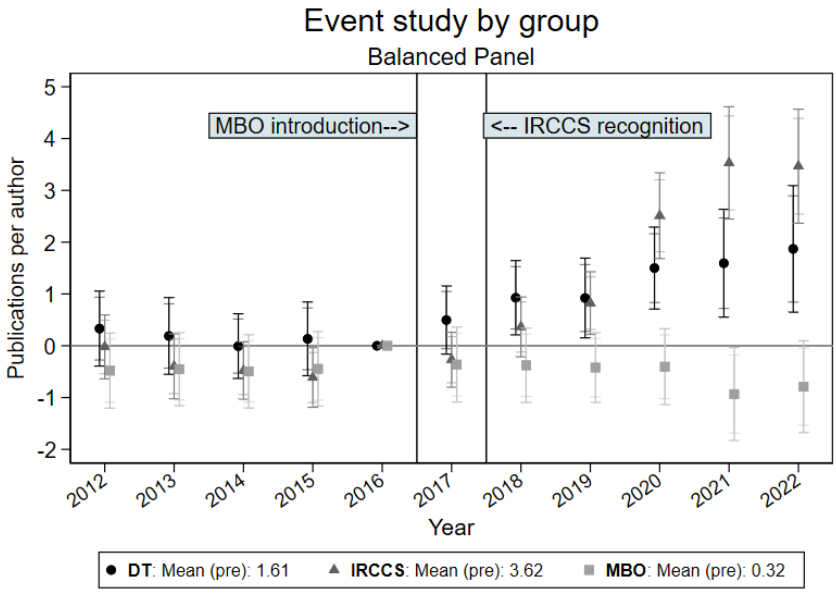


Figure 3: Event-study estimates of the effect of the MBO policy and IRCCS recognition on publications across physician groups, relative to the full control group and the pre-treatment mean of the outcome (balanced panel).

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 publication output 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 approximately one additional publication per researcher per year (about 28% relative to the pre-treatment mean). However, the effect increases steadily thereafter. By 2020, physicians included in the IRCCS perimeter exhibit an increase of roughly 2.5 publications per physician per year relative to their 2016 differential with respect to the control group, rising to more than 3 additional publications annually in the final years of the sample.

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, physicians simultaneously exposed to both policies also experience a significant increase in publication output after 2018. The effect reaches

approximately 1.5 additional publications per physician per year by 2020 ($\approx 90\%$ relative to pre-treatment mean), and remains close to 2 publications at the end of the panel. Combined with the absence of effects among MBO-only physicians, this pattern suggests that monetary incentives may be more effective when accompanied by an expansion in research capacity. While baseline publication levels differ across groups, the event-study estimates show no evidence of differential trends in the pre-treatment period, consistent with the parallel-trends assumption.

Finally, to assess whether the results are driven by sample selection arising from panel attrition, 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 limited effects of the MBO incentive scheme and the strong increase in scientific output observed among IRCCS-exposed physicians, we now focus on the impact of IRCCS recognition and the associated expansion in research capacity following the hospital's designation as an IRCCS institute in 2018.

5.1 Empirical Strategy

To estimate the impact of exposure to the IRCCS research environment on scientific productivity, we estimate a two-way fixed effects difference-in-differences specification 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 perimeter. In the baseline specification, the treatment group includes both IRCCS-only and double-treated physicians, as both groups gain access to the research resources, infrastructure, and funding associated with IRCCS recognition. The comparison group therefore consists of physicians not exposed to IRCCS-related research capacity, namely MBO-only physicians and the pure control group. 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.

Although the hospital updates the IRCCS research staff list annually, our baseline specification treats physicians who are ever included in the IRCCS perimeter as exposed from the beginning of the post-recognition period. The estimated coefficient should therefore be interpreted as an intention-to-treat (ITT) effect rather than as an average treatment effect on the treated (ATT). This choice reflects the institutional organization of research activ-

ity following IRCCS recognition. In practice, the expansion of research capacity was not allocated through a sequence of isolated individual treatments, but through the creation of a relatively stable research environment financed by IRCCS resources. Physicians formally added to the IRCCS staff list in later years were often already involved in IRCCS-related projects and collaborations before their official inclusion. Conversely, some physicians were subsequently removed from the list while continuing to participate in research activities supported by IRCCS-funded infrastructures and projects. We therefore model exposure to the IRCCS research environment as an absorbing condition, reflecting the organizational nature of the intervention rather than annual administrative changes in staff classification.

The coefficient θ captures the average ITT effect of exposure to the IRCCS research environment after 2018 for physicians belonging to the IRCCS research perimeter, relative to physicians not exposed to IRCCS-related research capacity. The main identification assumption is that, in the absence of IRCCS recognition and the associated expansion in research capacity, 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 the plausibility of this assumption through event-study specifications analogous to those presented in the previous section.

Because inclusion in the IRCCS research perimeter is determined by the institution and may partly reflect physicians' pre-existing research orientation and productivity, concerns about differential trends and endogenous selection naturally arise. Our identification strategy therefore relies on the assumption that, absent IRCCS recognition, the evolution of research productivity among physicians exposed and not exposed to the IRCCS research environment would have remained parallel. Beyond the visual inspection of pre-treatment dynamics in the event-study specifications, we assess the robustness of our results using the methodology proposed by Rambachan and Roth (2023). The *honestDiD* approach relaxes the strict parallel-trends assumption by allowing for bounded deviations from pre-treatment trends and constructs confidence sets that remain valid under moderate violations of the standard difference-in-differences identifying assumptions. Results of this analysis are discussed in Section 5.2.1.

Beyond the baseline specification, we investigate whether the effects of expanded research capacity differ across physicians facing different incentive structures. To this end, we estimate a series of alternative specifications comparing subsets of physicians exposed and not exposed to the IRCCS research environment. In particular, we consider the following pairwise group-comparisons: (1) double-treated versus MBO-only physicians; (2) double-treated versus the full control group; (3) IRCCS-only versus MBO-only physicians; (4) IRCCS-only versus the full control group; and (5) IRCCS-only versus double-treated physicians. These estimates allow us to assess whether the effects of research capacity vary according to exposure to

monetary incentives and underlying research orientation. All specifications are obtained by re-estimating Equation 2 using alternative reference groups.

We complement the baseline difference-in-differences estimates with dynamic event-study specifications that allow treatment effects to evolve flexibly over time and provide a visual assessment of pre-treatment trends. In the baseline event-study specification, the treatment indicator equals one for all physicians exposed to the IRCCS research environment, including both IRCCS-only and double-treated individuals. The control group consists of physicians not exposed to IRCCS-related research capacity, namely MBO-only physicians and the full control group. This specification therefore mirrors the baseline difference-in-differences design while allowing us to trace the dynamic evolution of research productivity around the 2018 IRCCS recognition. 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 event-study specifications typically normalize coefficients relative to the period immediately preceding treatment, we adopt the same normalization used in the previous section and omit 2016 as the reference year. This choice facilitates comparability across the different event-study specifications presented throughout the paper and allows all dynamic effects to be interpreted relative to a common pre-treatment benchmark.

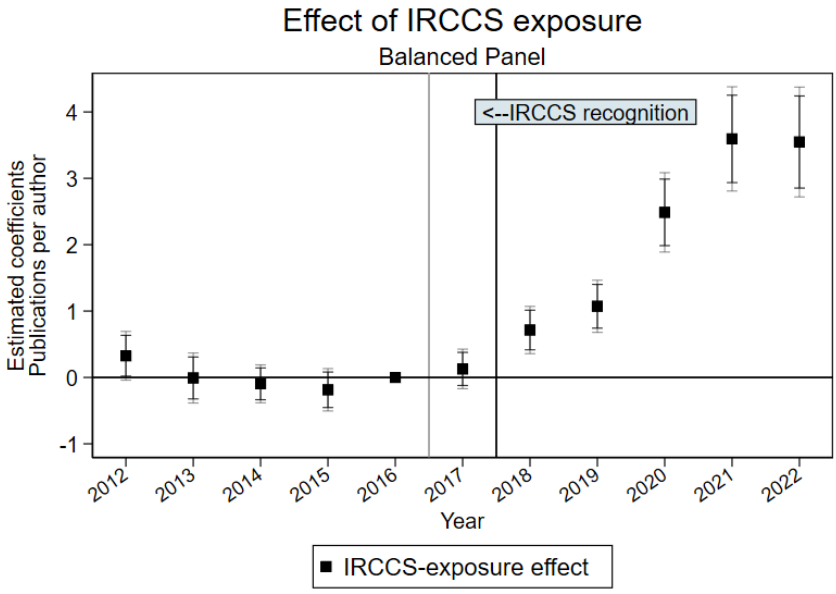
5.2 Results

	(1) IRCCS vs. others	(2) DT vs MBO	(3) DT vs PC	(4) IRCCS-only vs MBO	(5) IRCCS-only vs PC	(6) IRCCS-only vs DT
Publications	2.2557***	1.2398***	1.1365***	2.6706***	2.5569***	1.1125***
(SE)	(0.2425)	(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	3.12	1.65	1.65	1.65	3.64	3.64
Panel (Balanced)	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

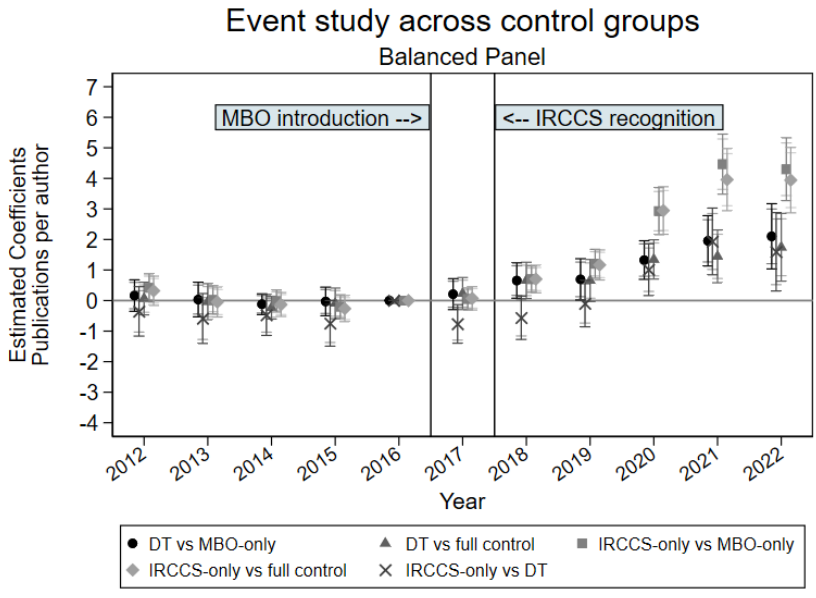
Table 3: Impact of IRCCS recognition on Annual Publications (Difference-in-Differences).

Table 3 reports the difference-in-differences estimates from Equation 2 and its alternative specifications. The results confirm the patterns documented in the initial event-study analysis and point to a substantial effect of exposure to the IRCCS research environment. In Column (1), the treatment group includes all physicians exposed to IRCCS-related research capacity, namely both IRCCS-only and double-treated individuals. The estimated coefficient

implies an increase of approximately 2.26 publications per physician per year following IRCCS recognition, corresponding to roughly 72% of the treated group's average pre-treatment publication rate. The estimate is statistically significant at the 1% level and indicates a sizeable expansion in scientific productivity associated with access to IRCCS funding, infrastructure, and research support. Columns (2)–(6) investigate heterogeneity across physicians facing different combinations of incentives and research capacity. Double-treated physicians publish approximately 1.24 more papers per year than MBO-only physicians and 1.14 more papers than the pure control group, confirming that research output increases substantially with monetary incentives only accompanied by access to IRCCS resources. At the same time, IRCCS-only physicians outperform all comparison groups. Relative to MBO-only physicians, they produce approximately 2.67 additional publications per year; relative to the pure control group, the difference amounts to 2.56 publications; and relative to double-treated physicians, IRCCS-only researchers publish about 1.11 additional papers annually. The evidence is consistent with research capacity being the main driver of the observed increase in scientific production. While the MBO scheme does not generate detectable effects on publication output when operating alone, exposure to the IRCCS research environment is associated with large productivity gains. The stronger effects estimated for IRCCS-only physicians hint that these gains are concentrated among workers already embedded in research-oriented professional trajectories.



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 reported in panel (a) of Figure 4. Treated units are physicians exposed to the IRCCS research environment, including both IRCCS-only and double-treated individuals. The comparison group consists of physicians not exposed to IRCCS-related research capacity, namely MBO-only physicians and the full control group. This specification mirrors the baseline difference-in-differences design in Table 3. Before 2018, the coefficients are small and statistically indistinguishable from zero, providing no evidence of differential pre-trends between IRCCS-exposed physicians and the comparison group. After IRCCS recognition, the estimates increase gradually. The effect is modest in 2018 and 2019, consistent with the lag between research activity and publication, but becomes substantially larger from 2020 onward. The coefficient rises sharply in 2020 and remains large in 2021 and 2022, indicating a persistent increase in publication output among physicians exposed to the IRCCS research environment.

While informative, the baseline specification combines individuals facing different incentive structures within the treated group. In particular, some physicians are exposed exclusively to the expansion in research capacity associated with IRCCS recognition, whereas others are simultaneously exposed to both IRCCS resources and the MBO incentive scheme. To assess whether responses differ across these groups, we estimate the additional pairwise event-study specifications discussed above. The corresponding estimates are reported in panel (b) of Figure 4. The results confirm the limited role of the MBO scheme in explaining the observed increase in scientific productivity. Double-treated physicians exhibit higher publication output than MBO-only physicians and the pure control group, but the magnitude of these effects remains substantially smaller than that observed for IRCCS-only researchers. The largest and most persistent increases in publication activity are consistently found among physicians exposed to the IRCCS research environment through academic research positions. Some differences in pre-treatment dynamics emerge when comparing IRCCS-only and double-treated physicians, suggesting that the two groups differ in baseline research orientation and career trajectories. While we return to this issue in the robustness analysis below, we compare our baseline estimates to a specification where we exclude individuals that had never published before 2017 from the sample, in order to make the units more comparable: Appendix Figure A5 displays that the results are basically unaffected. Importantly, this restriction does not substantially reduce sample size. As reported in Appendix Table A8, the resulting sample still contains 407 physicians, including 179 non-IRCCS physicians (57 academic, 99 non-academic) and 228 IRCCS-exposed physicians.

The larger effects observed among IRCCS-only physicians point to substantial heterogeneity in responses to research-support policies. However, since inclusion in the IRCCS perimeter is not random and may reflect prior research activity, concerns may arise regarding differential outcome dynamics across groups. We therefore assess the robustness of these

findings to potential violations of the parallel-trends assumption.

5.3 Robustness

5.3.1 Parallel Trends

Although the event-study estimates suggest no systematic violations of the parallel-trends criterion, we further assess the robustness of our results given the substantial magnitude of the estimated effects and the possibility of differential productivity dynamics across groups. Visual inspection of the baseline event-study (Figure 4) indicates that pre-treatment coefficients are generally small and statistically indistinguishable from zero. However, we observe a mild downward pattern among IRCCS-exposed physicians prior to 2016 when compared to the double-treated group, with the coefficient for 2015 approaching conventional significance levels. Importantly, this pattern goes in the opposite direction of the post-treatment effects and appears quantitatively modest relative to the magnitude of the estimated treatment effects. Nevertheless, it motivates a formal assessment of the sensitivity of our estimates to departures from the identifying assumption.

To address this issue, we implement the *HonestDiD* methodology proposed by Rambachan and Roth (2023). This approach constructs confidence sets that remain valid under bounded departures from the parallel-trends assumption. We consider two alternative restrictions. The first is the *bounded relative magnitude* (BM) approach, which allows post-treatment violations of parallel trends to be proportional to the largest deviation observed in the pre-treatment period. The second is the *smoothness restriction* (SR) approach, which limits the extent to which post-treatment trend deviations can depart from a smooth extrapolation of pre-treatment dynamics. Across both sets of assumptions, the estimated effects remain statistically distinguishable from zero for a wide range of permissible violations, indicating that the main conclusions are not driven by small deviations from common pre-trends. Detailed results and graphical representations of the sensitivity analysis are reported in Appendix Figures A7–A13.

5.3.2 Validity of the TWFE 2x2 DiD identification

Although IRCCS recognition occurred at the institutional level in 2018, the hospital updates the list of IRCCS researchers annually. While a large core group entered the perimeter immediately after recognition, additional physicians were formally added in subsequent years (Figure A2). This generates a partially staggered treatment assignment. Appendix Table A4 reports the composition of the IRCCS group, while Table A5 shows that late entrants differ significantly from the initial cohort along several observable dimensions. This feature raises the standard concern that TWFE estimators may be biased under staggered adoption when

treatment effects are heterogeneous across cohorts or over time. In our setting, however, staggered entry primarily reflects administrative updates to the IRCCS research staff list rather than distinct policy interventions, and most treated observations belong to the original 2018 cohort. Nevertheless, we assess the sensitivity of our results to treatment timing and cohort composition.

In the baseline analysis, treatment is defined according to an *ever-treated* criterion: physicians who are included in the IRCCS registry at any point after 2018 are considered treated from 2018 onward. The resulting estimand should therefore be interpreted as an intention-to-treat (ITT) effect:

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

where D_{ik} denotes the treatment status of individual i in period k.

We first re-estimate the event-study specification after excluding physicians who entered the IRCCS perimeter after 2018. Appendix Figures A17 and A18 show a dynamic pattern very similar to the baseline estimates, with substantially larger post-treatment effects. This suggests that the baseline ITT estimates likely understate the productivity gains experienced by the original IRCCS cohort. Conversely, when the original 2018 cohort is excluded and only late entrants are retained (Appendix Figures A19 and A20), publication output increases only from 2020 onward. This delayed response is consistent with both their later integration into the IRCCS research environment and the expansion of research opportunities during the COVID-19 period.

We then exploit the staggered inclusion of physicians into the IRCCS perimeter and decompose the TWFE estimator following Goodman-Bacon (2021). The decomposition reveals that 84% of the identifying variation originates from treated-versus-never-treated comparisons, while only 16% comes from comparisons across treatment cohorts. Moreover, both components yield positive effects of similar magnitude, suggesting that the baseline estimates are not driven by problematic comparisons between early and late adopters (Appendix Figure A16).

Finally, we compare the baseline TWFE estimates with the imputation-based estimator of Borusyak et al. (2024), which is robust to treatment-timing heterogeneity. As shown in Appendix Figure A21, the two approaches yield highly similar post-treatment effects. The negative pre-treatment drift visible under TWFE is substantially attenuated under the imputation-based estimator, while the coefficient immediately preceding treatment remains positive. This pattern is consistent with the institutional process governing entry into the IRCCS research staff, whereby formal inclusion may follow an already ongoing involvement in IRCCS-funded research activities. More importantly, the magnitude and evolution of post-treatment effects remain highly comparable across estimators, indicating that treatment-

timing heterogeneity does not materially affect the conclusions of the analysis. Similar conclusions emerge when the original 2018 cohort is excluded.

As a final check, we implement the *HonestDiD* procedure of Rambachan and Roth (2023) on the staggered event-study estimates. The resulting confidence sets confirm that both the TWFE and Borusyak estimates remain statistically distinguishable from zero under plausible departures from parallel trends (Appendix Figures A22 and A23).

5.3.3 Robustness Checks

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 A11. 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[7]. The resulting estimates, reported in Figures A31 and A32, remain qualitatively unchanged, suggesting that the observed effects are not driven by pandemic-related publication dynamics.

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 A12 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

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

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.

6 Who Responds to Research Capacity?

The previous section established that IRCCS recognition substantially increased scientific output. We now investigate whether these effects are concentrated among specific types of researchers or are broadly distributed across the physician population.

6.1 Baseline Productivity

We investigate whether the effects of IRCCS recognition vary across the pre-treatment productivity distribution. Physicians are classified according to terciles of their average publication output before 2017, and separate event-study specifications are estimated for each group. Figure 5 reports the corresponding estimates.

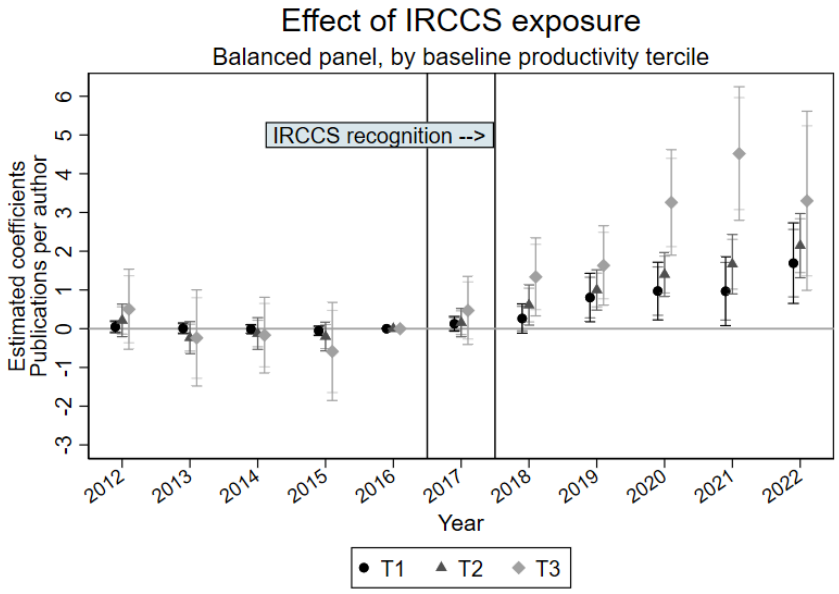


Figure 5: Event-study of the IRCCS effect by baseline productivity tercile (balanced panel).

The results reveal substantial heterogeneity across the productivity distribution, although the response to IRCCS recognition remains positive throughout. Researchers in all terciles experience an increase in publication output following the reform, suggesting that the expansion of research resources did not benefit only a narrow group of already highly productive scientists.

At the same time, the magnitude of the response differs across groups. Physicians in the upper tercile exhibit the largest gains, with publication output increasing sharply after 2019 and remaining persistently above pre-treatment levels. By contrast, researchers in the lower and middle terciles display more gradual increases. These patterns suggest that additional research resources are productive throughout the distribution, but generate larger absolute returns among individuals with stronger pre-existing research activity.

Importantly, the positive response observed among lower-productivity researchers indicates that the effects of IRCCS recognition are not confined to a small group of top performers. Rather, the policy appears to have expanded research activity across a broad segment of the hospital's research workforce. Additional analyses excluding physicians with no pre-treatment publications yield very similar results (Appendix Figure A6), indicating that the estimated effects are not driven by comparisons between active researchers and individuals with no prior scientific activity.

To better characterize the composition of the IRCCS research perimeter, we further examine physicians who were formally included only after the initial 2018 recognition. Appendix Table A9 shows that these late entrants were substantially less productive than the original 2018 cohort prior to the reform. On average, they produced 1.6 publications per year before 2017, compared with almost 4 publications among physicians included from the outset. Nevertheless, the vast majority of late entrants had already engaged in research activity before formal inclusion in the IRCCS perimeter: nearly 90% had at least one pre-2017 publication, a share comparable to that observed among the original cohort. This pattern suggests that subsequent expansions of the IRCCS perimeter primarily reflected the gradual incorporation of already active researchers rather than the recruitment of previously inactive physicians.

Appendix Table A10 further examines the correlates of late entry into the IRCCS registry. Physicians with lower baseline productivity were significantly more likely to enter the IRCCS perimeter after 2018, consistent with the descriptive evidence above. Although late entrants were more frequently eligible for the MBO scheme, the association becomes statistically imprecise once baseline research productivity is taken into account, which suggests that the expansion of the IRCCS research staff primarily involved the formalization of a pre-existing research workforce rather than the creation of new researchers through the incentive scheme.

6.2 Age Cohorts

We further investigate whether the effects of IRCCS recognition vary across the career distribution. Figure 6 reports static two-way fixed effects difference-in-differences estimates separately by age group, using physicians' age in 2017 to define six categories (35–39, 40–44, 45–49, 50–54, 55–59, and 60+ years old). The figure compares IRCCS-only, MBO-only, and double-treated physicians to the pure control group. rThe effects of IRCCS recognition

are remarkably stable across the age distribution. Positive and economically meaningful increases in publication output are observed in every age category, with estimated coefficients ranging from about two to three additional publications per physician per year. The largest point estimates emerge among physicians aged 50–59, although confidence intervals overlap substantially across cohorts: the expansion of research resources benefited both younger and more senior physicians rather than being concentrated at a specific career stage.

Double-treated physicians also display positive responses at younger and middle-career ages, although the estimates become substantially less precise among older cohorts. Given the relatively limited number of double-treated physicians within each age category, these differences should be interpreted cautiously[8]

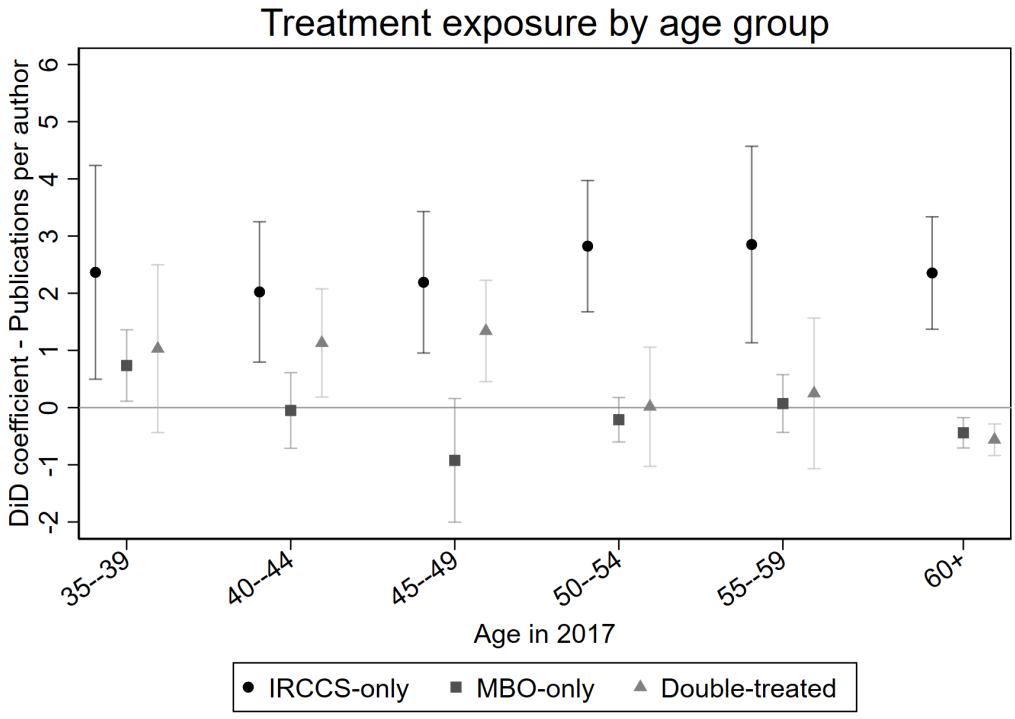


Figure 6: Treatment exposure by age group. Static TWFE DiD estimates of the effect of IRCCS recognition and MBO incentives on annual publications, by physicians' age in 2017. Estimates are reported separately for IRCCS-only, MBO-only, and double-treated physicians relative to the pure control group. Vertical bars represent 95% confidence intervals.

[8]Additional heterogeneity analyses by gender, tenure, and median age reveal no substantial differences in treatment responses across groups and are therefore omitted for brevity. Estimates are available upon request.

7 Research Scaling or Transformation?

7.1 Research Specialization

7.1.1 Volume by Field

A central question is whether IRCCS recognition altered the direction of research activity or primarily expanded existing research programs. To address this issue, we exploit the WoS-based classification of publications and construct, for each author-year observation, the number of publications produced within a given clinical-scientific area. Because the original WoS classification is highly granular, fields are aggregated into broader macro-specializations.[9]

Table 4 reports the estimated effects of IRCCS recognition on field-specific publication output. Each row corresponds to a separate difference-in-differences specification in which the dependent variable is the individual annual number of publications produced within a specialization. The reported coefficient represents the intention-to-treat (ITT) effect of IRCCS exposure on research output in that field. Standard errors are clustered at the author level. To account for the large number of outcomes examined, we report both conventional p-values and Romano–Wolf stepdown adjusted p-values, which correct for multiple hypothesis testing across the family of specializations (Romano and Wolf, 2005, 2016). We additionally report a joint test of pre-treatment coefficients to assess the presence of differential pre-trends.

The estimates indicate that the increase in research output following the IRCCS recognition is broad-based rather than concentrated within a small set of disciplines. Positive and statistically significant effects emerge across a wide range of scientific areas, including Biology, Cardiology, Diagnostics and Imaging, Infectious Diseases, Internal Medicine, Pharmacology, Public Health, and Oncology. Many of these effects remain statistically significant after Romano–Wolf adjustment, suggesting that the results are not driven by multiple testing concerns. The strongest and most robust effects are observed in several of the hospital's most research-intensive domains, where a substantial research infrastructure and scientific capacity possibly existed already. In particular, publication output strongly increases in Internal Medicine, Biology, Infectious Diseases, Cardiology, and Public Health, with no evidence of differential pre-treatment trends. Large positive effects are also observed in Oncology and Hematology as well as Diagnostics and Imaging. However, these estimates should be interpreted more cautiously, as the corresponding event-study specifications display some evidence of pre-treatment dynamics. While the post-treatment increases remain sizeable and economically meaningful, part of the observed growth may reflect trends which predate the formal IRCCS recognition.

The dynamic event-study estimates reported in Appendix Figures A24-A26 corroborate

[9]The methodology used to construct these macro-categories is reported in Appendix Table A6.

these findings. Across most disciplines, publication output begins to increase only after IRCCS recognition and follows a similar post-treatment trajectory, with little evidence of anticipatory behavior, thus further suggesting that the institutional change generated a significant scaling-up of research activity within existing areas of expertise.

Specialization	**ITT**			**Pre (joint)**		**Mean Pre**	**Comment**
	Coef.	**p-value**	**RW p-val.**	**F-stat**	**p-value**		
Biology	0.201***	0.001	0.008***	0.839	0.501	0.166	Expansion
	(0.058)						
Cardiology/Cardiovasc. Diseases	0.160***	0.002	0.008***	1.852	0.117	0.179	Expansion
	(0.051)						
Chemistry and Physics	0.070***	0.001	0.008***	0.354	0.841	0.022	Expansion
	(0.021)						
Diagnostics and Imaging	0.187***	0.000	0.006***	2.297	0.058*	0.143	Ambiguous
	(0.046)						
Infectious Dis./ Immunology	0.159***	0.001	0.008***	0.543	0.704	0.084	Expansion
	(0.046)						
Internal Medicine (Generic)	0.345***	0.000	0.001***	0.639	0.635	0.124	Expansion
	(0.054)						
Internal Medicine (Specialties)	0.344***	0.000	0.001***	1.806	0.126	0.259	Expansion
	(0.059)						
Maternal and Child Health	0.155**	0.022	0.033**	2.104	0.079*	0.139	Ambiguous
	(0.068)						
Neurosciences and Psych.	0.109**	0.037	0.033**	0.593	0.668	0.245	Expansion
	(0.052)						
Oncology and Hematology	0.455***	0.000	0.008***	2.834	0.024**	0.236	Ambiguous
	(0.123)						
Pharmacology	0.098***	0.001	0.008***	1.216	0.303	0.068	Expansion
	(0.029)						
Public Health and Health Systems	0.175***	0.000	0.001***	1.407	0.230	0.032	Expansion
	(0.034)						
Rehabilitation and Nursing	0.038***	0.002	0.008***	0.846	0.496	0.021	Expansion
	(0.012)						
Social Sciences and Humanities	0.009	0.421	0.475	1.164	0.325	0.014	Null
	(0.011)						
Surgery and Procedures	0.034	0.521	0.475	1.035	0.388	0.192	Null
	(0.053)						
Technology and Engineering	0.032	0.243	0.298	2.686	0.031**	0.074	Ambiguous
	(0.028)						

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.

7.1.2 Field Composition

The previous subsection shows that IRCCS recognition increased publication output across a broad range of scientific fields. An important question, however, is whether this expansion

was accompanied by a transformation in the composition of the physicians' research activity. From a policy perspective, one would expect research-intensive healthcare institutions not only to increase scientific production, but also to foster translational and interdisciplinary research across biomedical fields. If successful, these processes could generate changes in the distribution of research effort across scientific domains. To investigate the issue, Table A7 examines the effects of IRCCS recognition on the share of publications attributable to each specialization within physicians' publication portfolios. Because publications in the Web of Science are frequently assigned to multiple subject categories, changes in these shares should not be interpreted as physicians switching clinical specialties, but rather as changes in the relative scientific composition of their research output. The results provide little evidence of systematic reallocation. Across virtually all scientific areas, estimated effects on publication shares are economically small and statistically insignificant. Moreover, none of the estimated effects remains statistically significant after Romano–Wolf correction for multiple hypothesis testing. Physicians appear to publish more research after IRCCS recognition, but largely within the same areas of specialization in which they were already active.

The event-study estimates reported in Appendix Figures A27-A29 reinforce this interpretation. Publication shares remain broadly stable throughout the sample period, with no evidence of systematic structural breaks around the time of IRCCS recognition. Descriptive evidence on the aggregate composition of the hospital's research portfolio, reported in Appendix Figure A30, points to a similar conclusion. The figure reports the annual distribution of Web of Science (WoS) field mentions across macro-specializations. While some fluctuations are visible over time, these are largely confined to closely related medical classifications—particularly within Internal Medicine—rather than reflecting systematic shifts across broad scientific domains. Consistent with the results in Table A7, the overall composition of research activity remains remarkably stable throughout the period, suggesting that IRCCS recognition primarily generated research scaling rather than substantial research reorientation.

7.2 Research Organization

7.2.1 Team and network expansion

We next examine whether the increase in scientific output following IRCCS recognition reflects a transformation of research networks or, more simply, a scaling-up of existing research activity. To this end, we construct yearly collaboration outcomes at the researcher level. Our main measures count coauthorship appearances, separately for internal collaborators affiliated with the hospital and external collaborators. For each dimension, we distinguish between the total number of collaboration links and the number of links involving collabora-

tors who had not previously appeared with the focal researcher. We also compute the ratio between new and existing links, which captures whether newly formed relationships become more important relative to pre-existing ones[10].

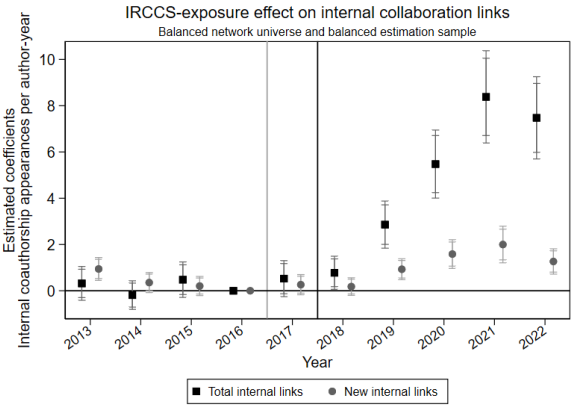
Figure 7 reports the dynamic effects of IRCCS exposure on these collaboration outcomes. Panel (a) shows a large and persistent increase in internal collaboration links after the recognition. The effect is already visible shortly after treatment and becomes sizeable from 2019 onward, indicating that treated researchers increasingly produce papers with internal coauthors. Panel (b) shows an even stronger expansion of external collaboration links. This pattern is consistent with a broader scale-up of research production: as publication activity increases, treated researchers appear more frequently in larger collaborative teams, both inside and outside the hospital.

The evidence on new links points in the same direction, but with an important qualification. The number of new internal and external links also rises after treatment, especially in the later post-treatment years. However, these increases are much smaller than the corresponding increases in total links. In other words, IRCCS recognition does not only reproduce exactly the same collaborations, but the dominant margin is the intensification and enlargement of existing collaborative activity rather than a sharp reconfiguration of researchers' networks.

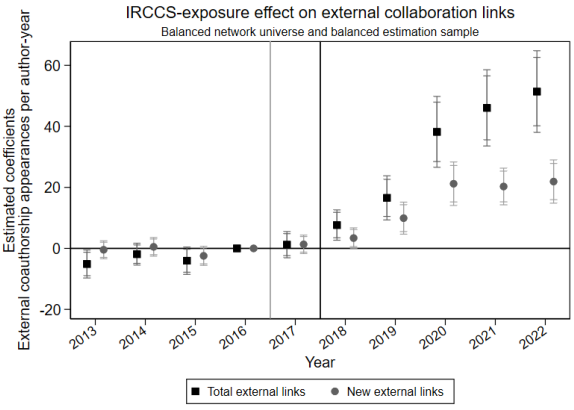
Panels (c) and (d) make this interpretation more direct. The ratio of new-to-existing internal links does not display a systematic post-treatment increase. The same is true for external links, where estimates are noisy and remain mostly statistically indistinguishable from zero. Hence, although treated researchers form additional collaborations in absolute terms, new relationships do not become more important relative to pre-existing ones. The collaboration response therefore appears proportional to the overall expansion of research output rather than indicative of a structural change in the organization of research.

The policy increased the volume of collaborative production and expanded research teams in absolute terms, but it did not fundamentally reshape collaboration portfolios. This interpretation is consistent with the broader evidence in the paper: the treatment raises scientific output, yet the increase operates mainly through the expansion of existing research capacity rather than through new research trajectories, new network structures, or a reorganization of the scientific production process.

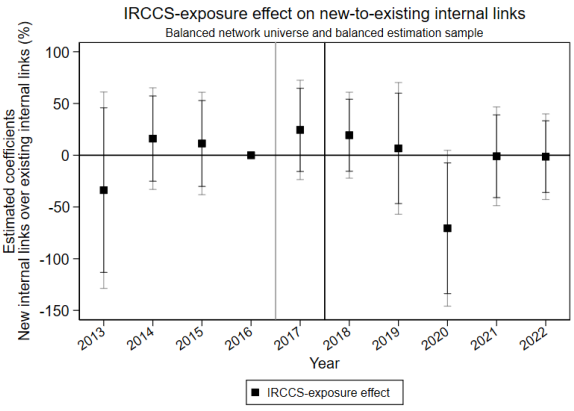
[10]The main estimates are based on the balanced physician panel and use coauthorship appearances as the unit of measurement. This captures the intensity of collaborative production rather than the number of distinct collaborators. Appendix Figure A33 reports analogous estimates using the unbalanced panel. Appendix Figure A34 reports robustness checks based on unique collaborators rather than coauthorship appearances. The qualitative conclusions are unchanged. Because 2012 is the first year in which collaborations are observed, new-link outcomes are estimated excluding 2012 to avoid mechanically classifying pre-sample collaborations as new.



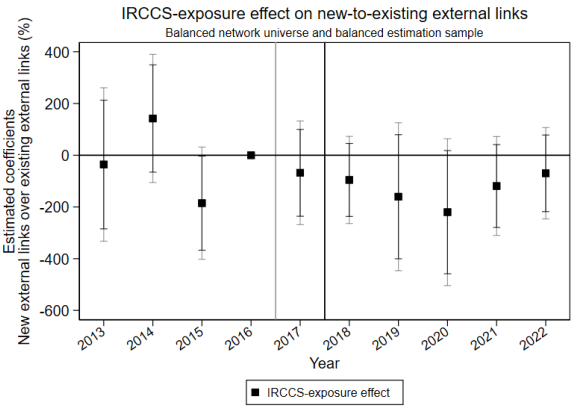
a) Internal collaboration links.



b) External collaboration links.



c) New-to-existing internal links.



d) New-to-existing external links.

Figure 7: Event-study estimates of the effect of IRCCS recognition on team and network expansion.

Cross- and within-group co-authorship

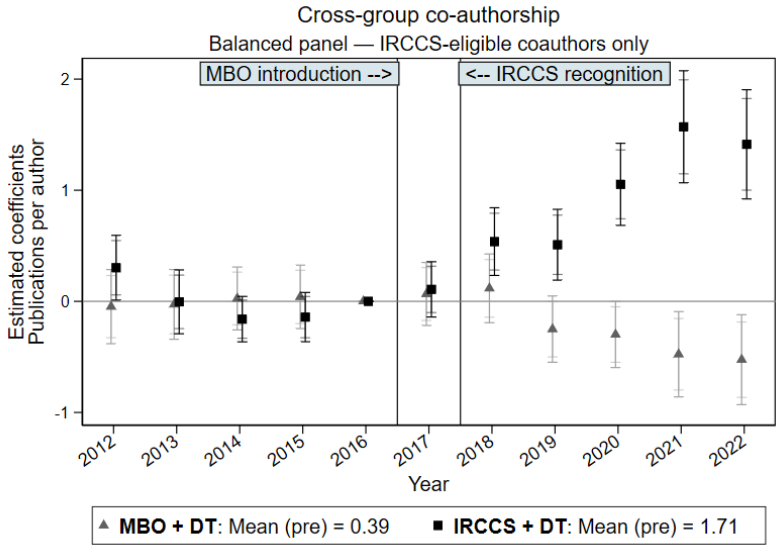
We next investigate whether the policies affected collaboration patterns across different groups within the institution. In particular, we assess whether the increase in research output was accompanied by greater collaboration between IRCCS-only researchers and MBO-exposed physicians, or whether it primarily reflected stronger production within the IRCCS research network.

We define as "networked" all publications featuring at least two co-authors affiliated with the institution and classify them according to the institutional group of the co-authors involved. We distinguish between: (*i*) publications involving IRCCS-only co-authors (*Type A*); and (*ii*) publications involving MBO-only co-authors (*Type B*). Double-treated physicians are retained in the estimation and enter both policy-specific exposure indicators. Thus, the IRCCS indicator comprises IRCCS-only and double-treated physicians, while the MBO indicator comprises MBO-only and double-treated physicians. The two sets of dynamic effects are estimated jointly following Equation 1, with physicians exposed to neither policy forming the reference group.

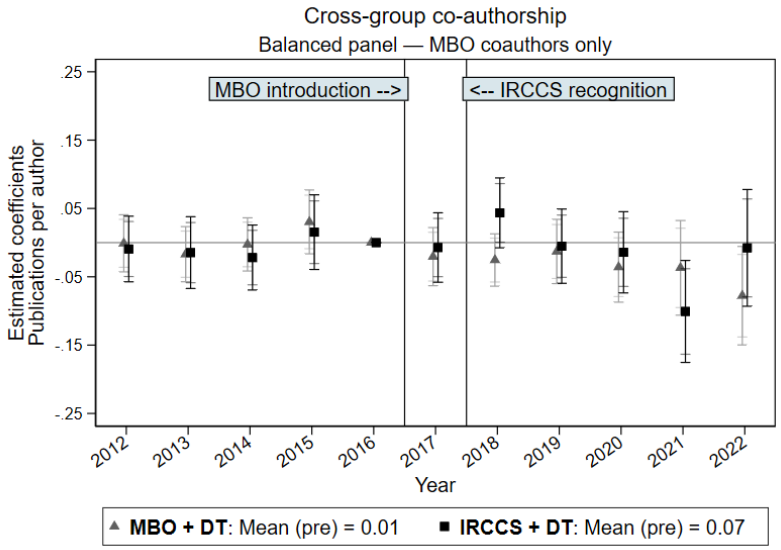
Figure 8 reports the results for the balanced panel. Panel (a) considers publications involving IRCCS-only co-authors, while panel (b) considers publications involving MBO-only co-authors. The corresponding estimates for the unbalanced panel are reported in Appendix Figure A43.

The pre-treatment coefficients are generally close to zero and do not display systematic trends. After IRCCS recognition, publications involving IRCCS-only co-authors increase substantially and persistently among IRCCS-exposed researchers. By contrast, the corresponding coefficients for the MBO-exposed group remain close to zero or become negative. Thus, the expansion of IRCCS research activity is concentrated within the IRCCS-only network: there is no evidence that MBO-exposed physicians, including double-treated physicians, increase their collaboration with IRCCS-only researchers.

Panel (b) provides complementary evidence. Publications involving MBO-only co-authors do not exhibit a sustained increase for either policy group. In particular, IRCCS-exposed researchers do not expand their collaboration with MBO-only physicians following recognition. The absence of a positive response in either direction therefore indicates that the two institutional groups remain largely separated.



(a) Outcome: publications involving IRCCS-only co-authors.



(b) Outcome: publications involving MBO-only co-authors.

Figure 8: Dynamic effects of the policy interventions on annual co-authored publications by co-author group in the balanced panel.

The evidence points to an expansion of research production within the IRCCS-only network rather than to greater integration across institutional groups. IRCCS recognition strengthens collaboration among researchers already connected to the research-oriented component of the institution, but does not generate comparable links with MBO-only or MBO-exposed physicians. This result is consistent with the SUTVA robustness analysis and supports an interpretation based on output scaling within established research structures, rather than a transformation of the institution's collaboration network.

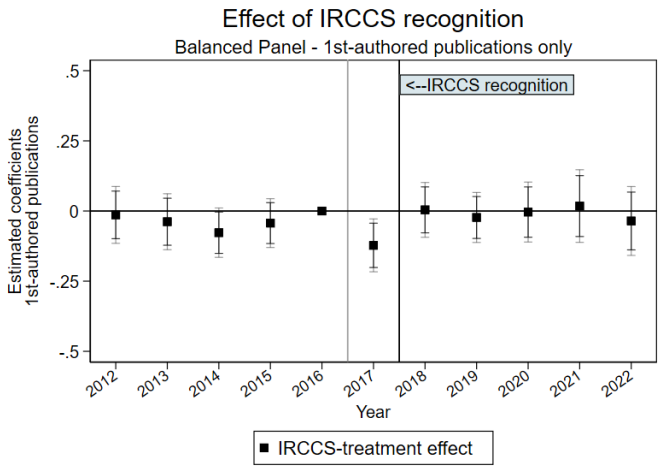
Co-authorship structure

Having established that IRCCS recognition expands research activity within the IRCCS network, we next examine whether this expansion is accompanied by changes in the organization of scientific production. We focus on researchers' positions in the author list, distinguishing between first-, middle-, and last-authored publications. In line with the main specification, the IRCCS-treated group includes both IRCCS-only and double-treated physicians, while non-IRCCS physicians form the comparison group. Figure 9 reports the corresponding event-study estimates for the balanced panel.

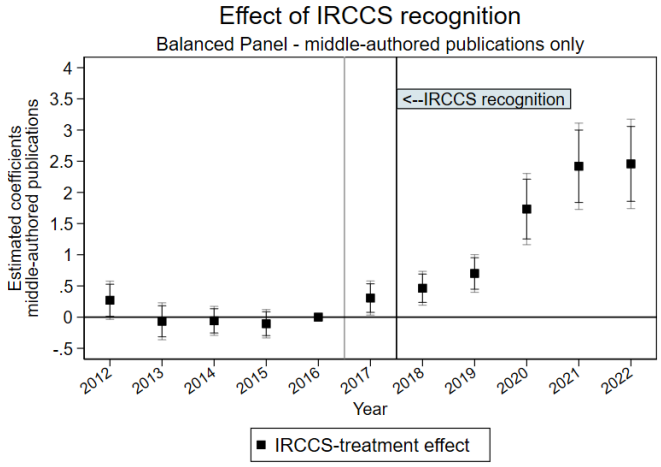
The results reveal substantial heterogeneity across authorship positions. First-authored publications display no systematic response to IRCCS recognition. The post-treatment coefficients remain close to zero and are imprecisely estimated, indicating that the increase in research output is not driven by a greater number of publications in which treated researchers occupy the leading authorship position.

Middle-authored publications, by contrast, increase substantially after recognition. The effect emerges immediately after 2018, becomes progressively larger over time, and reaches its highest magnitude in the final years of the sample. This is consistent with a broad expansion of treated researchers' participation across collaborative projects. Rather than producing a generalized shift toward first authorship, IRCCS recognition appears to increase the number of projects in which researchers contribute as members of larger research teams.

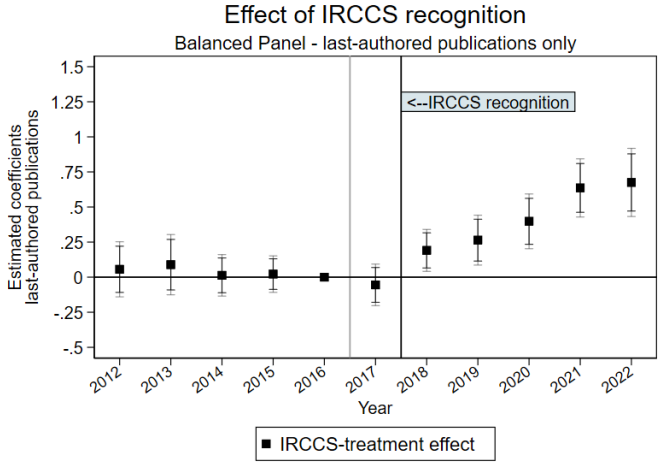
Last-authored publications also exhibit a positive and persistent increase. Although smaller in magnitude than the effect on middle authorship, the coefficients are statistically discernible from 2018 onward and grow over the post-treatment period. In the biomedical context, where last authorship commonly denotes a senior or supervisory role, this pattern suggests that IRCCS recognition also increases researchers' involvement in the coordination of scientific projects. Nevertheless, the dominant margin of adjustment remains the expansion of middle-authored contributions.



(a) Outcome: first-authored publications.



(b) Outcome: middle-authored publications.



(c) Outcome: last-authored publications.

Figure 9: Dynamic effects of IRCCS recognition on authorship positions among IRCCS-exposed researchers, including double-treated physicians (balanced panel).

Across the three outcomes, the pre-treatment coefficients do not display systematic trends. Therefore, we observe an increase in both participation and senior coordination, but not in first-authored leadership. This composition is consistent with a process of output scaling through larger and more active research teams, rather than a generalized transformation of individual researchers' roles within the authorship hierarchy. Specifications using alternative treatment and comparison groups yield similar qualitative patterns and are reported in Appendix Figure A44. Consistent with this interpretation, Appendix Figure A45 shows that IRCCS recognition also increases the average number of co-authors per publication, confirming that the expansion operates partly through larger research teams.

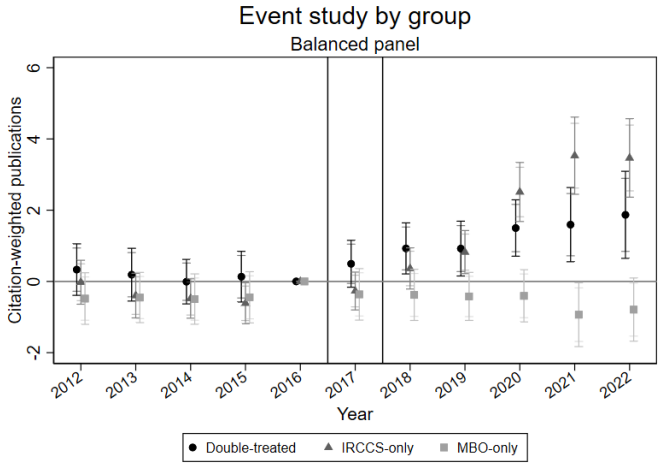
7.3 Research Impact

Citations

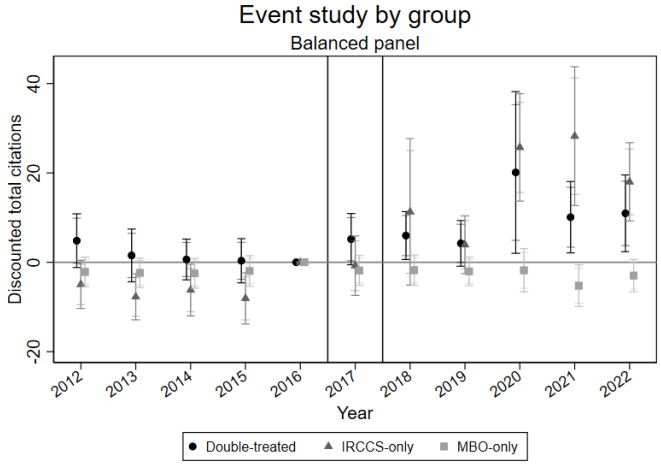
A further question is whether the increase in research output documented in the previous section translated into a proportional increase in scientific impact. To address this issue, we use citations as a proxy for research quality. Because citations are observed in Fall 2024 for papers published in different years, we construct age-adjusted citation measures by discounting citations according to the time elapsed between publication year t and 2024.

We consider three complementary outcomes. First, citation-weighted publications capture whether treatments increase publication output in journals that subsequently receive more citations. Second, discounted total citations measure the aggregate scientific impact generated by a physician in a given year. Third, discounted citations per publication measure the average impact of each paper and therefore provide our preferred indicator of research quality.

Unlike the baseline specifications, Figure 11 estimates treatment effects separately for MBO-only physicians, IRCCS-only researchers, and double-treated physicians, using individuals exposed to neither policy as the reference group. This decomposition allows us to distinguish whether the observed effects are driven by research capacity expansion, monetary incentives, or the combination of both.

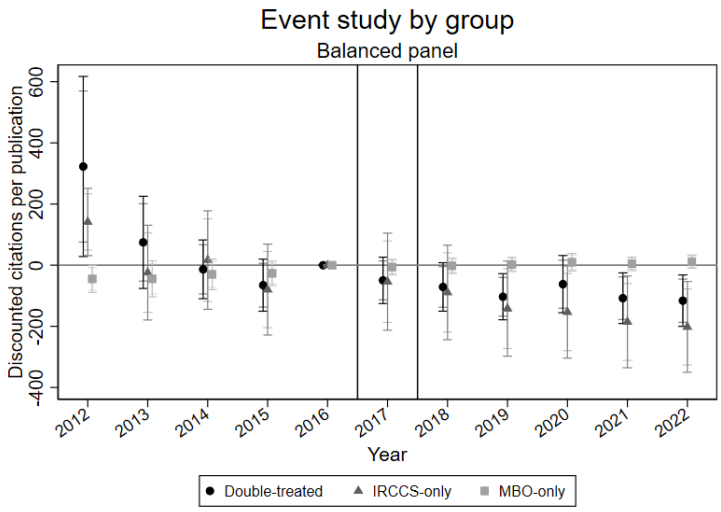


(a) Citation-weighted publications.



(b) Discounted total citations.

Figure 10: Event-study estimates of treatment effects on citation-weighted publications and discounted total citations. The omitted category consists of physicians exposed to neither MBO nor IRCCS recognition.



Discounted citations per publication.

Figure 11: Event-study estimates of treatment effects on discounted citations per publication. The omitted category consists of physicians exposed to neither MBO nor IRCCS recognition.

The three panels reveal a remarkably consistent pattern. Following IRCCS recognition, both IRCCS-only researchers and double-treated physicians experience sizeable increases in citation-weighted publications and discounted total citations. These results indicate that the expansion in publication output documented in the previous section also translates into a higher aggregate scientific impact. By contrast, MBO-only physicians display little evidence of changes in either outcome, consistent with the absence of substantial effects on publication volumes.

The picture changes when considering citations per publication. While IRCCS-only researchers and double-treated physicians exhibit an increasingly negative trajectory after recognition, MBO-only physicians remain close to zero throughout the post-treatment period. Thus, although IRCCS recognition generates more publications and more total citations, the average scientific impact of each publication declines.

The findings point to a potential quantity–quality trade-off associated with the IRCCS policy. The increase in aggregate scientific impact is driven primarily by a larger volume of publications rather than by improvements in the average influence of individual papers. In contrast, the MBO incentive scheme neither substantially expands research output nor generates systematic changes in average citation performance.

These results reinforce the interpretation developed throughout the paper. IRCCS recognition primarily induces *research scaling*: researchers publish more papers and accumulate more citations overall, but the organization and average impact of research change relatively little. The policy therefore expands existing research capacity rather than fundamentally transforming scientific production.

For completeness, Appendix Figure A46 reports the corresponding pooled event-study estimates, comparing all IRCCS-exposed researchers with non-IRCCS researchers and all MBO-exposed physicians with non-MBO physicians. The pooled estimates lead to the same qualitative conclusions while masking the heterogeneity across treatment groups.

Journal Quality and Strategic Responses

The previous section documented a quantity–quality trade-off following IRCCS recognition: researchers publish more papers and accumulate more citations overall, yet citations per publication decline. A natural concern is that this pattern simply reflects a shift toward lower-quality journals. Alternatively, physicians exposed to the MBO scheme may strategically manipulate the set of publications used to compute the performance bonus without substantially changing their underlying research activity. We investigate both mechanisms using administrative data on bonus calculations and publication-level bibliometric indicators.

Specifically, we exploit administrative records on individual payouts, Impact Factor scores, and the set of publications deemed eligible for the MBO evaluation over the period 2017–2022. Since historical Impact Factor information is unavailable before the introduction of the scheme, we estimate two-way fixed-effects models comparing MBO-treated physicians with untreated researchers after 2017. An important institutional feature is that the list of eligible publications is not generated mechanically from bibliometric databases. Instead, physicians can revise the list used for evaluation by correcting attribution errors or reporting omitted publications. Consequently, the number of eligible publications may differ from the complete publication record observed in bibliometric data.

This institutional setting allows us to examine two distinct margins of adjustment. First, we study whether physicians improve journal placement by considering the total Impact Factor score (adjusted for co-authorship). Second, we compare eligible and total publications to detect potential strategic reporting aimed at maximizing the bonus. Finally, we analyse the realized monetary payout associated with the MBO evaluation.

Table A13 provides little evidence that the decline in citations documented for IRCCS researchers reflects publication in systematically lower-ranked journals. Across specifications, Impact Factor scores remain broadly unchanged for MBO-treated physicians relative to comparable untreated researchers. Likewise, the number of eligible publications closely mirrors the total number of publications, and we find no robust evidence of systematic discrepancies between the two measures. The realized payouts also display no pattern consistent with strategic optimization of the incentive scheme.

Hence, physicians are seemingly not responding to the MBO incentive by manipulating the set of publications considered for evaluation, nor by systematically altering journal placement. Combined with the evidence from the previous subsection, these results indicate that the

decline in citations per publication observed after IRCCS recognition is unlikely to be driven by publication in lower-impact journals. Instead, research output expands while journal quality remains broadly stable, implying that the lower average citation performance reflects the marginal impact of the additional publications rather than a deterioration in publication outlets.

Additional Evidence on Hospital Specialization

As a final descriptive exercise, we examine whether IRCCS recognition coincides with changes in the hospital's clinical specialization. We use hospital-level indicators from the *Programma Nazionale Esiti* (PNE), focusing on two families of outcomes. The first includes standard quality indicators, such as mortality and readmission rates. The second includes volume-based measures of high-complexity and technology-intensive procedures, normalized by hospital beds.

The analysis compares the treated hospital with other hospitals in Lazio using hospital and year fixed effects. These estimates should be interpreted cautiously. The exercise is descriptive rather than causal: the pre-treatment period is short, several outcomes display visible pre-trends, and the timing of changes may partly reflect ongoing specialization processes rather than a discrete effect of IRCCS recognition.

Appendix Table A14 reports results for quality indicators. The evidence is mixed and does not point to a systematic improvement or deterioration in clinical quality. Some outcomes improve, others worsen, and several estimates are affected by significant pre-trends.

Appendix Table A15 reports results for procedure volumes. Here the pattern is more suggestive. Several high-complexity procedures increase after 2018, especially in oncology and minimally invasive surgery. However, some of these indicators also exhibit pre-existing trends, limiting causal interpretation.

Figures A47 and A48 display event-study estimates with extrapolated pre-treatment trends following (Roth, 2022). The figures confirm the main qualification: for several outcomes, post-2018 increases are difficult to separate from pre-existing trajectories. The clearest descriptive patterns are observed for minimally invasive lung cancer surgery and carotid endarterectomy, where pre-treatment dynamics are relatively flat and post-2018 increases are more pronounced.

IRCCS recognition seemed to be coinciding thus with a broader shift toward more specialized and technology-intensive clinical activity. However, given the short pre-period and the presence of pre-trends, these results should be read as descriptive institutional background rather than as causal evidence.

8 Conclusions

This paper studies how different forms of research support affect scientific production within a large healthcare organization. Exploiting two institutional reforms implemented in a major Italian teaching hospital—a publication-based monetary incentive scheme targeted at non-academic physicians and the hospital's recognition as an IRCCS institute in 2018—we compare policies operating through individual incentives with policies operating through research capacity.

Three main findings emerge. First, publication-based monetary incentives alone do not significantly increase scientific output among non-academic physicians. By contrast, exposure to the IRCCS research environment and the associated expansion in public funding, infrastructure, and organizational support generates large and persistent increases in publication activity. Physicians exposed to both interventions also experience productivity gains, suggesting that monetary incentives may become effective when researchers simultaneously benefit from adequate research capacity. A key takeway is that relaxing resource constraints may be more important than increasing private rewards in settings where research competes with demanding clinical responsibilities.

Second, we document substantial heterogeneity in responses to research-support policies. Academic physicians and researchers already embedded in research-oriented environments react strongly to the expansion of research capacity, whereas non-academic physicians display little response to monetary incentives alone. These findings complement the evidence of Checchi et al. (2021), who show that responses to academic incentives depend on researchers' underlying characteristics and career prospects. They also qualify the results of De Philippis (2021), who find that stronger research incentives increase scientific production in academic environments. While we similarly observe positive effects on research output, these arise primarily when researchers operate within an organizational setting that provides dedicated resources and research support. This suggests that the effectiveness of research incentives depends on the severity of multitasking constraints and on the relative importance of competing professional activities. Finally, our results resonate with Lepinteur and Nisticò (2025), who emphasize the importance of access to productive research inputs. In our setting as well, the benefits of expanded research capacity are concentrated among physicians already integrated into research-oriented environments, highlighting the joint importance of organizational context, research opportunities, and individual career incentives in shaping the effectiveness of research policies.

Third, we show that increases in scientific output need not imply a transformation of scientific activity. While IRCCS recognition substantially increases publication volumes, the additional research is largely produced within existing teams, collaboration networks, and

scientific fields. We find little evidence of major changes in research leadership, collaboration structures, or field composition. Consistent with this interpretation, total citations increase because of higher publication volumes, while citations per publication decline. The expansion of scientific activity therefore reflects research scaling rather than research transformation. Policies can substantially increase scientific output without inducing comparable changes in the organizational structure of research itself. Whereas De Philippis (2021) emphasizes how research incentives affect effort deployment and the composition of the research workforce in academic environments, our findings suggest that such responses depend critically on the organizational context in which researchers operate. In settings where research is central to career progression, incentives may reshape both individual behavior and workforce composition. In healthcare organizations, by contrast, research often competes with demanding clinical responsibilities, making access to resources and research capacity a more binding constraint than the marginal return to publication. As a result, research output may grow substantially without reshaping organizational trajectories. More broadly, at least in such contexts, policies aimed at increasing research output may generate more science without necessarily generating different science.

These results have important implications for research policy. Investments in stable funding, infrastructure, and organizational support can substantially increase scientific production, particularly among workers already oriented toward research activities. However, greater research capacity does not automatically generate new collaborations, broader integration across professional groups, or a reconfiguration of an organization's scientific profile. Policies that seek not only to increase the quantity of research but also to transform how research is conducted may therefore require interventions beyond augmenting capacity alone.

Finally, our analysis focuses on a single research-intensive healthcare institution, and the external validity of the results may depend on organizational characteristics and baseline research capacity. Future research should examine whether similar patterns emerge in other hospitals, universities, and research organizations, and should further investigate the conditions under which research-support policies generate heterogenous outcomes.

References

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.

Babina, T., He, A. X., Howell, S. T., Perlman, E. R., & Staudt, J. (2023). Cutting the Innovation Engine: How Federal Funding Shocks Affect University Patenting, Entrepreneurship, and Publications*. *The Quarterly Journal of Economics*, *138*(2), 895–954.

Banal-Estañol, A., Jofre-Bonet, M., Iori, G., Maynou, L., Tumminello, M., & Vassallo, P. (2023). Performance-based research funding: Evidence from the largest natural experiment worldwide. *Research Policy*, *52*(6), 104780.

Barrenho, E., Gautier, E., Miraldo, M., Propper, C., & Rose, C. (2025). Innovation Diffusion Among Coworkers: Evidence from Senior Doctors. *Management Science*.

Barrenho, E., Miraldo, M., Propper, C., & Walsh, B. (2021). The importance of surgeons and their peers in adoption and diffusion of innovation: An observational study of laparoscopic colectomy adoption and diffusion in England. *Social Science & Medicine*, *272*, 113715.

Baruffaldi, S. H., Marino, M., & Visentin, F. (2020). Money to move: The effect on researchers of an international mobility grant. *Research Policy*, *49*(8).

Benavente, J. M., Crespi, G., Figal Garone, L., & Maffioli, A. (2012). The impact of national research funds: A regression discontinuity approach to the Chilean FONDECYT. *Research Policy*, *41*(8), 1461–1475.

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

Bloom, 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.

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.

Chandra, A., & Staiger, D. O. (2007). Productivity Spillovers in Healthcare: Evidence from the Treatment of Heart Attacks. *Journal of Political Economy*, *115*, 103–140.

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.

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

De Philippis, M. (2021). Multi-Task Agents and Incentives: The Case of Teaching and Research for University Professors. *The Economic Journal*, *131*(636), 1643–1681.

Eijkenaar, F., Emmert, M., Scheppach, M., & Schöffski, O. (2013). Effects of pay for performance in health care: A systematic review of systematic reviews. *Health Policy*, *110*(2), 115–130.

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

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

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

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.

Holmstrom, B., & Milgrom, P. (1991). Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *The Journal of Law, Economics, and Organization*, *7*(special_issue), 24–52.

Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, *95*(9-10), 1168–1177.

Lepinteur, A., & Nisticò, R. (2025). 'Based on Admin Data!': How Administrative Data Fosters Young Economists' Career. *IZA Institute of Labor Economics*, (IZA DP No. 17906).

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.

Murphy, K. M., & Topel, R. H. (2006). The Value of Health and Longevity. *Journal of Political Economy*, *114*(5), 871–904.

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-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.

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.

9 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).

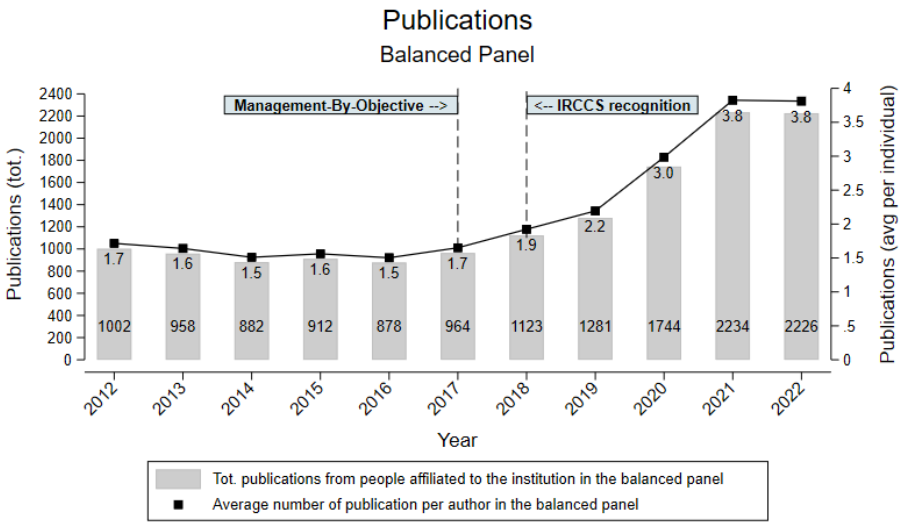


Figure A1: Total number of publications attributed to employees of the institution in the balanced sample (2012–2022).

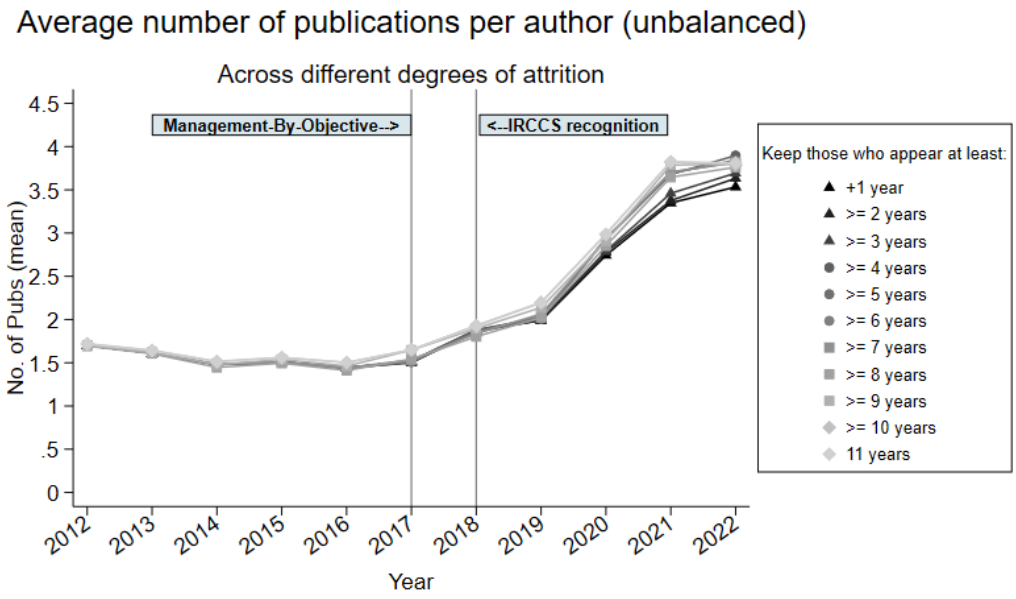


Figure A2: Average publications per researcher per year (2012–2022) across different degrees of panel attrition. Each line represents the mean trend obtained after progressively excluding individuals who are observed for fewer years in the dataset.

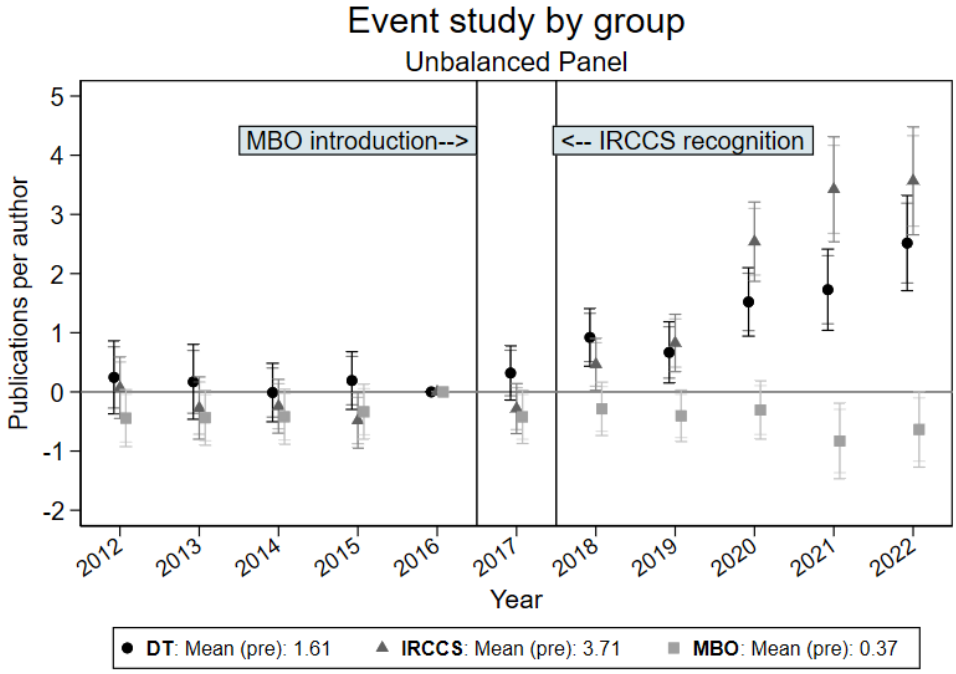
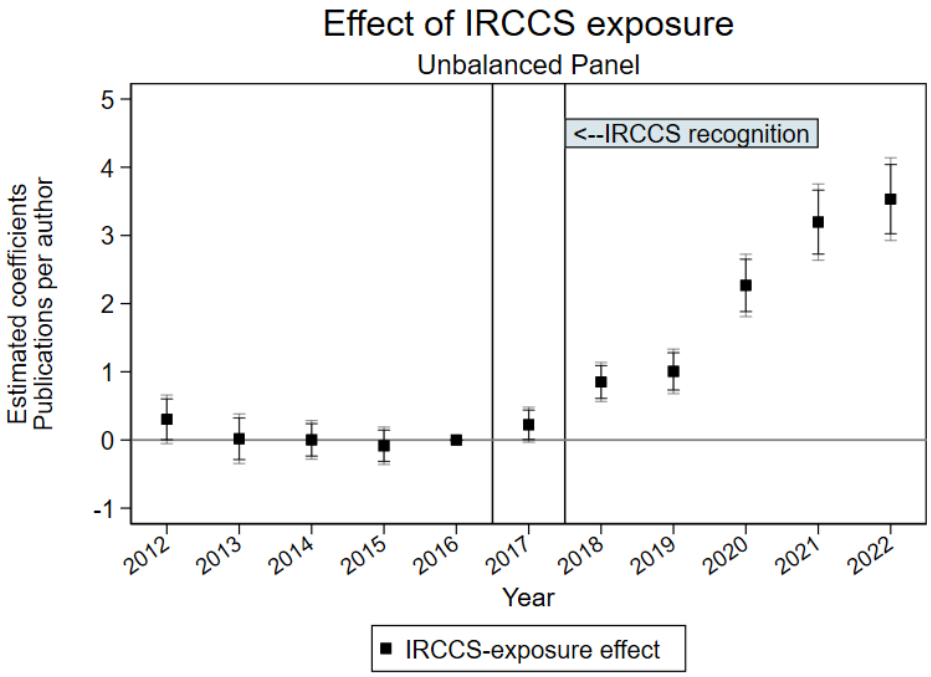


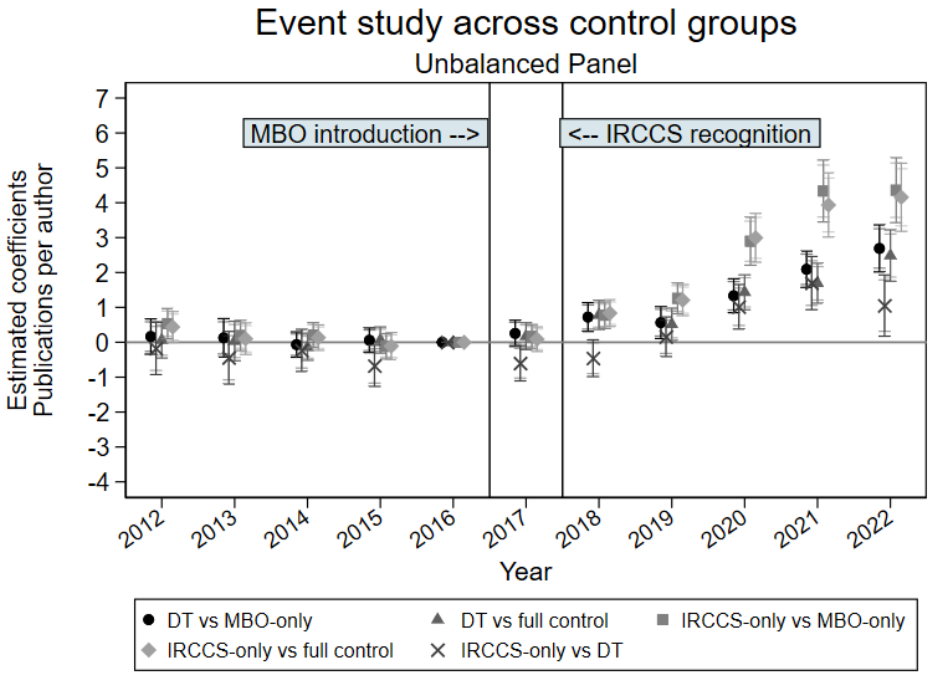
Figure A3: Event-study of the MBO joint with IRCCS recognition policy effect on publications across the different groups, and with respect to the full control group. - Full unbalanced panel.

	(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.

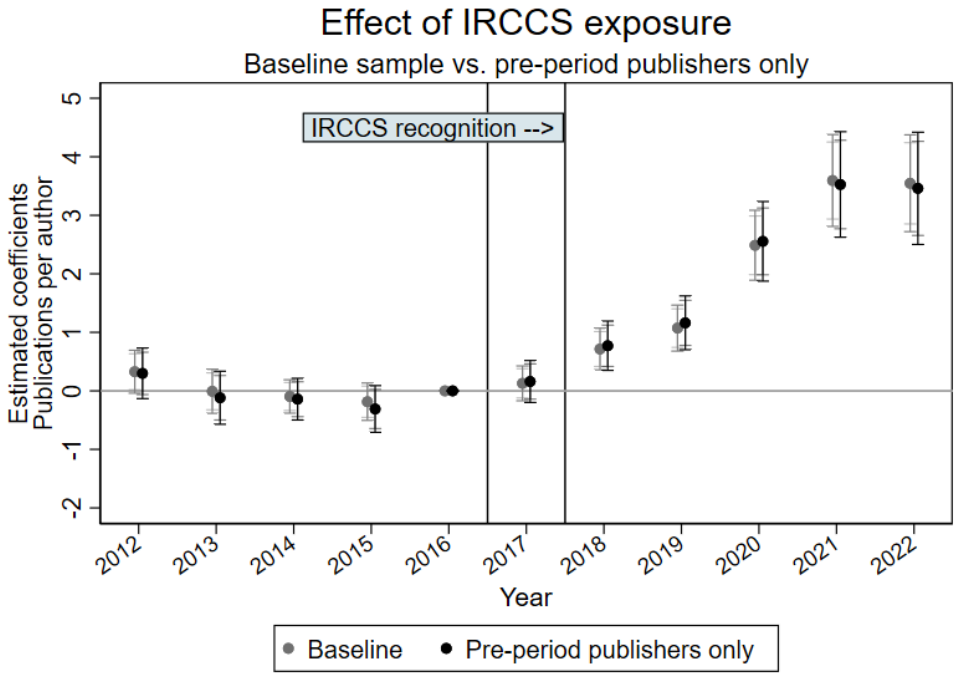


Figure A5: Event-study estimates: comparison between baseline specification and the one excluding never-publishers.

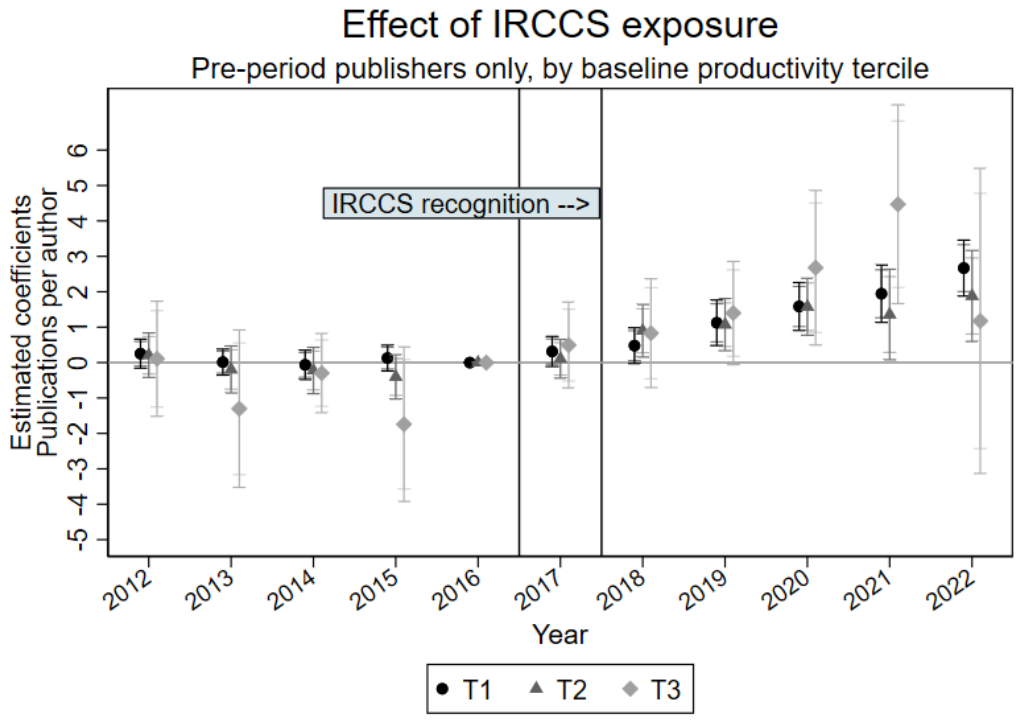
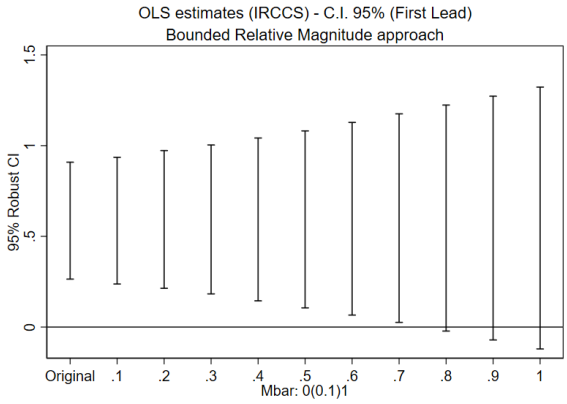
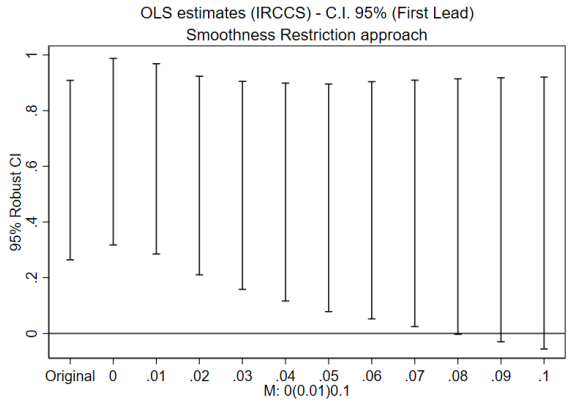


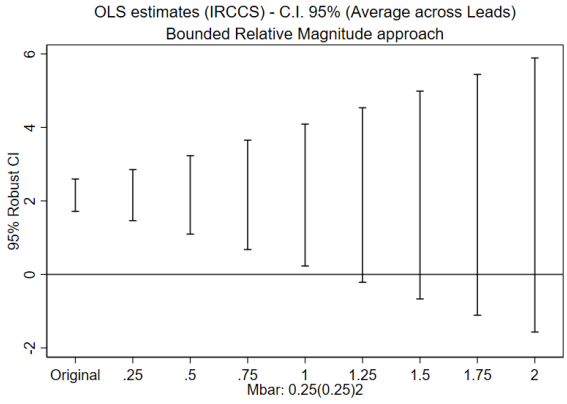
Figure A6: Event-study of the IRCCS effect by baseline productivity quartiles (balanced panel, including those who never published before 2017).



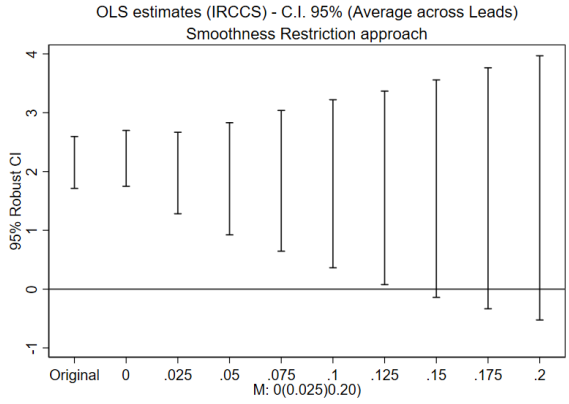
a) BM approach for the significance of the first lead.



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

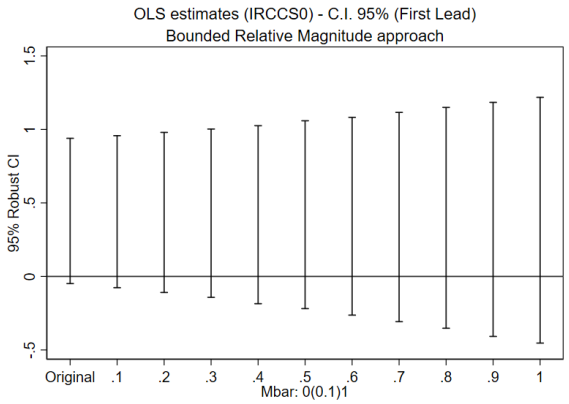


c) BM approach for the significance of the average across all leads.

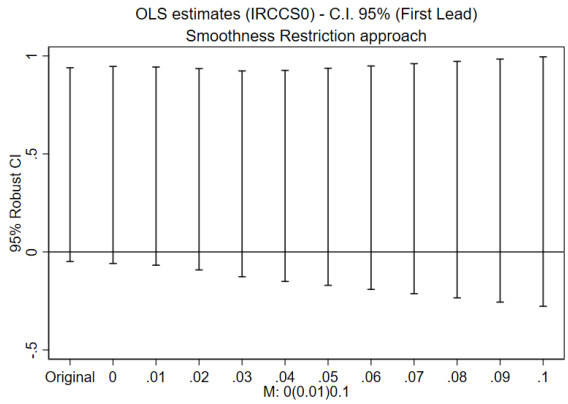


d) SR approach for the significance of the average across all leads.

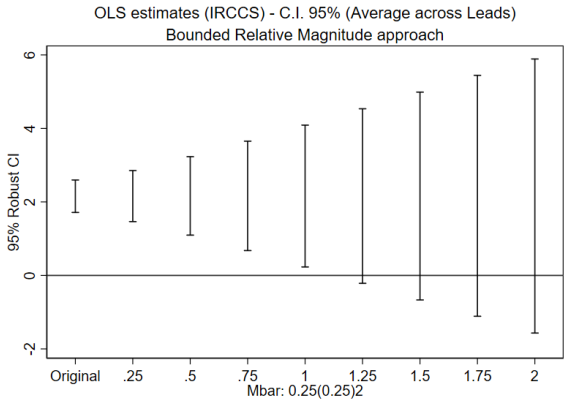
Figure A7: Honest DiD robust confidence sets for overall IRCCS effect estimated with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average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



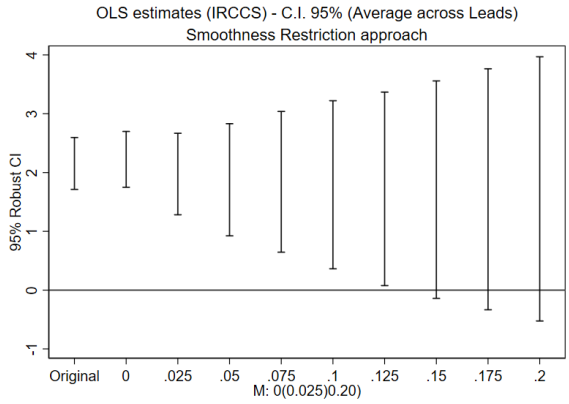
a) BM approach for the significance of the first lead.



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

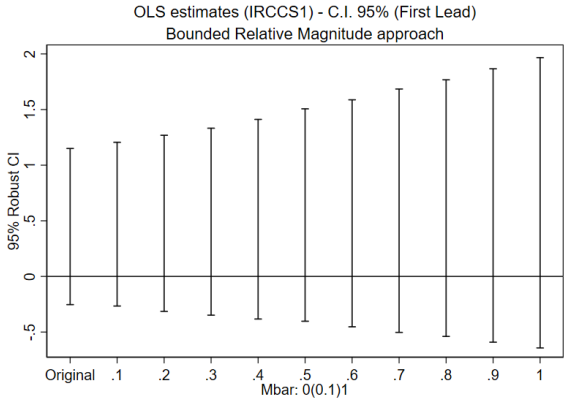


c) BM approach for the significance of the average across all leads.

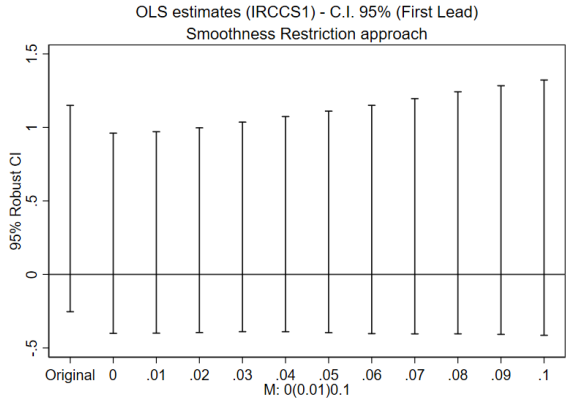


d) SR approach for the significance of the average across all leads.

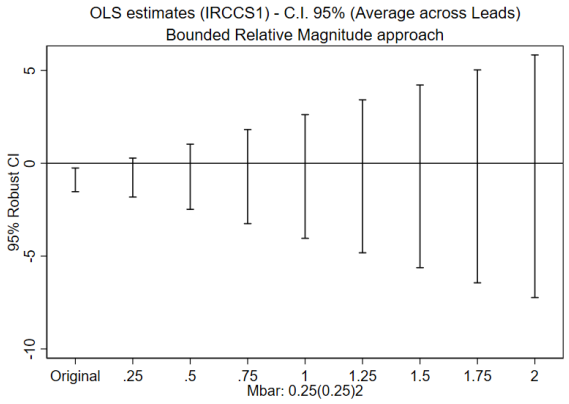
Figure A8: Honest DiD robust confidence sets in the comparison between double-treated and MBO-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of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



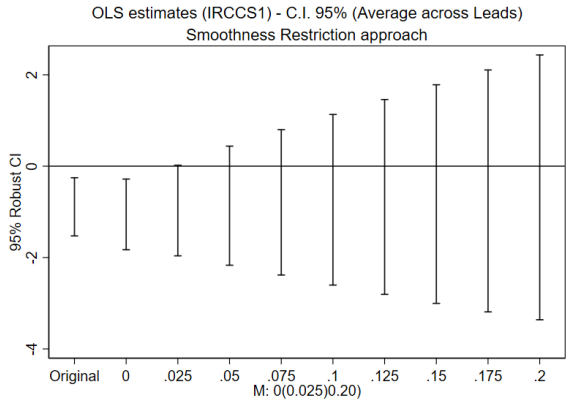
a) BM approach for the significance of the first lead.



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

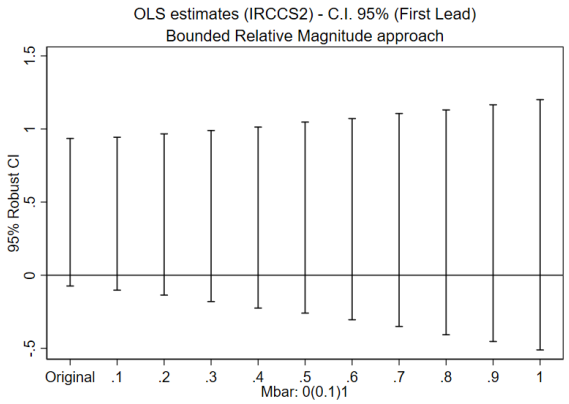


c) BM approach for the significance of the average across all leads.

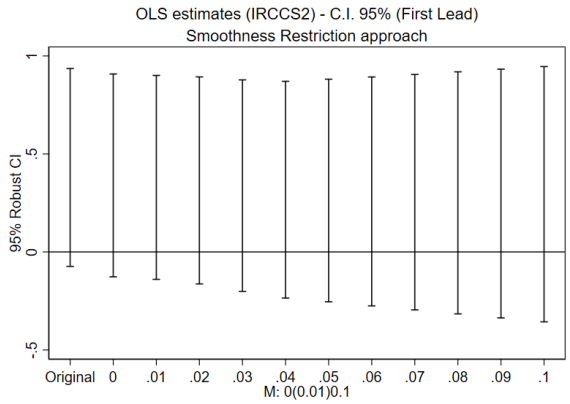


d) SR approach for the significance of the average across all leads.

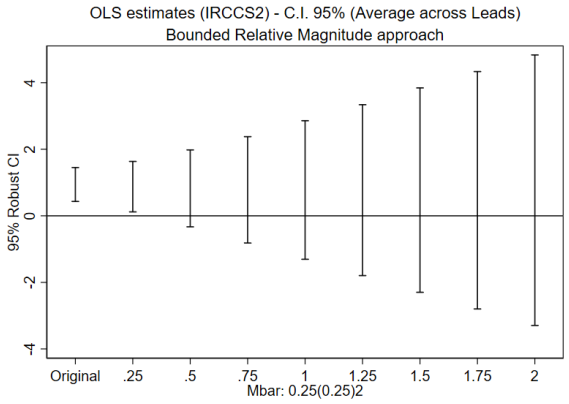
Figure A9: Honest DiD robust confidence sets in the comparison between double-treated and IRCCS-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities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



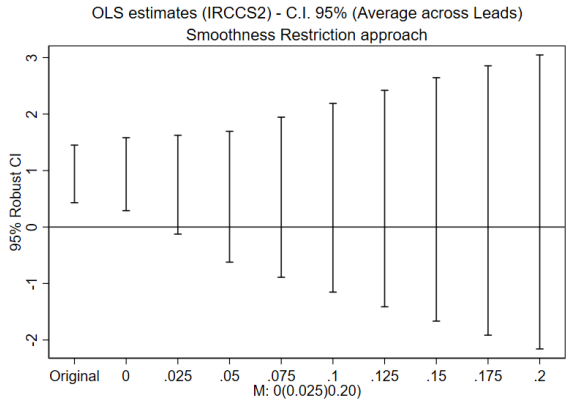
a) BM approach for the significance of the first lead.



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

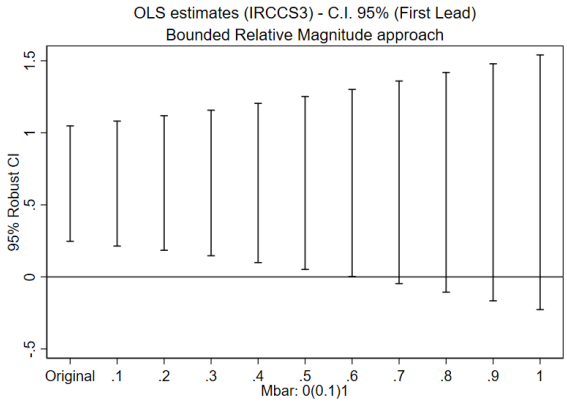


c) BM approach for the significance of the average across all leads.

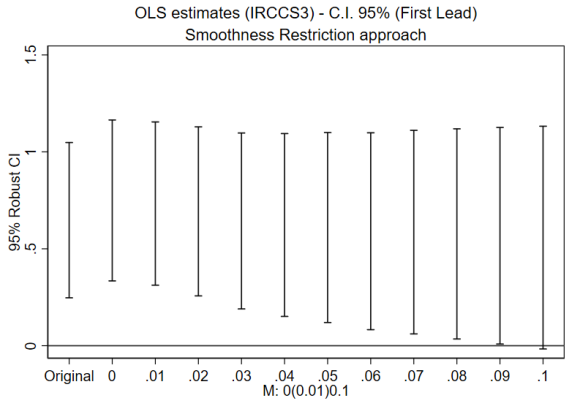


d) SR approach for the significance of the average across all leads.

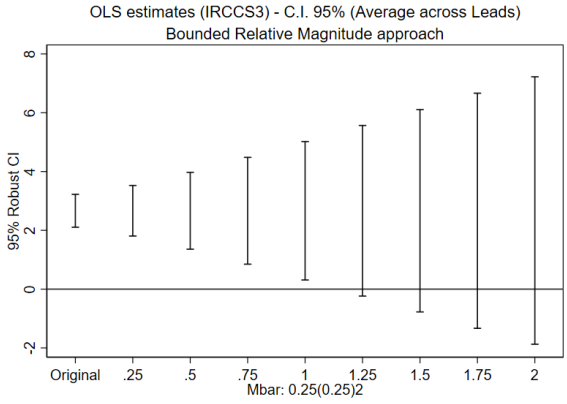
Figure A10: Honest DiD robust confidence sets in the comparison between double-treated and pure control units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities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



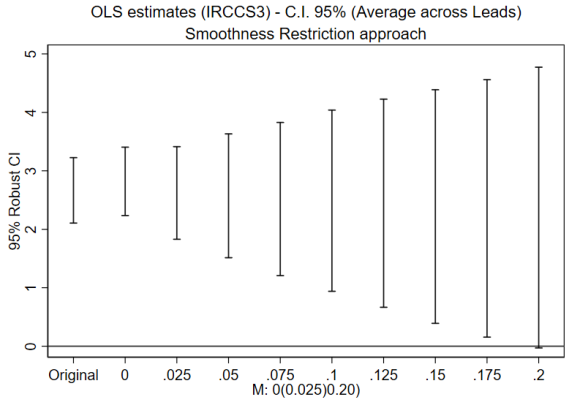
a) BM approach for the significance of the first lead.



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

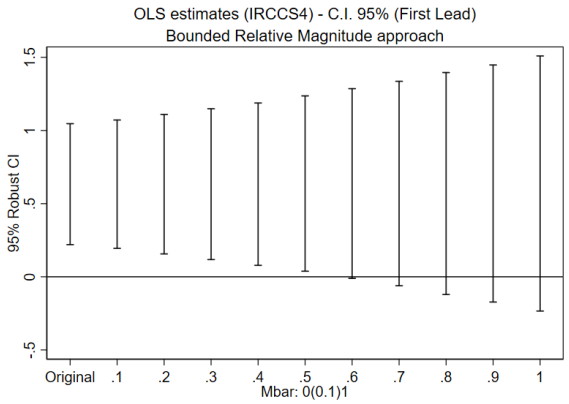


c) BM approach for the significance of the average across all leads.

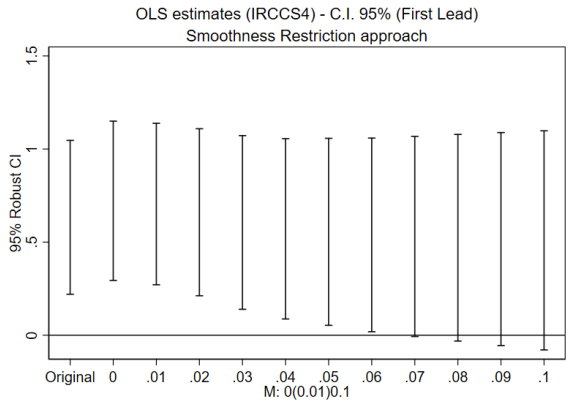


d) SR approach for the significance of the average across all leads.

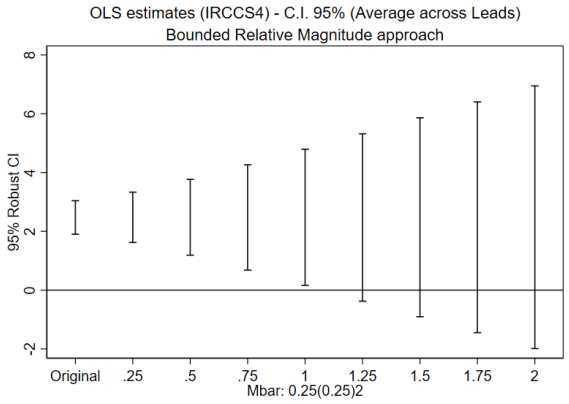
Figure A11: Honest DiD robust confidence sets in the comparison between IRCCS-only and MBO-only units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



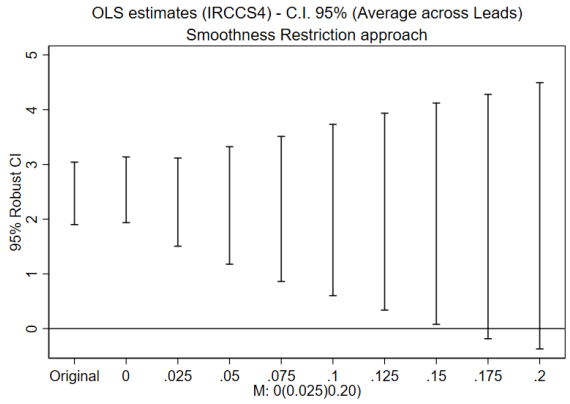
a) BM approach for the significance of the first lead.



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

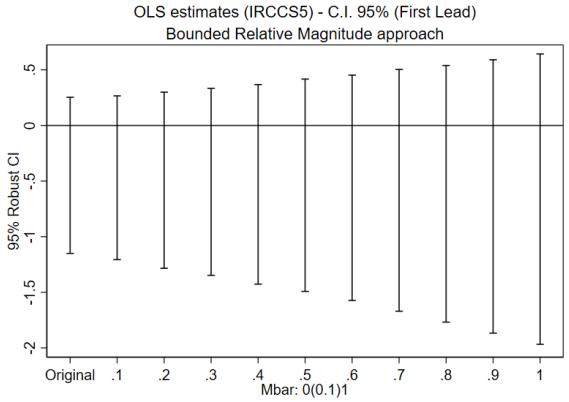


c) BM approach for the significance of the average across all leads.

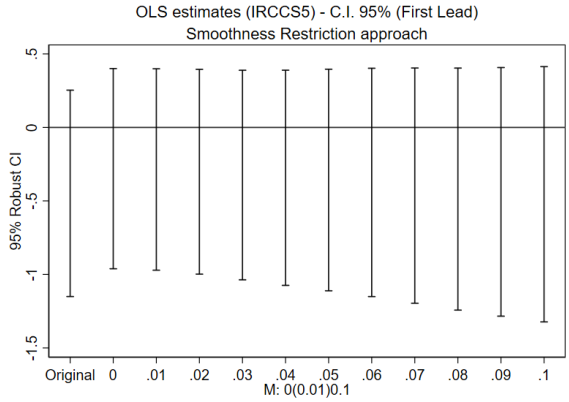


d) SR approach for the significance of the average across all leads.

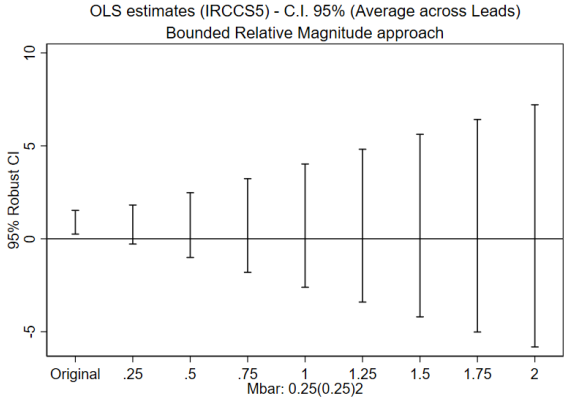
Figure A12: Honest DiD robust confidence sets in the comparison between IRCCS-only and pure control units, estimated the or overall IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



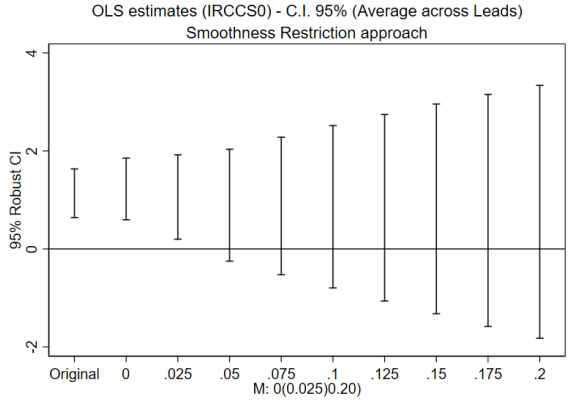
a) BM approach for the significance of the first lead.



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 A13: 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)).

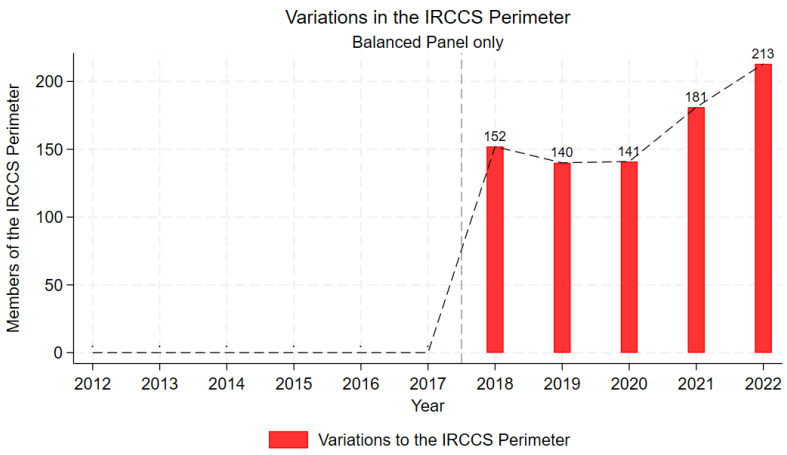


Figure A14: Evolution of the units "lately selected" into the IRCCS perimeter.

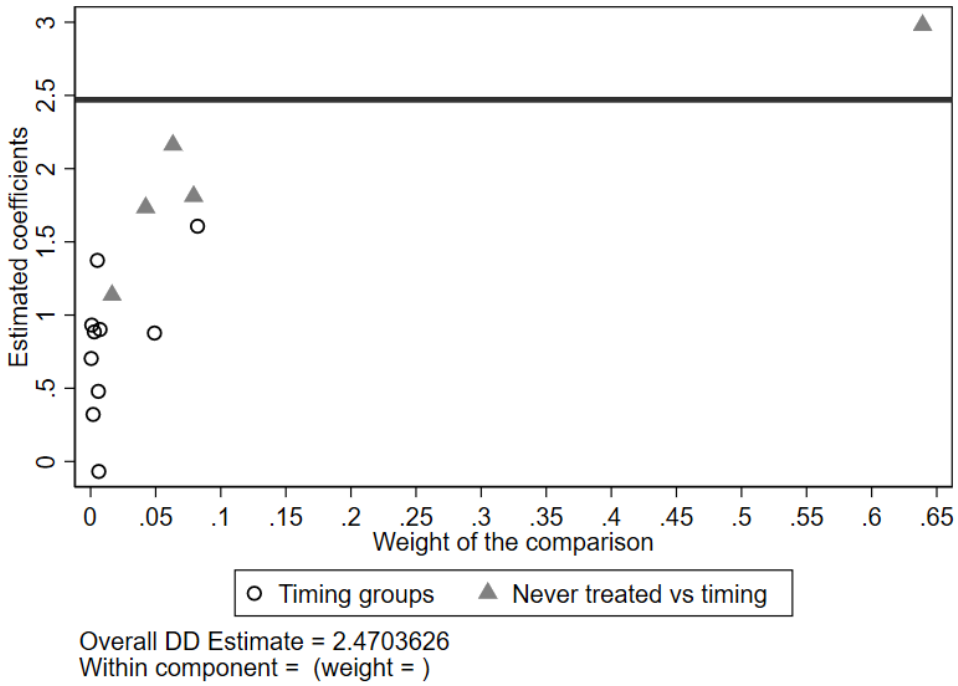


Figure A15: Bacon Decomposition of TWFE Estimate

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

Figure A16: 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.

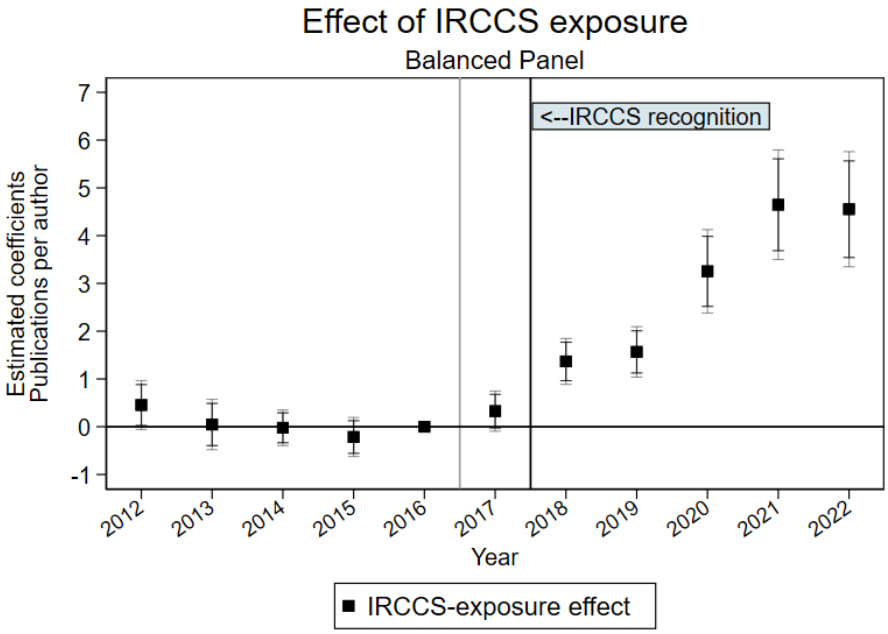


Figure A17: Event-study of IRCCS recognition effect on publications. Units included in the IRCCS perimeter in any year after 2018 are excluded from the estimates.

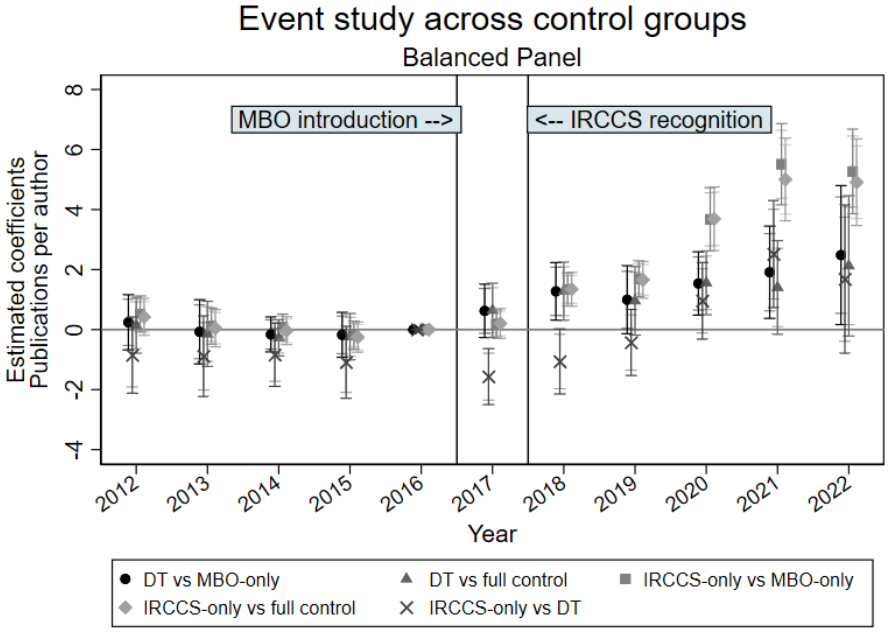


Figure A18: Event-study of IRCCS policy effect on publications for comparisons across different groups. Units included in the IRCCS perimeter in any year after 2018 are excluded from the estimates.

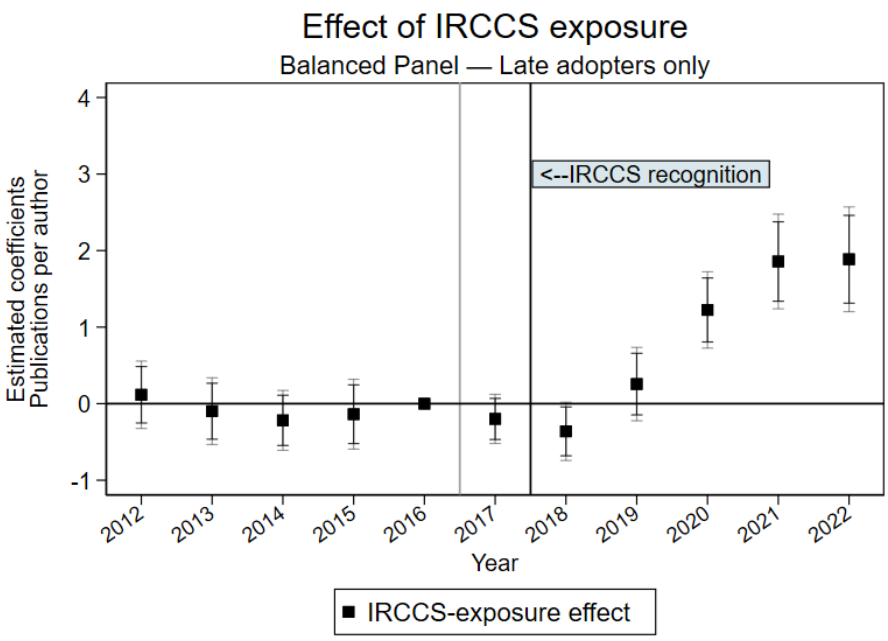


Figure A19: Event-study of IRCCS recognition effect on publications of late-enrolled. Units included in the IRCCS perimeter in 2018 are excluded from the estimates.

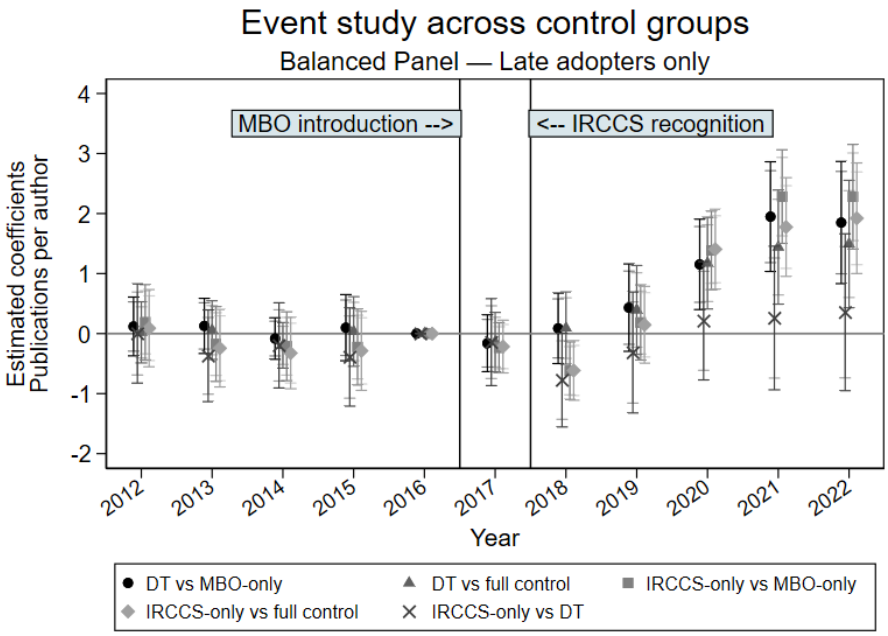
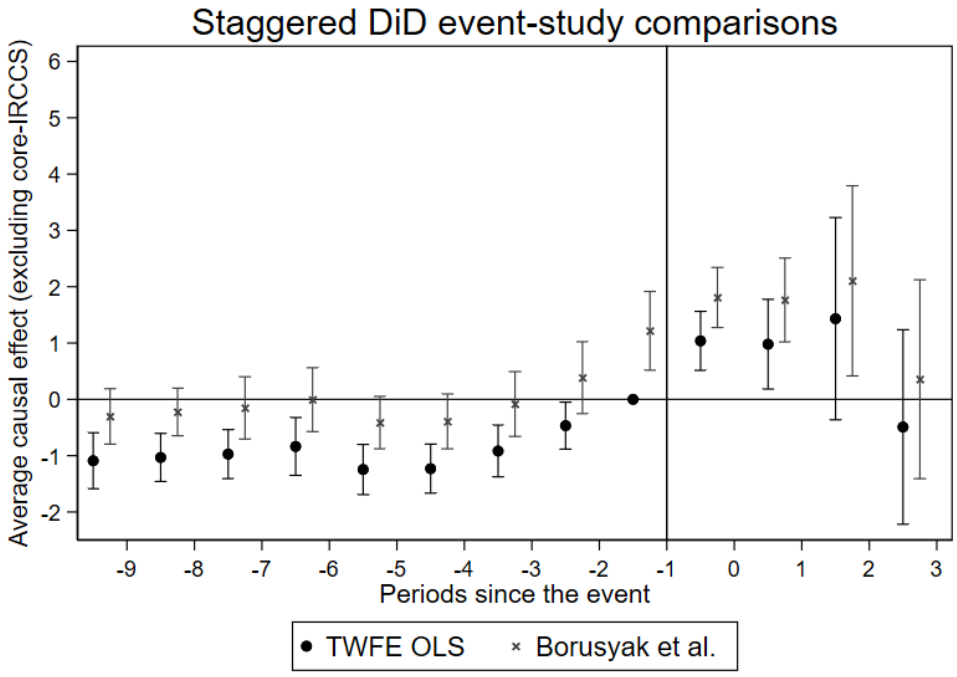
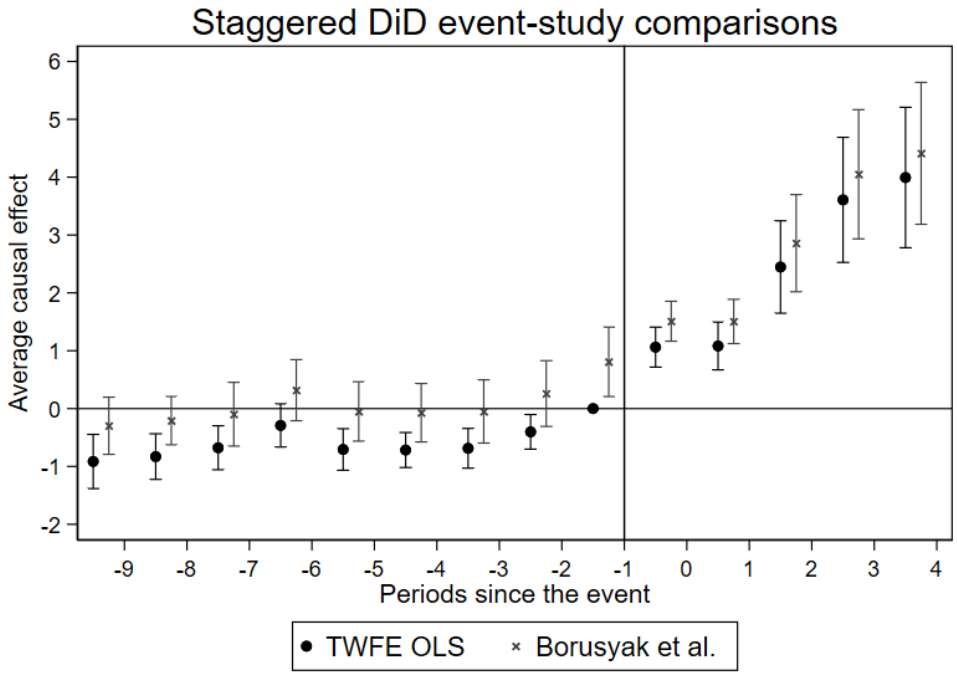


Figure A20: Event-study of IRCCS policy effect on publications late-enrolled for comparisons across different groups. Units included in 2018 are excluded from the estimates.

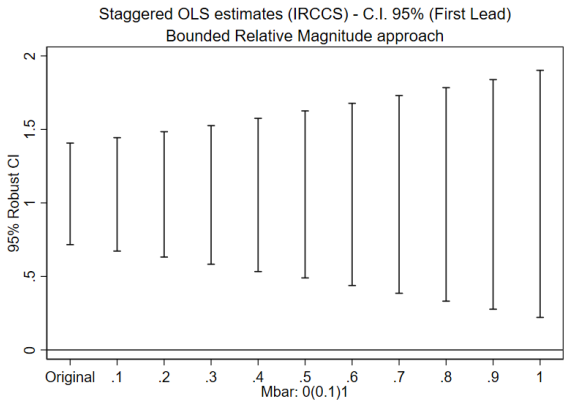


a) Event-study of the staggered inclusion in the IRCCS perimeter of late-enrolled only, excluding the "core" researchers treated in 2018.

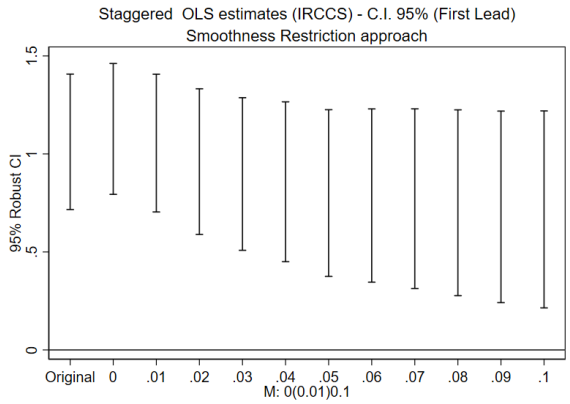


b) Event-study of the staggered inclusion in the IRCCS perimeter of all treated units, including the "core" researchers treated in 2018.

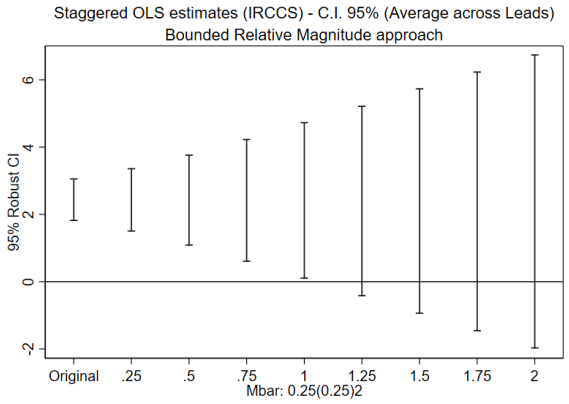
Figure A21: Event-study of IRCCS policy effect on publications for comparisons across different groups.



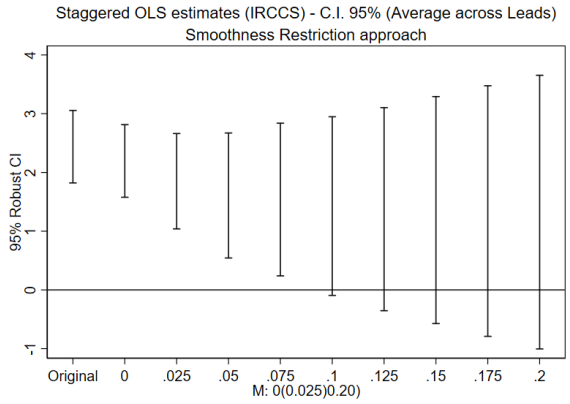
a) BM approach for the significance of the first lead.



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

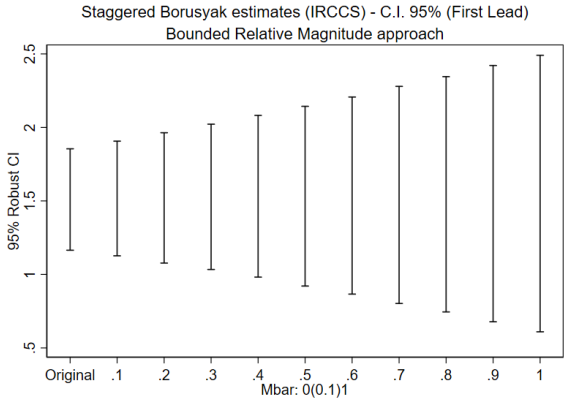


c) BM approach for the significance of the average across all leads.

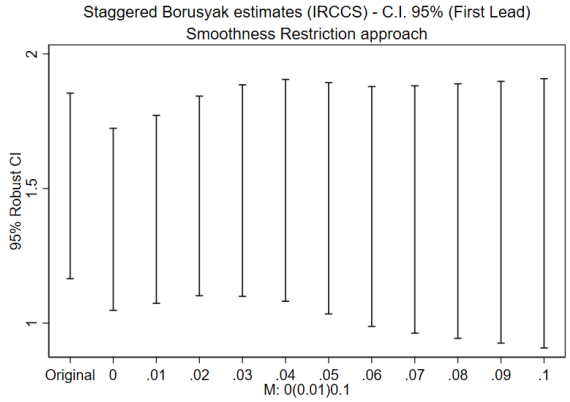


d) SR approach for the significance of the average across all leads.

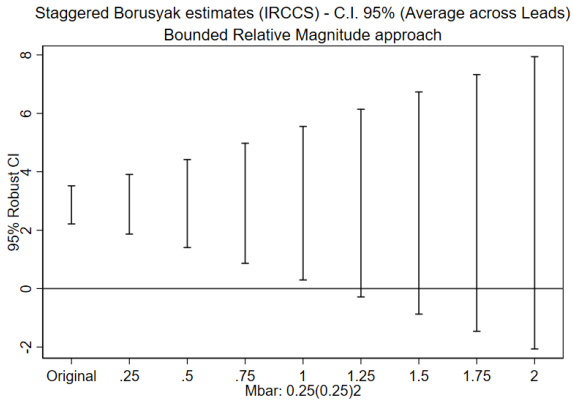
Figure A22: Honest DiD robust confidence sets in the OLS staggered event-study estimates on the IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different quantities of interest (coefficient on the first lead of the event study, a) and b). Average across all post-treatment leads, c) and d)).



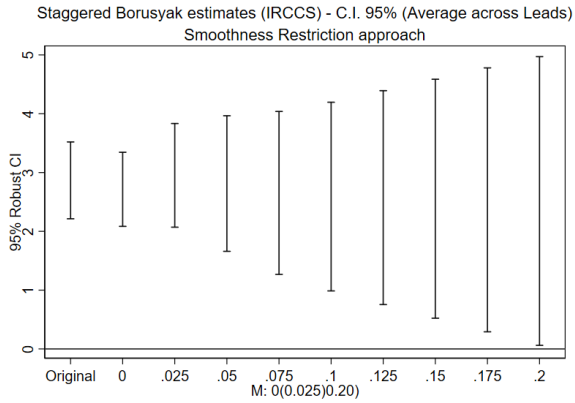
a) BM approach for the significance of the first lead.



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



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

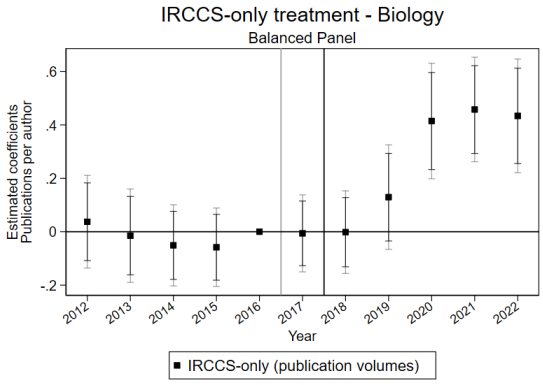


d) SR approach for the significance of the average across all leads.

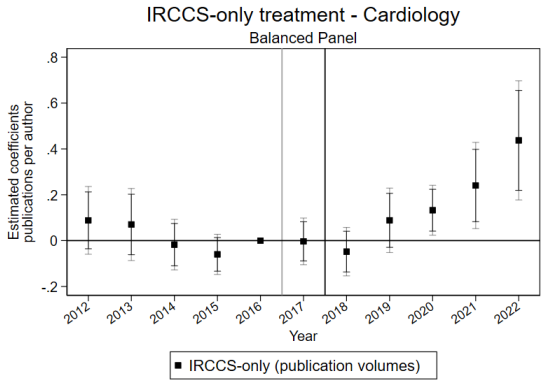
Figure A23: Honest DiD robust confidence sets in the staggered event-study estimates obtained via the procedure by Borusyak et al., 2024 on the IRCCS effect by adopting with different methodologies (Bounded Relative Magnitude - BM: a) and c). Smoothness Restriction - SR: b) and d)) and for different 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
Cardiology/Cardiovascular Diseases	Cardiology
Chemistry/Physics	Chemistry; Electrochemistry; Crystallography; Materials Science; Physics; Polymer Science; Spectroscopy; Acoustics; Optics
Diagnostics/Imaging	Radiology; Pathology; Medical Laboratory Technology
Infectious Diseases/Immuno-Allergology	Infectious Diseases; Virology; Parasitology; Tropical Medicine; Mycology; Immunology; Allergy
Internal Medicine (Generic)	General and Internal Medicine; Emergency Medicine; Integrative and Complementary Medicine; Research and Experimental Medicine
Internal Medicine (Specialties)	Gastroenterology; Endocrinology; Respiratory System; Rheumatology; Geriatrics; Dermatology; Nutrition and Dietetics
Mother/Child	Obstetrics; Pediatrics; Reproductive Biology
Neuroscience/Psychiatry/Psychology	Neurosciences; Psychiatry; Psychology
Oncology/Hematology	Oncology; Hematology
Pharmacology	Pharmacology; Toxicology; Substance Abuse
Public Health/Health Systems	Health Care Sciences and Services; Public, Environmental and Occupational Health; Business and Economics; Public Administration; Biomedical Social Sciences; Environmental Sciences
Rehabilitation/Care	Rehabilitation; Nursing; Sport Sciences; Audiology
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
Surgery/Procedural	Surgery; Anesthesiology; Orthopedics; Otorhinolaryngology; Urology; Transplantation; Ophthalmology; Dentistry
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

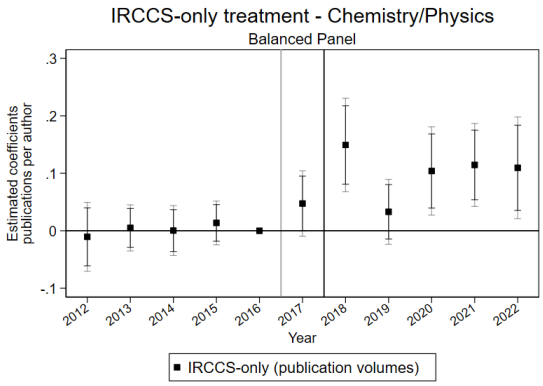
Table A6: Classification of research areas into macro-categories based on Web of Science (WoS) subject categories.



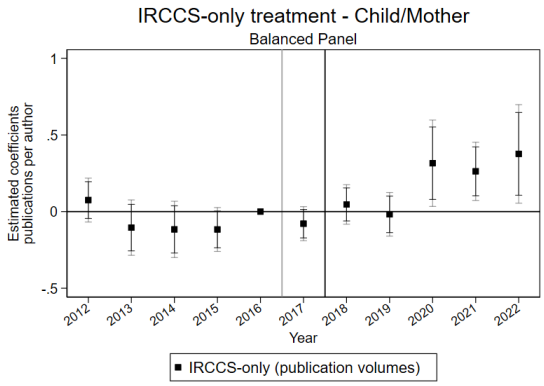
(a) Biology



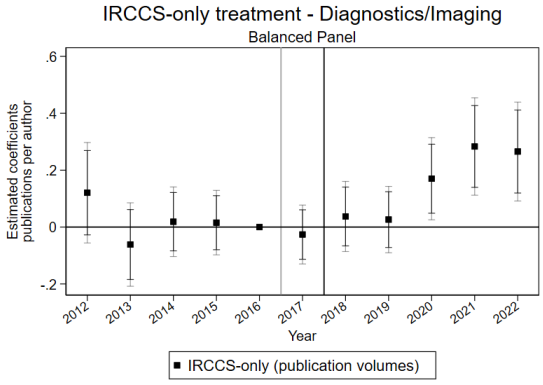
(b) Cardiology and Cardiovascular Systems



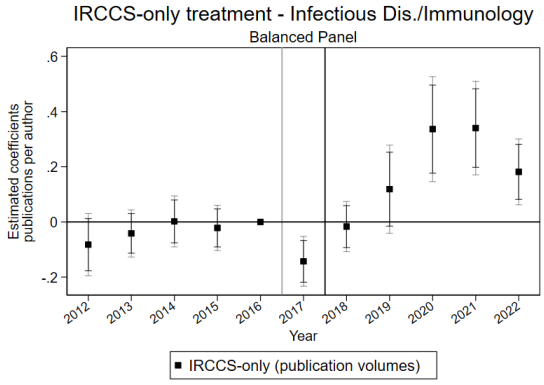
(c) Chemistry and Physics



(d) Child and Maternal Health

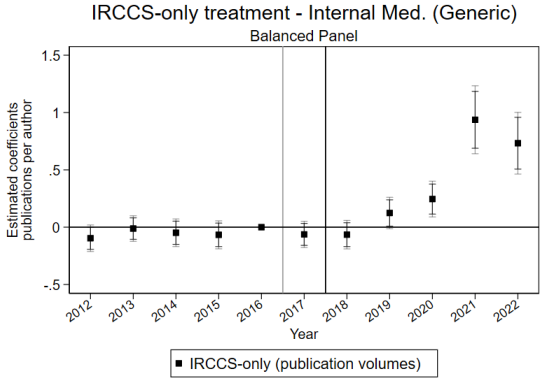


(e) Diagnostics and Imaging

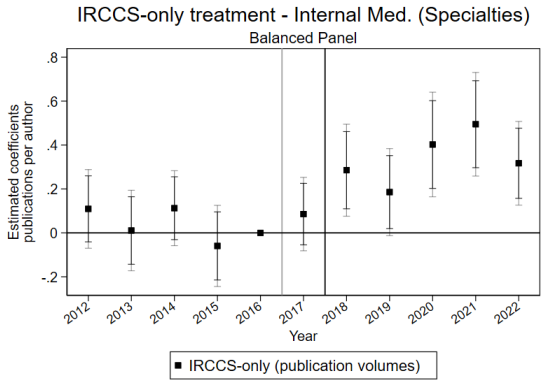


(f) Infectious Diseases, Immunology, Allergology

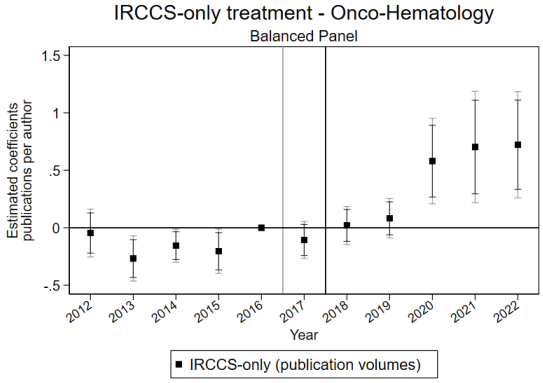
Figure A24: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the number individual publications in a specific area.



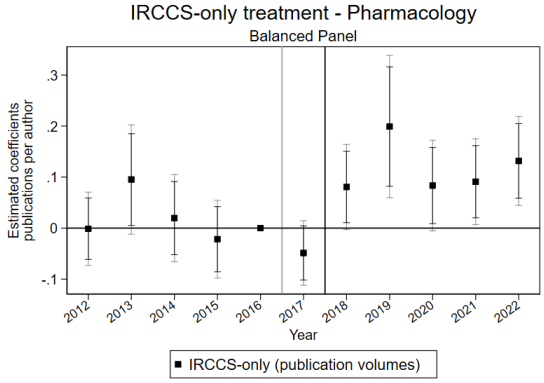
(a) Internal Medicine (General and Emergency Medicine)



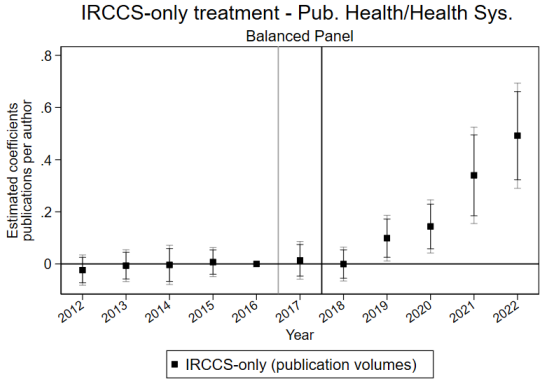
(b) Internal Medicine (Specialties)



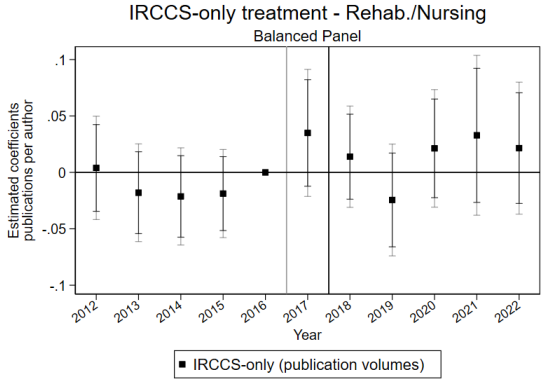
(c) Oncology and Hematology



(d) Pharmacology and Toxicology

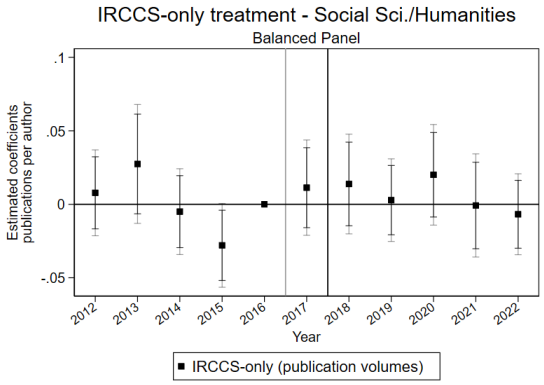


(e) Public Health and Health Systems

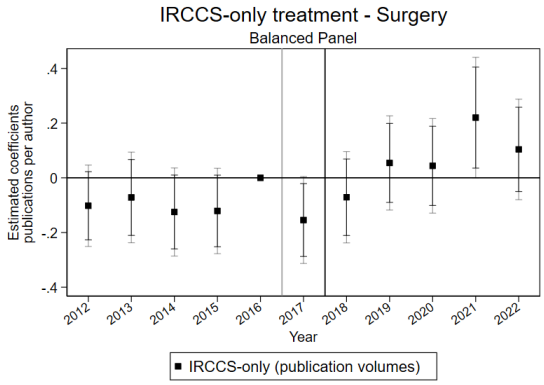


(f) Rehabilitation and Nursing

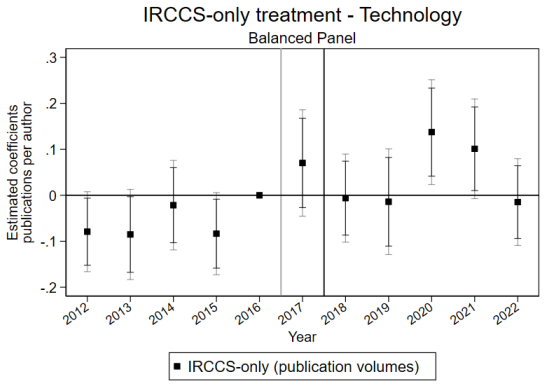
Figure A25: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the number individual publications in a specific area.



(a) Social Sciences and Humanities



(b) Surgery, Dentistry, Procedures

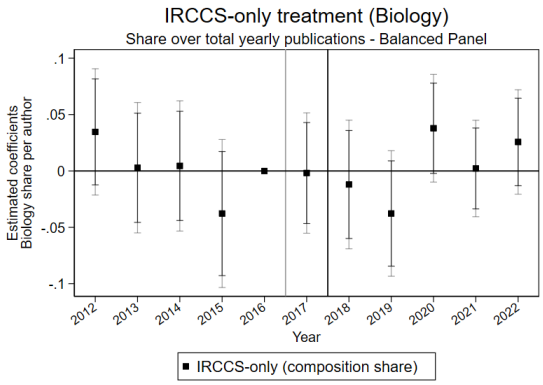


(c) Technology, Engineering, Computer Science

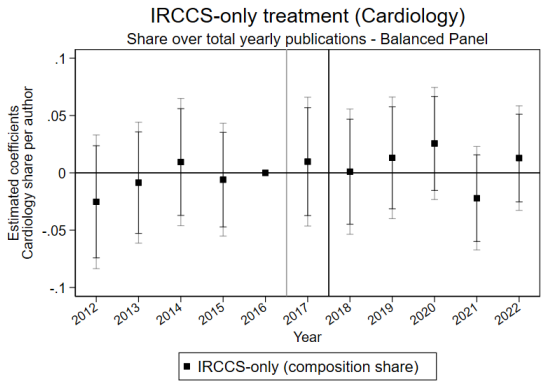
Figure A26: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the number individual publications in a specific area.

Specialization	ITT			Pre (joint)		Mean Pre	Comment
	Coef.	p-value	RW p-val.	F-stat	p-value		
Biology	-0.003	0.802	0.998	0.560	0.692	0.047	Null
	(0.013)						
Cardiology/Cardiovasc. Diseases	0.002	0.860	0.998	0.596	0.665	0.041	Null
	(0.013)						
Chemistry and Physics	0.001	0.910	0.998	0.592	0.668	0.006	Null
	(0.006)						
Diagnostics and Imaging	0.003	0.748	0.998	1.948	0.101	0.041	Null
	(0.011)						
Infectious Dis./ Immunology	-0.000	0.981	0.998	1.691	0.150	0.025	Null
	(0.011)						
Internal Medicine (Generic)	0.009	0.578	0.991	0.514	0.725	0.046	Null
	(0.016)						
Internal Medicine (Specialties)	0.012	0.438	0.976	2.750	0.028**	0.083	Ambiguous
	(0.016)						
Maternal and Child Health	0.008	0.470	0.976	0.408	0.803	0.039	Null
	(0.011)						
Neurosciences and Psych.	-0.018	0.280	0.884	3.295	0.011**	0.069	Ambiguous
	(0.017)						
Oncology and Hematology	0.012	0.414	0.973	0.579	0.678	0.061	Null
	(0.014)						
Pharmacology	-0.004	0.721	0.998	1.007	0.403	0.021	Null
	(0.011)						
Public Health and Health Systems	0.005	0.429	0.976	2.004	0.092*	0.011	Ambiguous
	(0.006)						
Rehabilitation and Nursing	0.003	0.594	0.991	0.355	0.840	0.011	Null
	(0.006)						
Social Sciences and Humanities	0.006	0.350	0.948	1.641	0.162	0.006	Null
	(0.006)						
Surgery and Procedures	-0.035*	0.052	0.109	0.271	0.896	0.073	Null
	(0.018)						
Technology and Engineering	-0.016	0.142	0.477	2.255	0.062*	0.020	Ambiguous
	(0.011)						

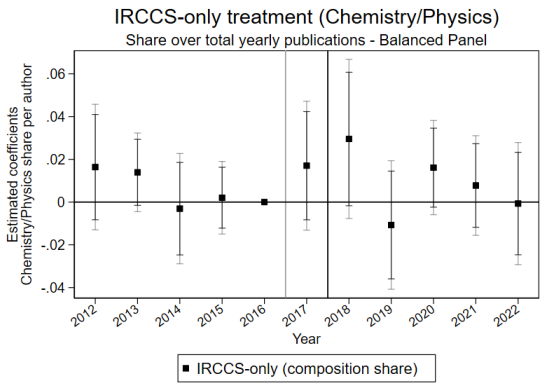
Table A7: Each row reports results from a separate regression of the yearly share of publications in a given specialization on a post-treatment indicator for IRCCS-only researchers, controlling for unit and year fixed effects. Standard errors, reported in parentheses beneath the ITT coefficient, are 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.



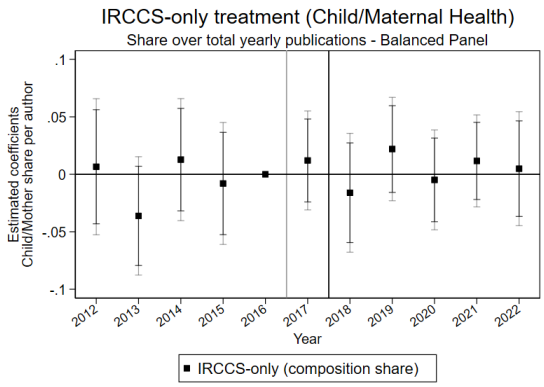
(a) Biology



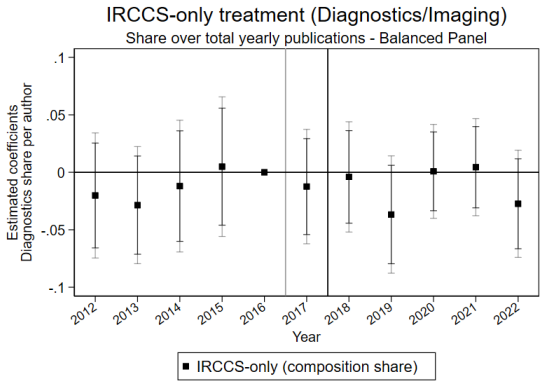
(b) Cardiology and Cardiovascular Systems



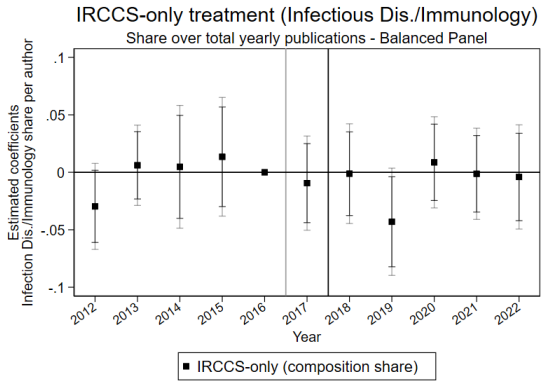
(c) Chemistry and Physics



(d) Child and Maternal Health

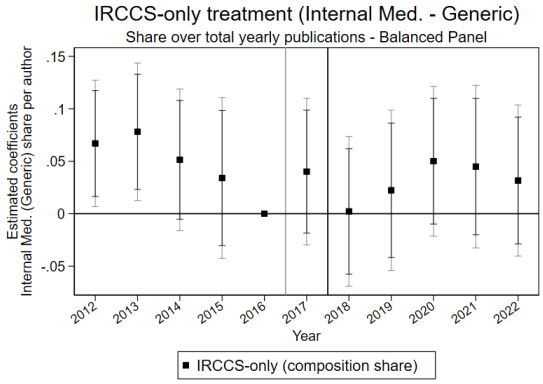


(e) Diagnostics and Imaging

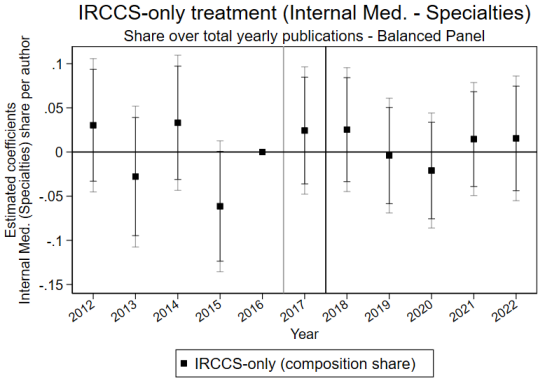


(f) Infectious Diseases, Immunology, Allergology

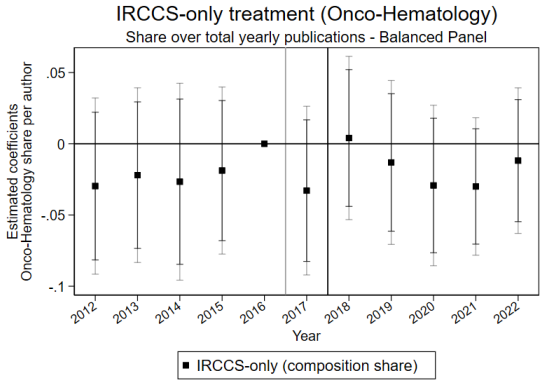
Figure A27: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the share of individual publications in a specific area over the total number of one's annual publications.



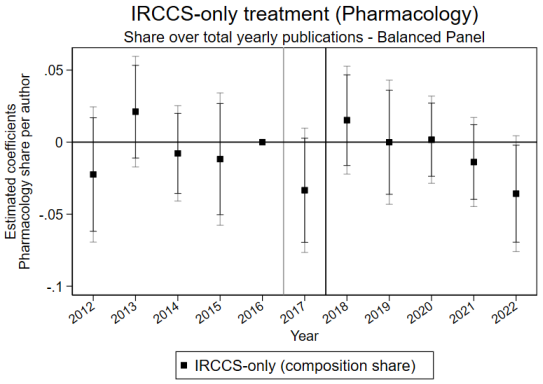
(a) Internal Medicine (General and Emergency Medicine)



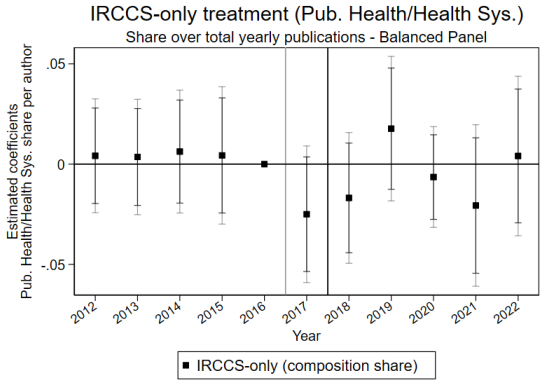
(b) Internal Medicine (Specialties)



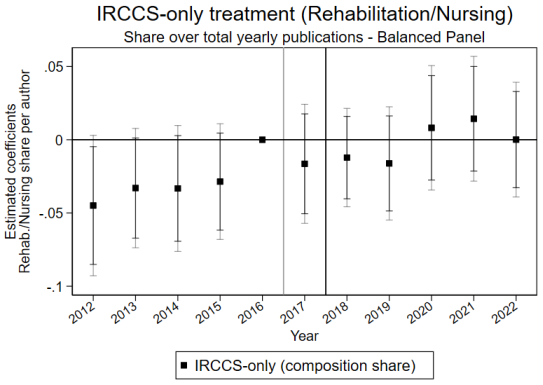
(c) Oncology and Hematology



(d) Pharmacology and Toxicology

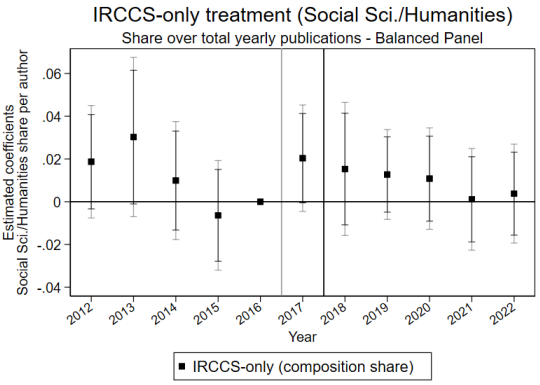


(e) Public Health and Health Systems

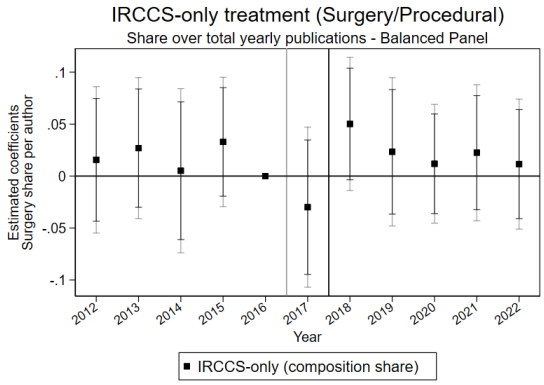


(f) Rehabilitation and Nursing

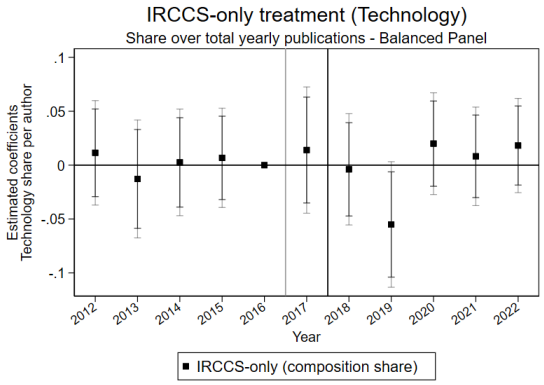
Figure A28: The figure reports event-study coefficients from regressions at the author-year level, where the dependent variable is the share of individual publications in a specific area over the total number of one's annual publications.



(a) Social Sciences and Humanities



(b) Surgery, Dentistry, Procedures



(c) Technology, Engineering, Computer Science

Figure A29: 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.

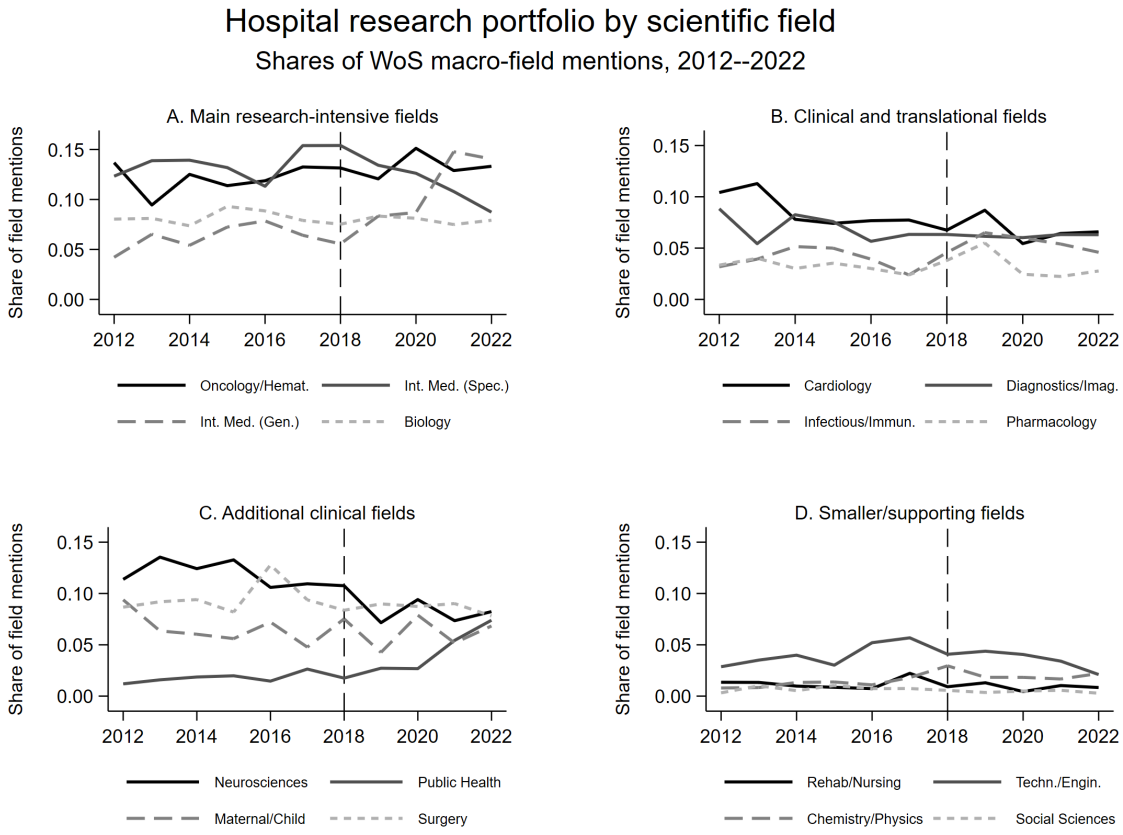


Figure A30: Descriptive evolution of the hospital's research portfolio across scientific fields, 2012–2022. The figure reports the annual distribution of Web of Science (WoS) macro-field mentions aggregated at the hospital level. Because WoS subject categories are not mutually exclusive, a publication may contribute to multiple scientific fields; the reported shares therefore describe the relative composition of field classifications rather than mutually exclusive publication shares. The vertical dashed line denotes the 2018 IRCCS recognition. The composition of the hospital's research portfolio remains broadly stable throughout the period, with the most visible fluctuations occurring within closely related medical classifications—particularly between generic and specialty-related Internal Medicine categories—rather than across distinct scientific domains.

Table A8: Sample composition after excluding physicians with no pre-treatment publications

	Control	MBO	IRCCS	DT
Publishers pre-2017	101	78	168	60
Never-publishers pre-2017	57	99	14	7
% publishers	63.9	44.1	92.3	89.6

Table A9: Baseline Characteristics of Early and Late IRCCS Entrants

	Core 2018	Late entrants	Difference	p-value
Mean publications (2012–2016)	3.978	1.640	2.338***	0.000
Total publications (2012–2016)	19.890	8.202	11.688***	0.000
No pre-2017 publications	0.071	0.106	-0.035	0.332
MBO physician	0.213	0.362	-0.149*	0.010

Table A10: Predictors of Late Entry into the IRCCS Research Perimeter

	Dependent variable: Late entrant		
	(1)	(2)	(3)
Mean publications per year	-0.0356***	-0.0359***	-0.0329***
	(0.00706)	(0.00736)	(0.00723)
No publications before 2017		-0.0134	-0.0111
		(0.118)	(0.117)
Eligible for MBO scheme			0.109
			(0.0709)
Constant	0.488***	0.490***	0.451***
	(0.0399)	(0.0429)	(0.0491)
Observations	249	249	249
R-squared	0.083	0.083	0.093

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(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 A11: Impact of IRCCS recognition on Annual Publications, by changing the outcome to log-transformation and inverse hyperbolic sine (Difference-in-Differences).

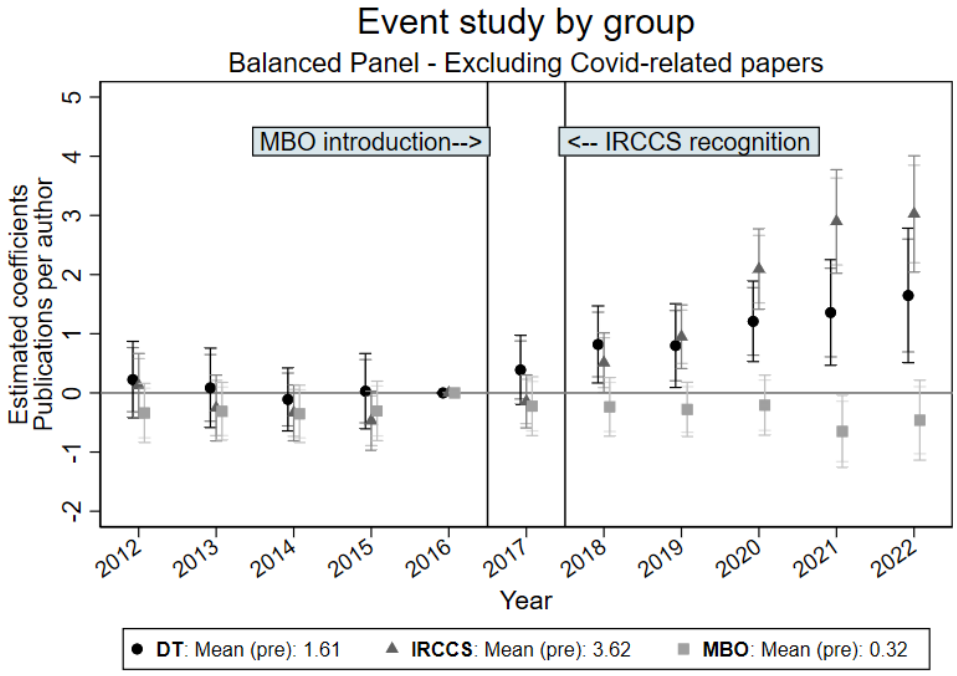


Figure A31: Event-study of the IRCCS after dropping COVID-related publications (balanced panel).

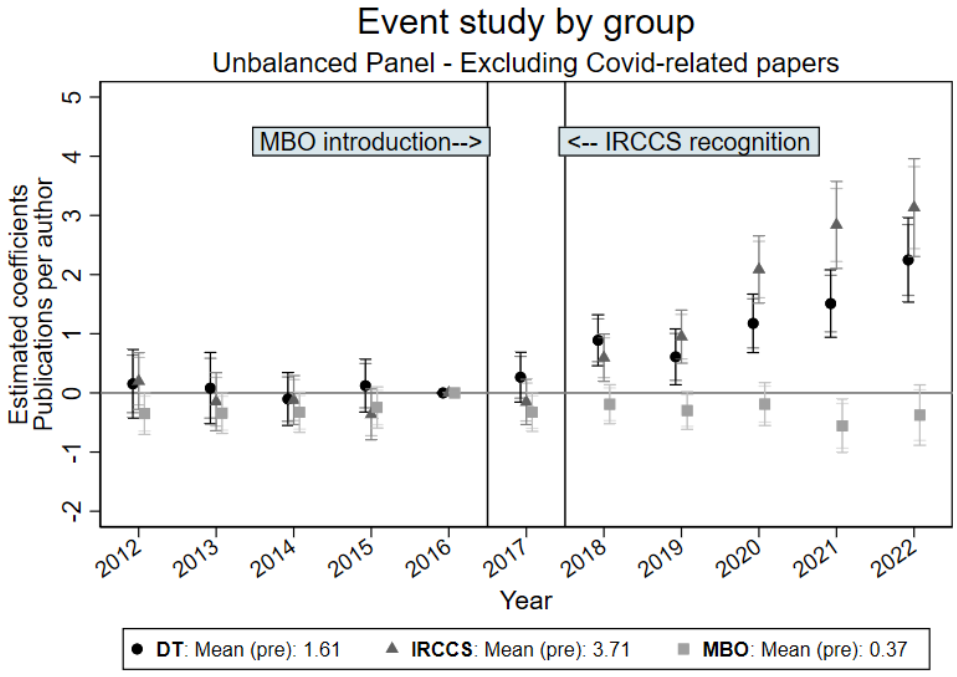
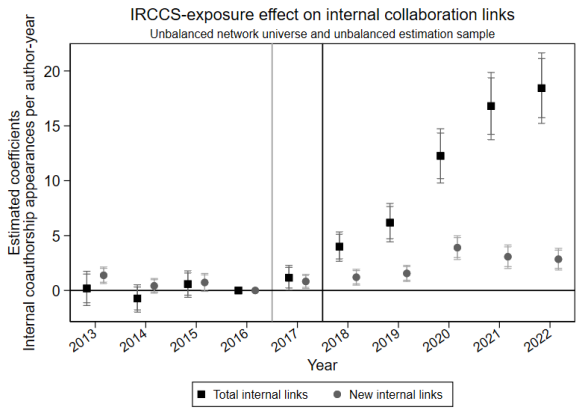
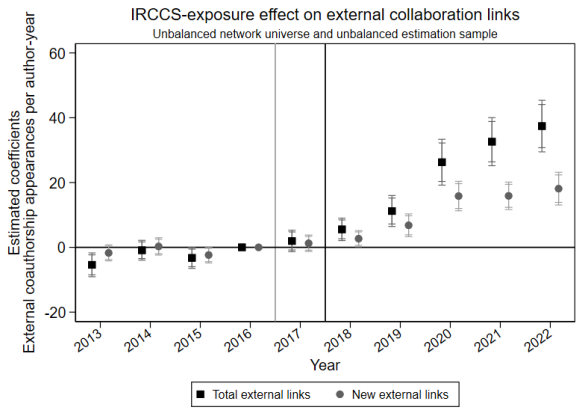


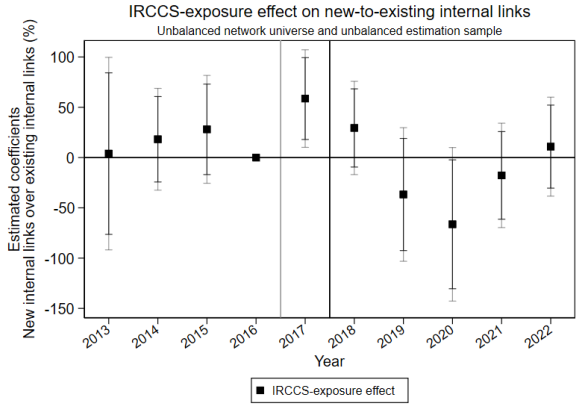
Figure A32: Event-study of the IRCCS after dropping COVID-related publications (unbalanced panel).



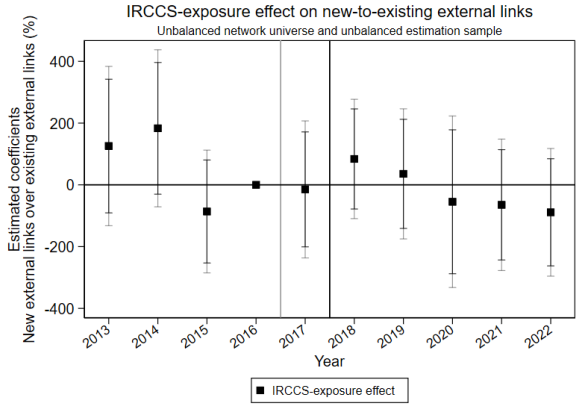
a) Internal collaboration links.



b) External collaboration links.

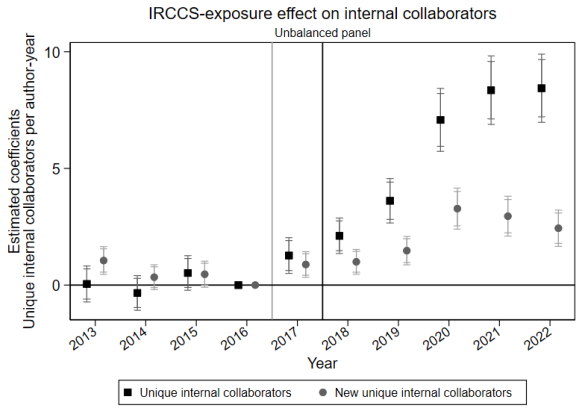


c) New-to-existing internal links.

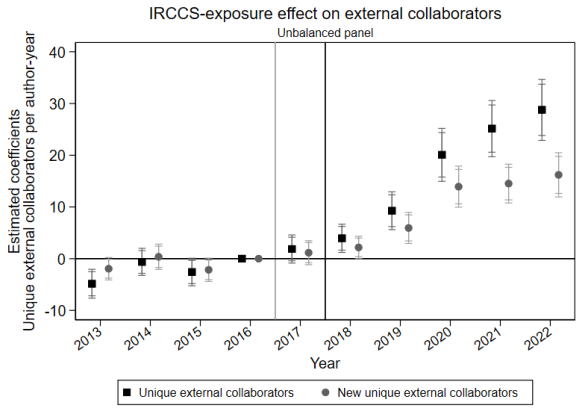


d) New-to-existing external links.

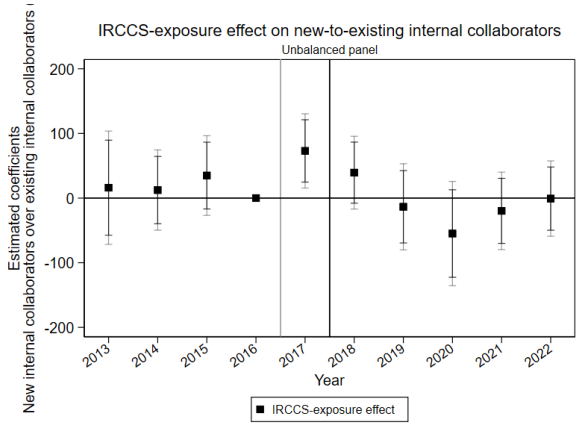
Figure A33: Event-study estimates of the effect of IRCCS recognition on collaboration links, unbalanced panel.



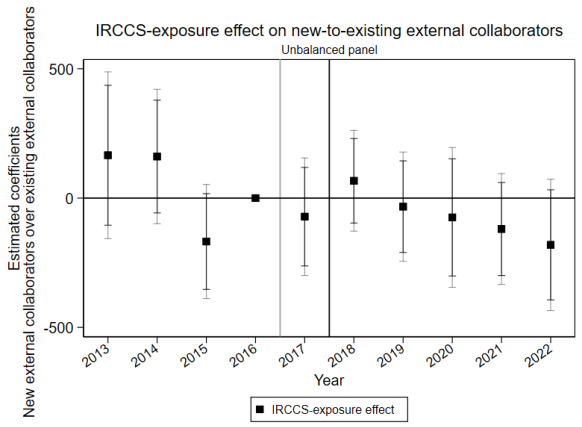
a) Unique internal collaborators.



b) Unique external collaborators.



c) New-to-existing unique internal collaborators.



d) New-to-existing unique external collaborators.

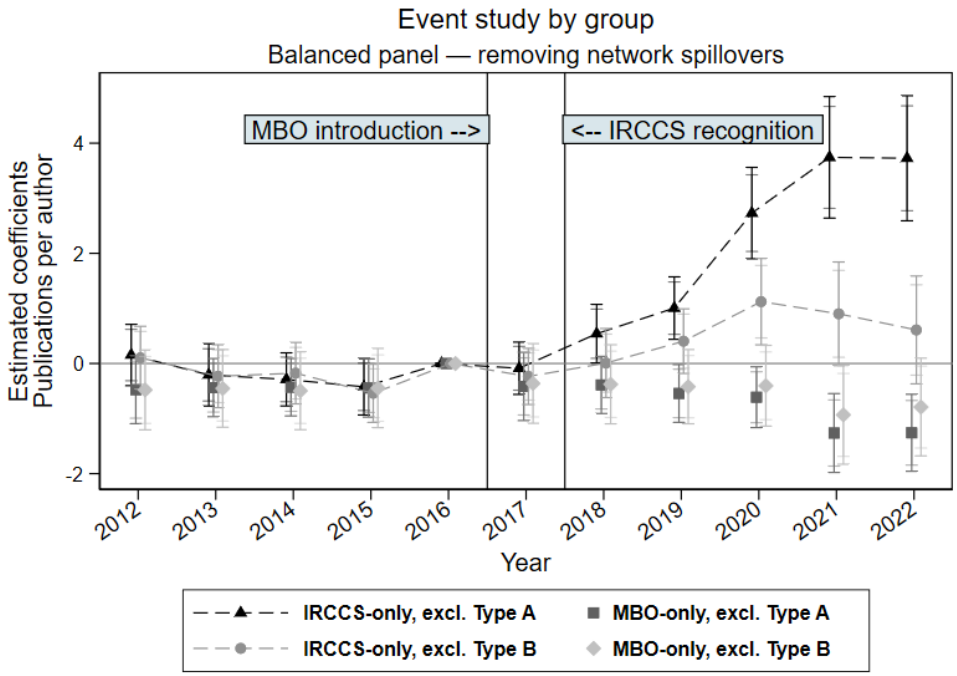
Figure A34: Event-study estimates of the effect of IRCCS recognition on unique collaborators, unbalanced panel.

2012-2022; SEs clustered at individual level

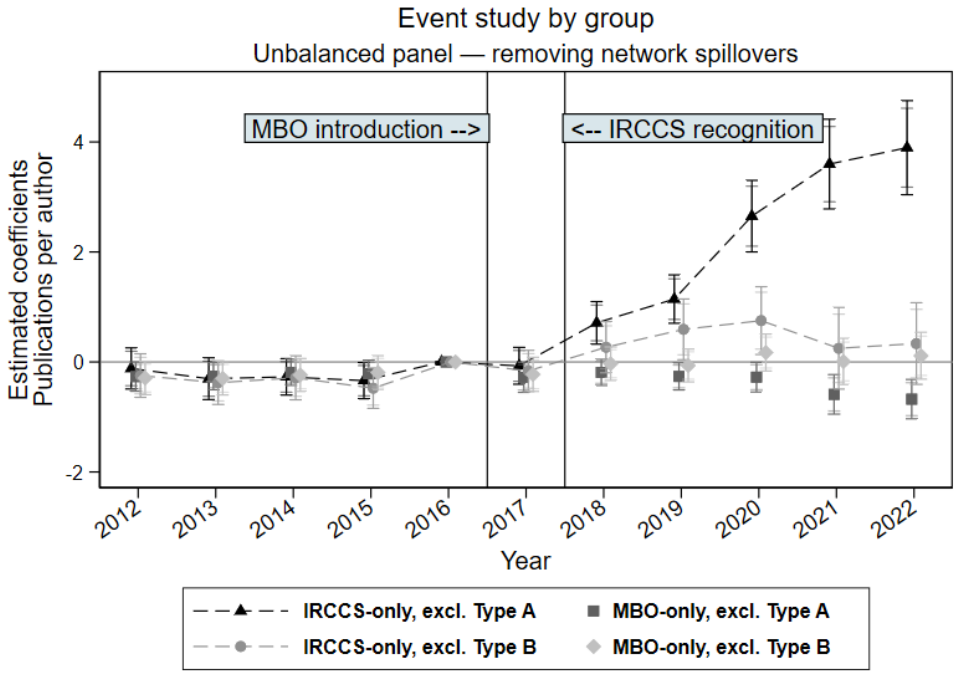
	(1) Pubs	(2) Pubs	(3) Pubs	(4) Pubs	(5) Pubs	(6) Pubs	(7) Pubs	(8) Pubs
Post 2017*MBO-only	-0.30943 ***	-0.18027			-0.13506 **	0.19353 ***		
	[0.10852]	[0.13296]			[0.06280]	[0.07404]		
Post 2017*MBO-only	2.11063 ***	0.49499 ***			2.19877 ***	0.59396 ***		
	[0.29400]	[0.17523]			[0.23569]	[0.15897]		
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 A12: Impact of IRCCS recognition on Annual Publications, by changing the outcome to log-transformation and inverse hyperbolic sine (Difference-in-Differences).

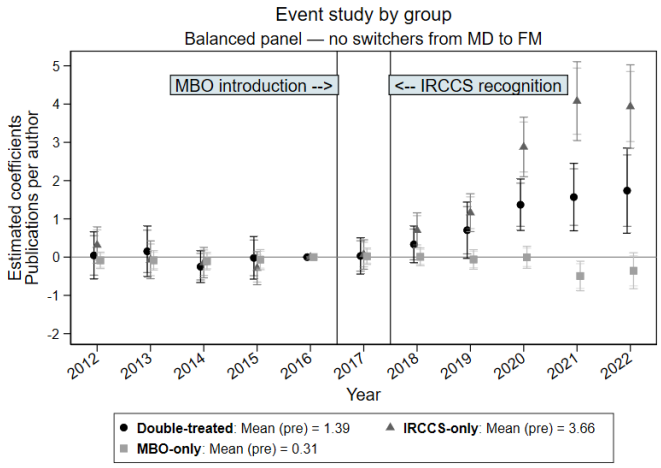


a) Balanced panel - Event-study of the two policies' effect on publications of the IRCCS-treated and MBO-treated only groups once excluded double-treated units and *Type A* researchers (MBO-treated units who *ever* co-authored with IRCCS researchers) in the first case, and double-treated units and *Type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

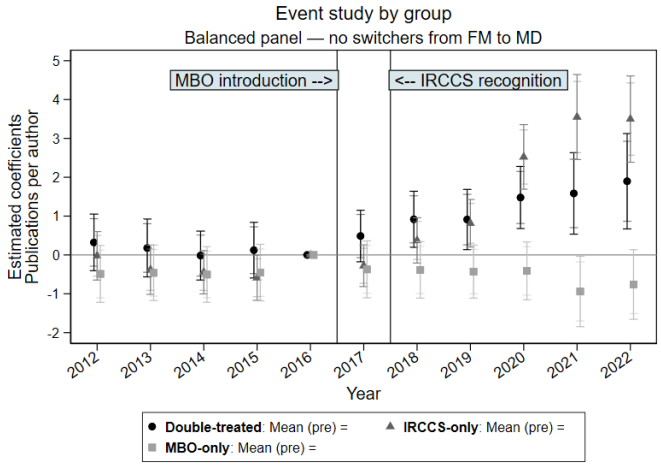


a) Unbalanced panel - Event-study of the two policies' effect on publications of the IRCCS-treated and MBO-treated only groups once excluded double-treated units and *Type A* researchers (MBO-treated units who *ever* co-authored with IRCCS researchers) in the first case, and double-treated units and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

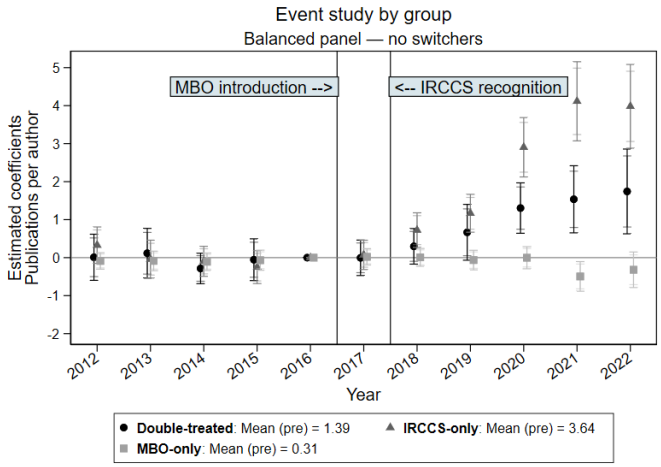
Figure A35: Event-study of for the dynamic effects of the MBO policy and the IRCCS recognition effect on publications in different samples subset based on collaboration dynamics. Double-treated units are always excluded. a) is the Balanced panel, b) the unbalanced one.



a) Removing switchers from Medical Direction to Faculty Membership.

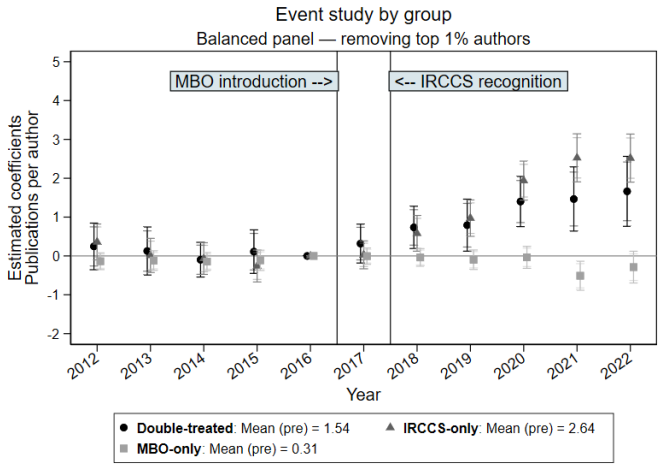


b) Removing switchers from Faculty Membership to Medical Direction.

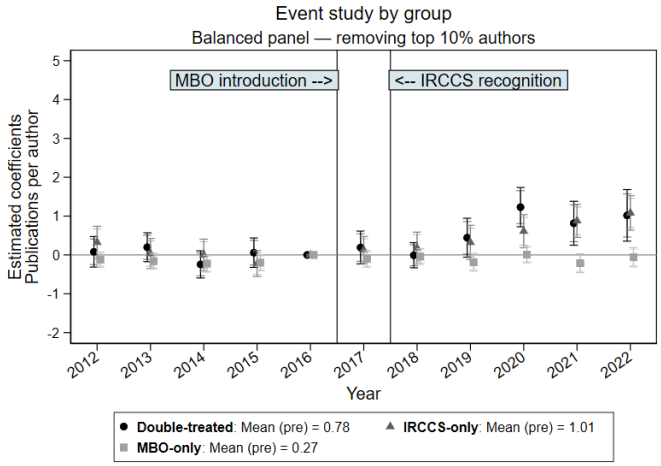


c) Removing all switcher.

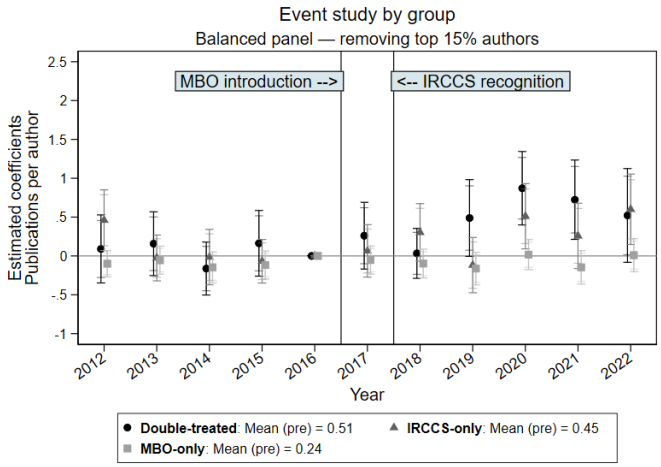
Figure A36: Dynamic mpact of MBO-implementation and IRCCS recognition on Annual Publications in the sub-samples obtained by excluding those who switch across groups.



a) Removing top 1% researchers.

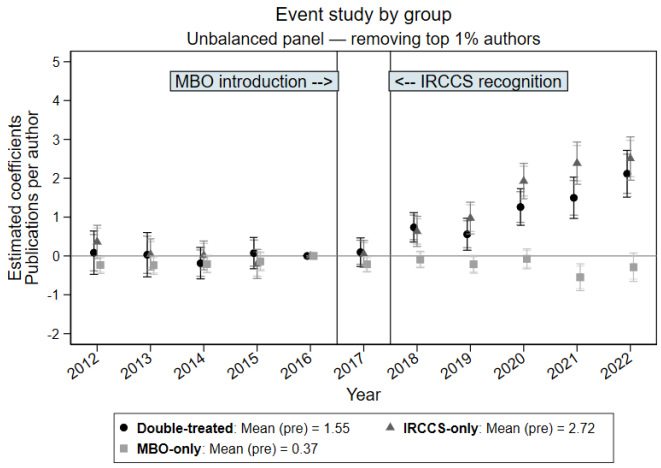


b) Removing top 10% researchers.

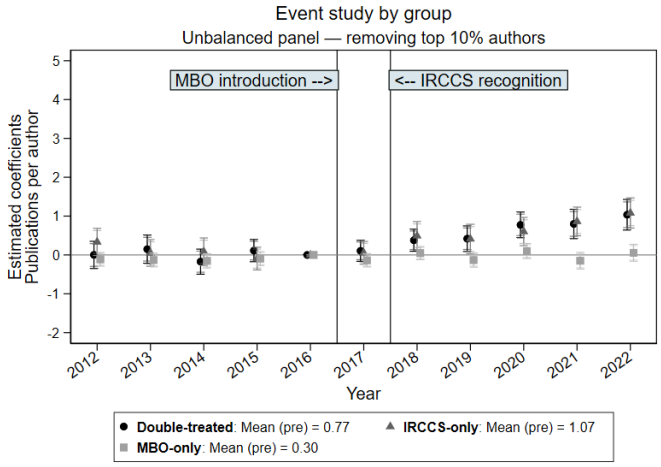


c) Removing top 15% researchers.

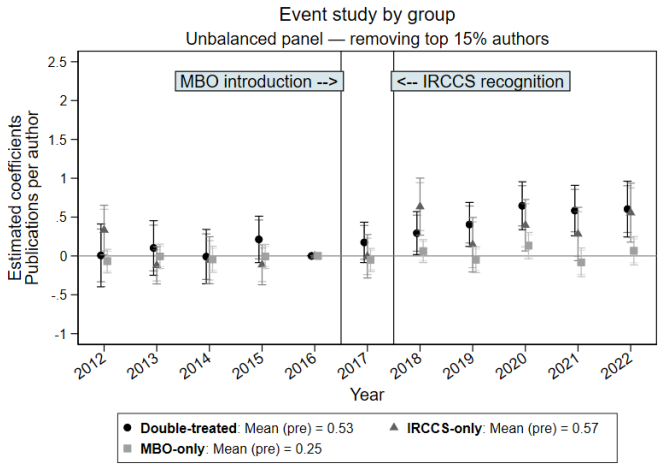
Figure A37: Balanced panel - Dynamic impact of MBO-implementation and IRCCS recognition on annual publications in the sub-samples obtained by excluding most prolific authors.



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.

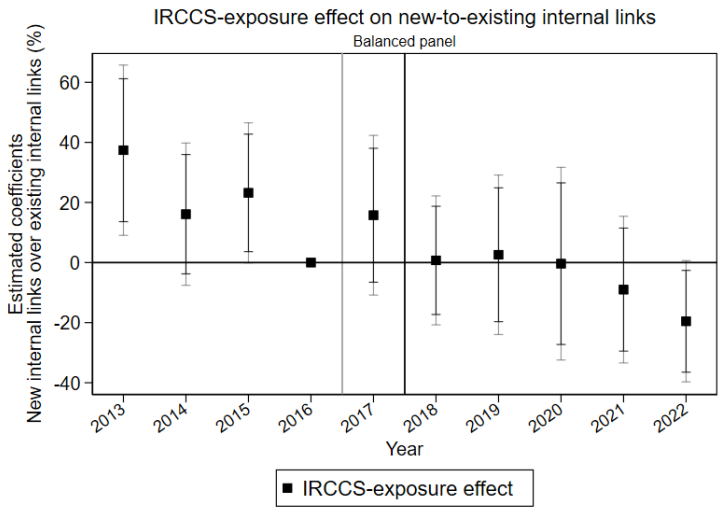


Figure A39: Event-study estimates for the share of joiners over total internal co-authors.

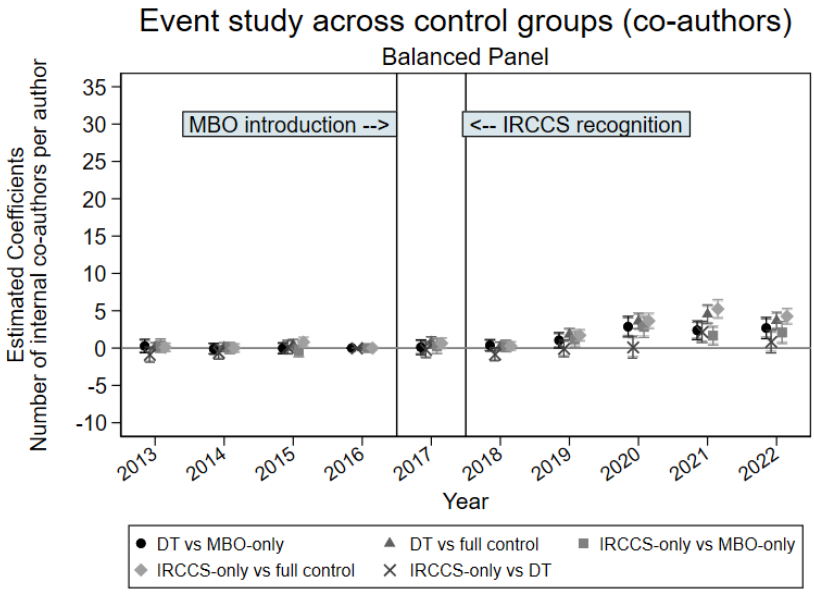


Figure A40: Event-study estimates for the total internal co-authors across alternative treatment group comparisons.

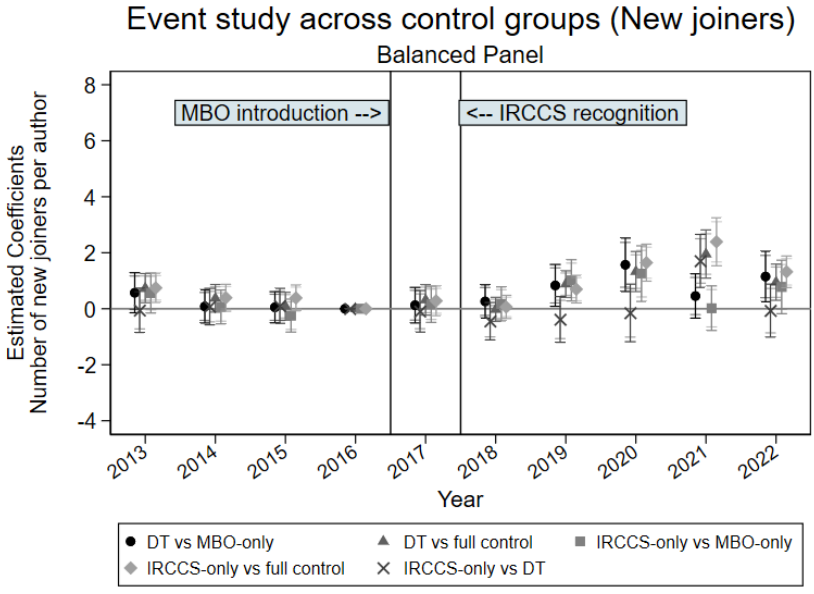


Figure A41: Event-study estimates for the number of joiners across alternative treatment group comparisons.

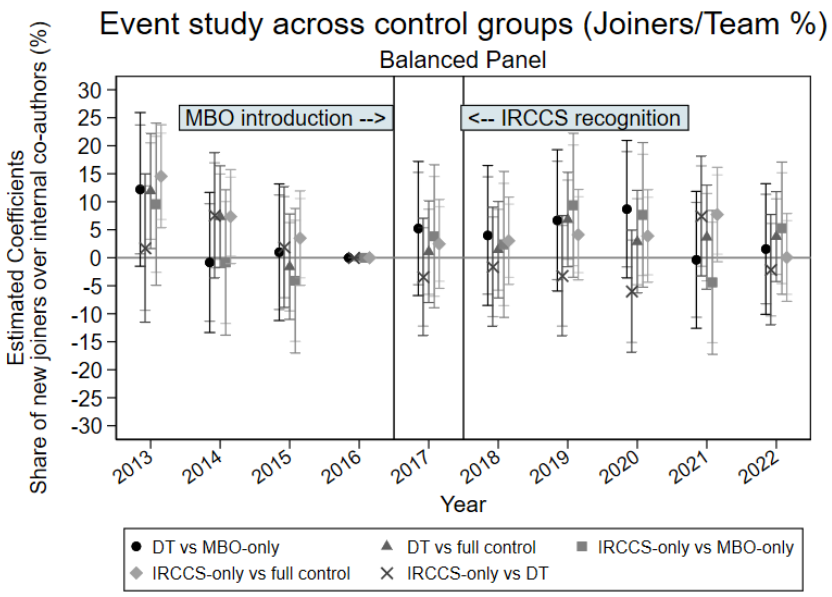
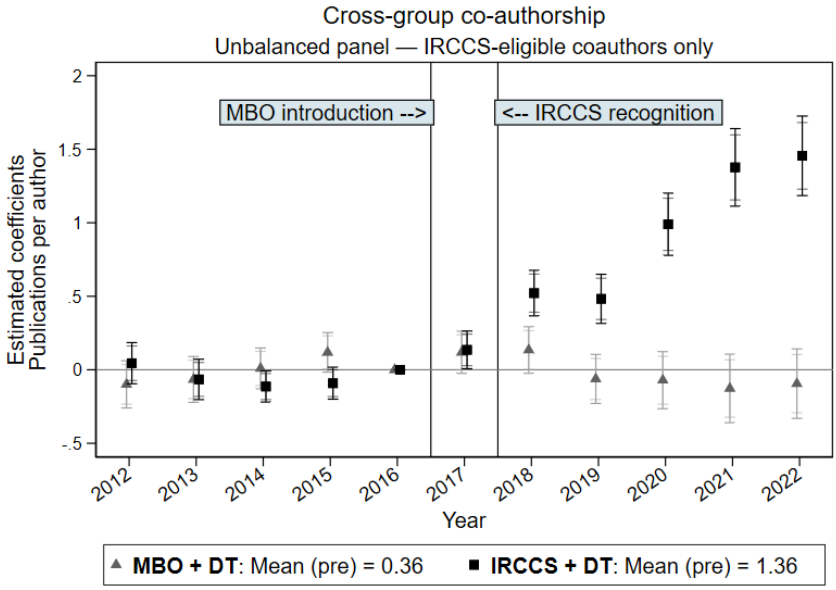
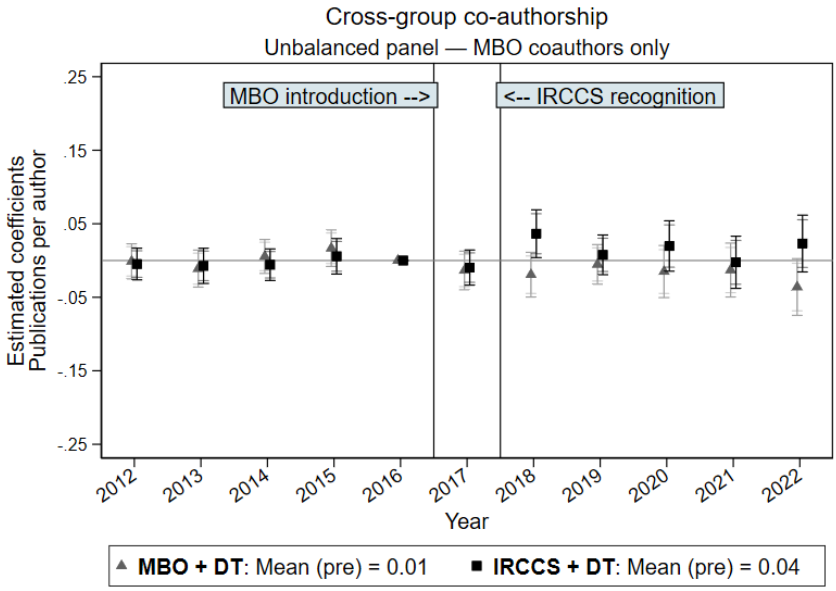


Figure A42: Event-study estimates for the nshare of joiners over total internal co-authors across alternative treatment group comparisons.

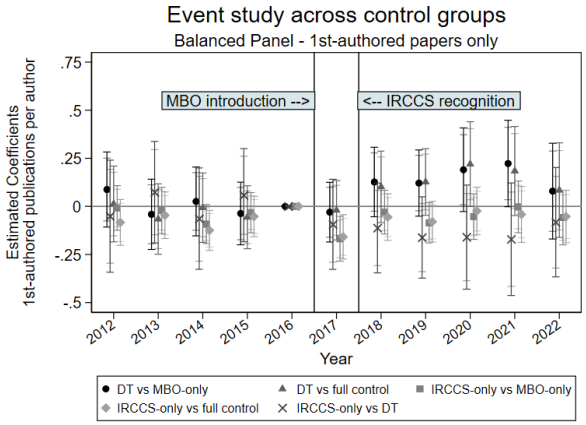


a) Outcome: publications co-authored with IRCCS units only.

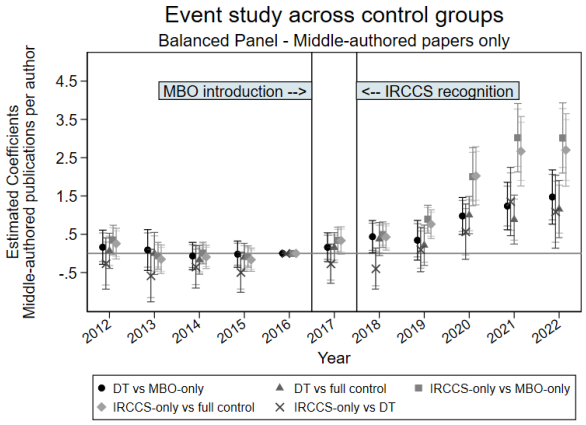


b) Outcome: publications co-authored with MBO-eligible units only.

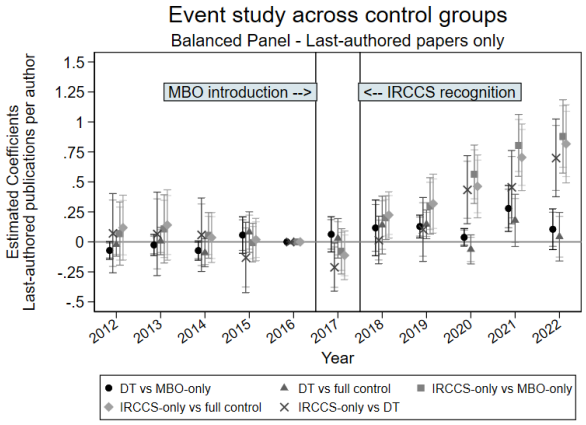
Figure A43: Dynamic impact of the policy implementation on annual cross-group publications in the balanced panel.



a) First-authored publications.



b) Middle-authored publications.



c) Last-authored publications.

Figure A44: Event-study estimates across alternative treatment group comparisons for authorship positions.

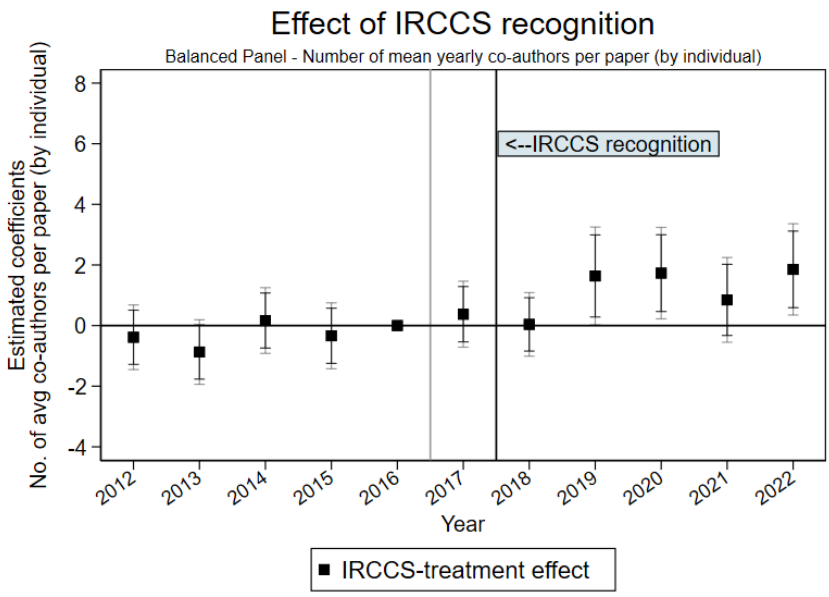
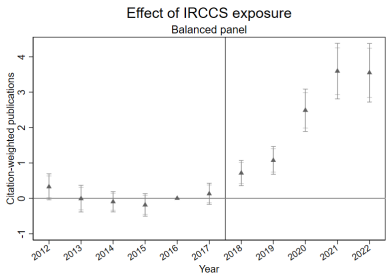
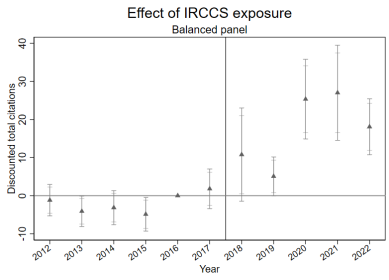


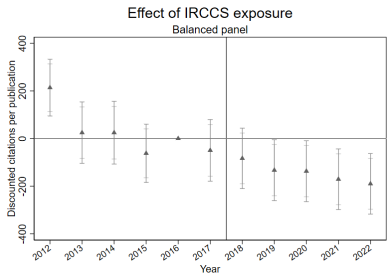
Figure A45: Dynamic effect of IRCCS recognition on the average number of co-authors per publication in the balanced panel. The IRCCS-treated group includes both IRCCS-only and double-treated physicians.



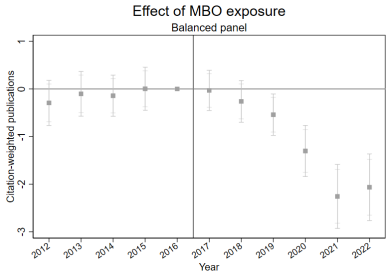
a) IRCCS exposure: citation-weighted publications.



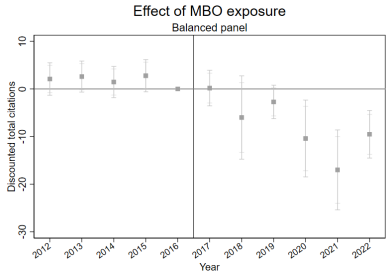
b) IRCCS exposure: discounted total citations.



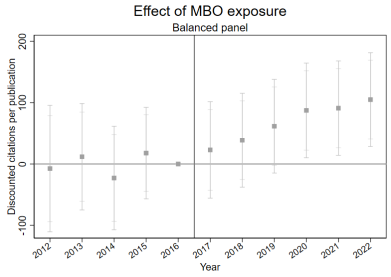
c) IRCCS exposure: citations per publication.



d) MBO exposure: citation-weighted publications.



e) MBO exposure: discounted total citations.



f) MBO exposure: citations per publication.

Figure A46: Pooled event-study estimates of citation-based outcomes. IRCCS exposure includes IRCCS-only and double-treated physicians, while MBO exposure includes MBO-only and double-treated physicians.

	(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

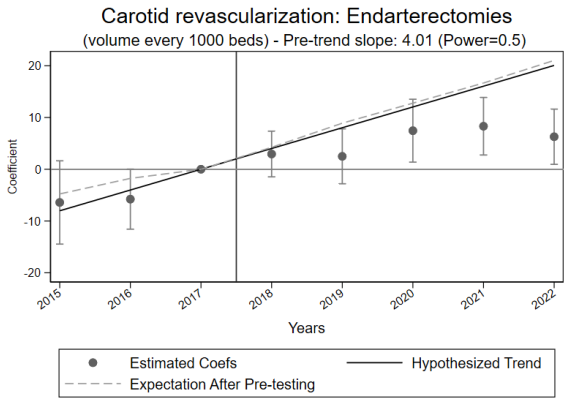
Table A13: Regressions between enrollment into IRCCS perimeter and several factors.

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)						

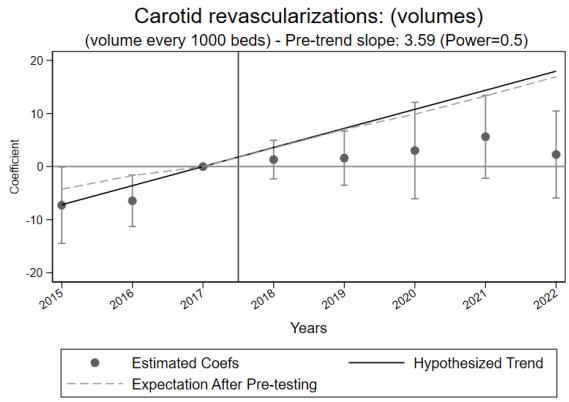
Table A14: Each row reports results from a separate regression of yearly PNE outcomes on a post-treatment indicator, for the studied hospital controlling for unit and year fixed effects, compared to a set of control hospitals within the same region. Standard errors, reported in parentheses beneath the 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)						

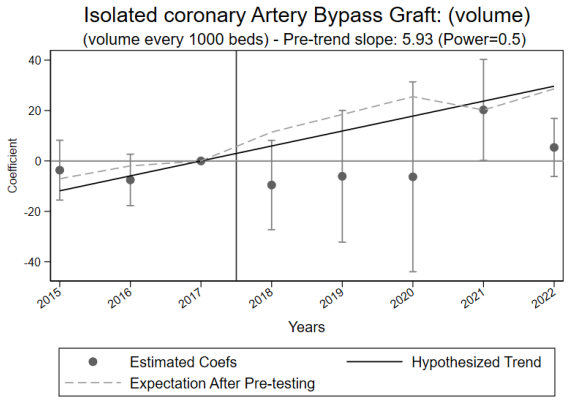
Table A15: Each row reports results from a separate regression of yearly PNE outcomes on a post-treatment indicator, normalized for 1000 beds, for the studied hospital controlling for unit and year fixed effects, compared to a set of control hospitals within the same region. Standard errors, reported 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.



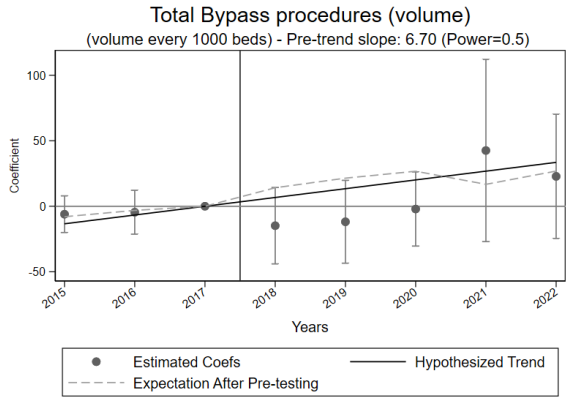
a) Carotid revascularization: endarterectomies



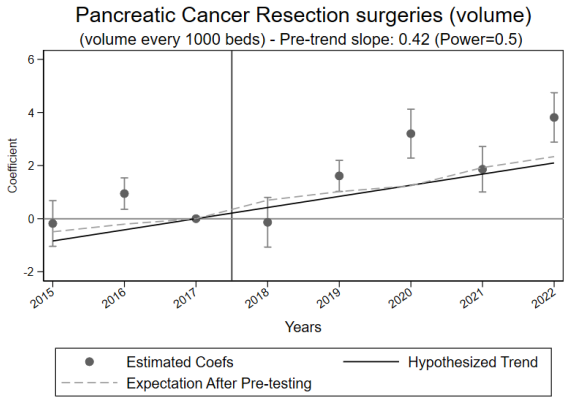
b) Carotid revascularizations (total).



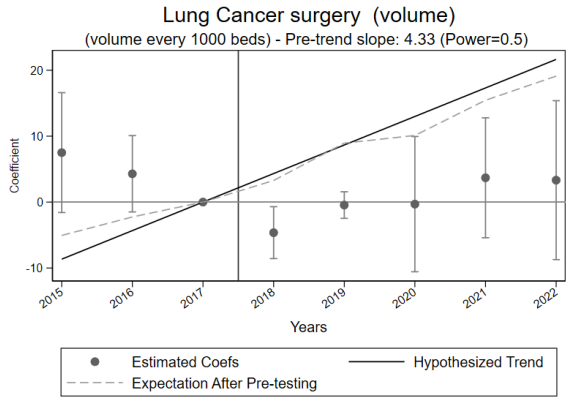
c) Isolated coronary artery bypass graft .



d) Total bypass procedures.

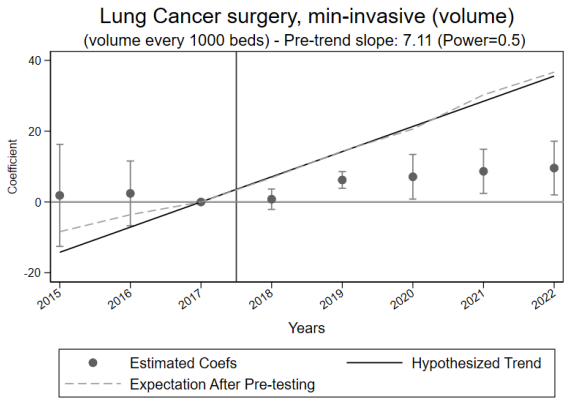


e) Pancreatic cancer resection surgeries.

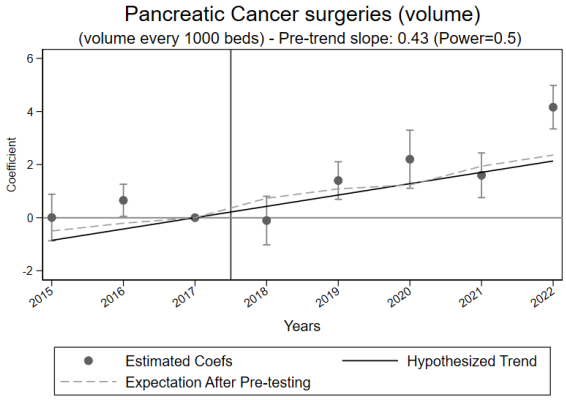


f) Lung cancery surgery procedures.

Figure A47: 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 A48: 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.