‘Returning Brains’: Tax Incentives, Migration, and Scientific Productivity

By Pier Paolo Creanza*

This paper investigates how tax incentives for high-skill immigrants affect productivity. I collect data covering 90 percent of Italian faculty between 2000 and 2020 and use it to evaluate a 2004 tax break targeting researchers. First, the program induced substantial migration and positive selection of beneficiaries. Second, higher-productivity hires significantly increase their academic group's average productivity, roughly split between their direct contribution and indirect responses of local faculty. Third, this indirect effect is largely explained by higher-productivity local researchers sorting into the treated group, rather than by productivity spillovers on incumbent researchers.

[Click here for the most recent version]

I. Introduction

The migration of highly skilled workers is a critical policy issue in knowledge-driven economies. Faced with concerns about ‘brain drain’, many countries have experimented with policies intended to induce immigration (or return migration) of high-skill individuals. Tax-based incentives are particularly common and have proven successful in encouraging high-skilled migration (Kleven et al., 2020). However, an important question remains unanswered: do these policies generate productivity benefits that justify their substantial fiscal costs? This question has been left open by seminal contributions in the literature (Kleven, Landais and

* Princeton University, pcreanza@princeton.edu. Financial support from the Princeton Industrial Relations section is gratefully acknowledged.

Saez, 2013; Kleven et al., 2014). Specifically, understanding beneficiaries' direct contributions to productivity within their new organizations and whether they generate significant spillovers on incumbent workers is essential for assessing the overall effectiveness of these policies. Notably, most available causal evidence on productivity effects in knowledge-intensive professions relies on exceptional, crisis-induced shocks, such as the collapse of the USSR (Borjas and Doran, 2012), the rise of the Nazi Party in Germany (Waldinger, 2012; Moser, Voena and Waldinger, 2014), or sudden deaths (Azoulay, Zivin and Wang, 2010). While these studies provide valuable insights into peer effects, their findings may not generalize to the routine policy interventions that governments actually deploy, highlighting the need for evidence on productivity effects in standard policy environments.

This paper provides new evidence on these questions by examining Italy's *rientro dei cervelli* (returning brains) tax scheme that targets researchers and university professors. This policy took effect in 2004. Provided that they maintain their fiscal residence in Italy and work as researchers—either in universities or in the private sector—beneficiaries are granted a 90 percent reduction in their taxable income. For a professor with a gross annual salary of €120,000 in 2011, the scheme results in tax savings of approximately €45,700, virtually eliminating the entire tax obligation. The eligibility requirements are minimal: the scheme applies to individuals of any nationality who have worked abroad as researchers for at least two years before moving to Italy.

Focusing on this setting allows me to address two key challenges in estimating the productivity effects of preferential tax schemes. First, contrary to most organizational settings, scholars' productivity can be measured by their scientific publications—authorship maps contribution and citation counts capture impact. By contrast, most other knowledge-intensive professions do not have such clear measures of individual productivity. Thanks to data sources unique to Italy, I was able to gather comprehensive information covering nearly all Italian faculty and their publications between 2000 and 2020. Second, this setting features demand constraints that ease concerns about endogeneity of the timing of a hire. Budget cuts, strict turnover rules, and occasional hiring freezes constrained demand by public universities between 2003 and 2015. This significantly reduced an university's ability to perfectly manipulate the timing of new hires.

As a first step, I examine the migration and selection effects of the tax scheme on the inflow of foreign-based scholars in Italian academia. I show that the share of scheme-eligible scholars in total hires saw a substantial increase of more than ten percentage points, from 2 to 12.5 percent. By contrast, the proportion of ineligible hires—those moving from abroad who did *not* qualify for the tax break—only rose marginally, by less than two percentage points. Next, I find that, on average, eligible hires exhibit higher productivity than other entrants.

I then use a dynamic Difference-in-Differences design to estimate how this inflow of more productive scholars affected scientific productivity. My findings reveal substantial positive effects on overall group productivity. However, these effects are largely explained by the direct contributions of highly productive hires and by higher-productivity local researchers sorting into treated groups, rather than by productivity spillovers on incumbent researchers.

In my main empirical strategy, I compare units (academic groups or individuals) that are exposed to hires with productivity significantly above the group average to units with hires who are closer to the average. This comparison by treatment intensity—rather than a treated-untreated contrast—reduces concerns about selection because it conditions on groups that have all selected into making an eligible hire, thereby controlling for factors that lead departments to hire eligible scholars in the first place. This approach isolates the effect of an eligible hire's greater relative productivity, rather than the effect of making a hire *per se*.

Though the ability to attract more or less productive eligible hires may still be endogenous to a department's trajectory, I address this concern in two ways. First, I examine pre-trends and find that groups treated by hires of different productivity levels were not on divergent trajectories in the five years preceding the event. This supports the parallel trends assumption necessary for identification and suggests that differences in outcomes are not driven by pre-existing trends. Second, I define treatment intensity based not on the observed productivity at the time of hire—which departments might select on—but on an estimate of lifetime *permanent* productivity. This measure is derived from an individual fixed effects model, following Mas and Moretti (2009), that accounts for all output and controls for peer effects, field-wide trends, and institutional moves. Since this measure incorporates future productivity and adjusts for various factors, it is not directly observable at the time of the hire and is thus less likely to be perfectly selected by the hiring department.

To further address concerns about contemporaneous shocks, I focus on academic groups within public universities that see an eligible hire during the period of constrained demand (up to 2015). This restriction ensures that the events considered happen at a time when universities could not perfectly manipulate the timing of new hires, as this was determined by ever-changing turnover rules and hiring freezes. Moreover, by comparing units that are treated in the same relative time I can further control for confounders that correlate with the hire's timing. To the extent that mean reversion, leadership, or university policy changes may drive both the hire and the observed effect, this comparison nets out common shocks around the arrival of an eligible hire.

Lastly, I show that a secondary empirical strategy that controls for contemporaneous department-level shocks yields consistent results. In this alternative approach, I compare top treated groups to their broader departments, giving further confidence that my main results are not confounded.

To fully characterize the impact of eligible hires, I consider several outcomes. First, I estimate the overall effect of scheme-induced hires on the average scientific productivity of the receiving group—the *Total Effect*. This estimate compounds both the primary mechanical effect of having a more productive researcher join the group (the *Direct Effect*) and the productivity responses of local researchers (the *Indirect Effect*). To tease these forces apart, I estimate the overall indirect effect using changes in leave-out average productivity, i.e. the group average after excluding eligible hires. In turn, this indirect effect can be decomposed into changes in faculty composition through exits and entries—a *Composition Effect*—and within-person differences in productivity induced by the new hire, a *Spillover Effect*.

Spillover and composition effects are harder to pin down because the eligible hire may affect both initial incumbents and subsequent entrants. Thus, I approach this question in two ways. At the group level, I estimate the composition effect using changes in the average *individual* fixed effect (FE) of active faculty.[1] The fraction of the indirect effect *not* explained by sorting suggests the likely extent of spillover effects, but can be confounded by movements in and out of the group. To formally estimate spillovers, I conduct an individual-level analysis where I restrict to incumbents at the time of the hire, follow them across career movements, and control for individual FEs.

The results reveal productivity gains that operate primarily through talent reallocation rather than spillovers.

At the group level, the analysis reveals positive direct, indirect and composition effects from eligible hires whose permanent productivity is significantly above the group average. After three years—the typical duration of a pre-tenure contract—overall group productivity (*Total Effect*) increases by more than 0.14 standard deviations of citation-weighted output (SDs), while leave-out productivity excluding eligible hires (*Indirect Effect*) increases by slightly less than 0.08 SDs. Critically, I tease-out the contribution of compositional changes as higher-productivity local researchers sort into groups with top eligible hires: from the 0.08 SDs *Indirect Effect*, 0.05 SDs are explained by sorting (*Composition Effect*).

[1]This can be interpreted as a measure of average individual permanent productivity. The estimation of individual permanent productivity is detailed in Section IV.B. Fundamentally, it amounts to individual fixed effects (FEs) from a regression that controls for peer effects as done by Mas and Moretti (2009). Notice that again I only consider non-eligible faculty, which includes both incumbents and non-eligible new entrants.

This decomposition already suggests limited scope for individual-level spillovers. Average effects over the whole six-year post-period are smaller because magnitudes build up gradually. Furthermore, heterogeneity analyses suggest that the strongest effects accrue to groups starting with intermediate levels of scientific productivity and whose faculty, before 2004, had high levels of international mobility.

At the individual level, my findings confirm the intuition from group-level results and show that the within-person *Spillover Effect* is zero and insignificant, and homogeneously so along different cuts of the data. Confidence bounds at the 95-percent level rule out changes greater than about 0.03 SDs in either direction.

Taken together, the evidence presented here offers a mixed picture on the effectiveness of the 'returning brains' policy. On the one hand, the tax-scheme has attracted a sizeable number of foreign-based scholars at a time when Italy was seeing large outmigration of researchers (OECD, 2017). These scholars have on average higher productivity than local researchers and contribute to their host universities. On the other hand, the scheme is loose enough that a large fraction of beneficiaries—about one third—fall below the average productivity of the receiving academic group. Further, the apparent productivity increase in groups exposed to the top hires is largely a result of composition effects—at the individual level, incumbent scholars do not systematically increase their productivity.

Thus, the evidence suggests that the tax scheme produces a reallocation of talent, likely a success from the point of view of the local policy-maker. It is important to acknowledge that the beneficiaries may be contributing in other dimensions not considered in this paper, like teaching, prestige and funding. However, from a global efficiency point of view, there are clear costs to this policy—such as tax competition and distortions in the international academic labor market—that are not offset by individual level spillovers.

This paper contributes to the literature on tax-induced mobility of high-skilled workers.[2] Top professionals exhibit high mobility in response to tax changes (Kleven, Landais and Saez, 2013; Akcigit, Baslandze and Stantcheva, 2016). Kleven et al. (2014) specifically explore preferential tax schemes, focusing on a Danish policy targeting high-income earners, and report substantial mobility responses. This trend extents to within-country mobility (Agrawal and Foremny, 2019; Moretti and Wilson, 2017) and tax policies targeting broader income segments (Timm, Giuliodori and Muller, 2025; Bassetto and Ippedico, 2023). Notably, Bassetto and Ippedico (2023) present strong migration responses to a similar Italian tax break aimed at highly educated workers.

To the best of my knowledge, this is the first paper that causally estimates the productivity effects of preferential tax schemes affecting migration. The closest related work is by Prato (2025), who employs a structural model to demonstrate that such tax schemes can mitigate brain drain and boost local productivity in the short term. However, Prato (2025)'s model also suggests that such policies may reduce international collaboration in the long run.

Beyond preferential tax policy, this article speaks to broader debates on migration and innovation. Empirical work has often found that immigration benefits innovation, both historically (Moser, Voena and Waldinger, 2014; Moser, Parsa and San, 2020) and in contemporary settings (Kerr and Lincoln, 2010; Hunt and Gauthier-Loiselle, 2010; Kerr et al., 2016; Bernstein et al., 2022; Burchardi et al., 2020; Azoulay et al., 2022). However, Borjas and Doran (2012) and Doran, Gelber and Isen (2022) offer counterexamples where immigrant inflows do not raise productivity, but rather displace native workers.

Lastly, this study contributes to the ongoing debate on peer effects in the workplace. Mas and Moretti (2009) observed significant effort spillovers among supermarket workers, while Azoulay, Zivin and Wang (2010) found evidence of knowledge spillovers among medical researchers. However, Waldinger (2012) did not observe significant spillover effects in 1930s German academia. Similarly, Cor-

[2]For a comprehensive review on this topic, refer to Kleven et al. (2020).

nelissen, Dustmann and Schönberg (2017) found no significant wage peer effects among knowledge workers.

The paper is organized as follows. Section II provides background information on the 'returning brains' tax scheme and Italian academia, Section III describes the data, Section IV shows migration and selection effects of the policy, Section V sets out the empirical strategy whose results are reported in Sections VI and VII. Section VIII concludes.

II. Background

A. The 'returning brains' preferential tax schemes

Rientro dei cervelli (returning brains) policies encompass a range of initiatives aimed at encouraging the return migration of high-skilled Italian expatriates and attracting similarly skilled foreign professionals. Most of these initiatives are fiscal in nature, including various preferential tax schemes.[3] This paper focuses on the most generous and long-established scheme, targeting researchers and university professors.

The 'returning brains' tax break for researchers benefits scholars who have worked abroad and moved to Italy. The scheme applies to researchers and university professors of any nationality, not just Italian citizens. The benefit consists in a 90 percent reduction in taxable personal income for a period of time, originally three fiscal years, and as long as beneficiaries maintain their fiscal residence in Italy. In order to be eligible, individuals must: (*i*) have resided outside of Italy for at least two consecutive years, (*ii*) have worked as researchers or university professors for foreign university or research centers for the same period, (*iii)* have an appropriate university degree, (*iv*) move their fiscal residence to Italy. Though

[3]For instance, a 2015 reform introduced a 70 percent reduction in taxable personal income for workers moving to Italy for at least two years, under certain conditions (see law *DL n. 147/2015*). A 2016 reform offered a €100,000 substitute tax on foreign income for individuals moving their fiscal residence to Italy after residing abroad for at least nine years (see law *Legge n. 232/2016*). A no longer active provision gave a 70 percent income tax reduction to EU citizens with specific criteria (see law *Legge n. 238/2010*). Additional benefits could be obtained based on gender, children, and location within Italy.

not spelled out in the letter of the law, the scheme only applies to research-related income.[4]

The tax scheme was initially passed in late 2003, with 2004 being the first treated year. See Table A1 in the Appendix for precise legislative references. This policy was initially conceived as temporary, but it has been renewed several times and made permanent in 2016. Benefit duration has also been changed to four years since 2015 and six years since 2020. Amendments passed in 2019 to take effect from 2020 have further expanded benefit duration when some conditions are met, like having under age children or purchasing a house.

Overall, the tax break is exceptionally generous. For instance, in 2011 an ineligible professor with a gross salary of €80,000 would have faced a marginal tax rate of 43 percent and an average tax rate of 35 percent.[5] Were she eligible, given a no tax area around €8,000, she would not have had any personal income tax liability. Thus, over the original three-year benefit period she would have saved around €82,000 in taxes. Table A2 in the Appendix exemplifies benefit levels in greater detail.

The 'returning brains' scheme operates at significant cost, as I estimate around one billion euros of foregone tax revenue between 2004 and 2019. It should be noted that the tax break also applies to researchers in the private sector and to both wages and self-employment income as long as they pertain to research activities. Thus, the pool of beneficiaries is wider than just eligible faculty and potentially affects more income than just their academic salaries. Researches in university make up roughly 15.5 percent of all beneficiaries. Using official statistics starting in 2012 and my estimates of the eligible population before then, back of the envelope calculations suggest a total cost of nearly €960 million in my sample period (2004-2019), with around €150 million accounted for by beneficiaries in universities. See the Appendix, Table A3, for more details. Given

[4]See *Circolare n. 22/2004, Agenzia delle Entrate.*

[5]Then, net income would have been around €52,700, abstracting from deductions and other special circumstances.

a yearly cost around €70-75 million in the last few years, by now total cost has likely exceeded €1.2 billion. For a benchmark, in 2019 ordinary government funding for public universities (which does not include individual research grants and EU funds) totalled €7.1 billion.[6]

B. Italian academia

The Italian tertiary education system bears close resemblance to those in the US and other advanced economies. Research and teaching are primarily organized through universities, supplemented by research centers and specialized institutions. In terms of faculty, a *Professore ordinario* corresponds to a Full Professor, a *Professore associato* to an Associate professor and a *Ricercatore* to an Assistant Professor. An *Assegnista di ricerca* roughly corresponds to a Post-doc. A variety of fixed-term teaching or research collaboration contracts exists.[7]

Notably, Italian academia is predominantly public, with over 96 percent of faculty employed in public universities in 2019. Despite this homogeneity, there is a wide spectrum in terms of university sizes, academic prestige, and scientific output. All public universities and their faculty are mandated to engage in both teaching and research. Private universities are overall much smaller and more specialized than public universities.

A significant feature of the period studied in this paper is the pronounced constraint on labor demand in the academic job market. As part of broader fiscal austerity measures, since the early 2000s the Italian government has introduced various limits on hiring and strict turnover rules. These restrictions also included two major hiring freezes in 2003 and 2009. Not only the inflow of new hires decreased, but turnover rules that did not allow for one-to-one replacement of

[6] See the *Decreto Ministeriale n. 738 dell'8/8/2019 relativo ai criteri di ripartizione del Fondo di Finanziamento Ordinario 2019 - Università Statali.*

[7] The main differences in terms of faculty concern promotion and tenure. Firstly, until 2010 *Ricercatori* could be tenured—hired permanently in this position and never advance to higher ranks. A reform then introduced two three-year terms for *Ricercatori* that, similarly to the tenure clock, lead to either promotion to associate or termination (the corresponding law is *Legge n. 240/2010*). Secondly, the same 2010 reform introduced the requirement of obtaining a national scientific habilitation (ASN) to the rank of associate or full professor in order to be considered for these positions.

positions opened up by retirements and cessations reduced the total number of faculty employed in public universities. Section A.3 in the appendix reviews the provisions constraining demand between 2003 and 2015, and shows trends in total faculty and new hires (Figure A1).

III. Data

A. Sources

Faculty employment. — I reconstruct the universe of Italian faculty from publicly available faculty rolls.[8] Compiled by the Ministry of Education, University and Research (MIUR), these lists are based on mandatory administrative reports from all universities, both public and private, and represent total employment as of December 31st each year. Beginning in 2000, these rolls include the full legal name, gender, university affiliation, and academic field of each faculty member, though they lack unique identifiers. After disambiguating the vast majority of individuals, the final faculty panel encompasses 88,923 academics, resulting in 1,129,054 person-year observations.

Furthermore, I gather additional career data on all individuals who joined Italian academia between 2001 and 2019. At a first approximation, I identify new hires by comparing changes in faculty composition annually, yielding a sample of 41,501 entries.[9] This extensive data collection effort was performed manually, and entailed searching for each hire online and extracting information from possibly varied sources like CVs, institutional websites, LinkedIn pages, biographies, and so forth. See Section B.1 in the Appendix for the algorithms used to find and code information. To minimize the costs associated with manual search, I only collected information that was strictly necessary to assess eligibility for the tax break. Panel A in Table 1 illustrates summary statistics on key career data.

[8]These records are located on a government website, `https://cercauniversita.cineca.it`.
[9]This is an upperbound on the number of entries, as it includes individuals switching universities or reentry after spells outside Italian academia. This first definition of new hires was intentionally loose in order to allow manual disambiguation.

Scientific output. — I obtain research output data from institutional repositories (IRs) that the vast majority of Italian universities maintain. These online platforms document a wide range of scientific activity, not just journal articles but also books, chapters, conference material, and so forth. Each record has information on its authors, year of publication, and rich but non-homogeneous metadata. Critically, whenever possible records are linked to Scopus, Web of Science and PubMed—popular indexing services that allow to reconstruct citation counts.[10] My scientific output dataset is based on the IRs of universities accounting for 93% of total faculty in 2019. Table B4 in the Appendix provides details on university coverage, faculty employment in 2019, and the volume of IR records available.

I exactly match faculty and output datasets on names, surnames, their abbreviations and permutations thereof. This approach is made necessary because faculty rolls and IRs lack comprehensive and homogeneous individual identifiers. This matching technique has a few limitations. First, exact matching is sensitive to spelling errors and typos, though these are unlikely in the authorship of academic articles. Second, not all available output that we know to have been produced by at least one Italian-affiliated faculty will be linked and used in the analysis. Despite this, exact matching of name variations links 88 percent of all publication records to at least one faculty in my disambiguated faculty panel. Given attrition from disambiguation, the maximum matching rate would be around 93 percent.

Additional department characteristics are from the government-mandated 2001-2003 research quality assessment report (*Valutazione triennale della ricerca* or VTR).[11]

[10] At the time of data collection in summer 2022, the IRs of only five small universities did not have citation data.

[11] Available online. Working link as of January 2024: `https://www.unipd.it/valutazione-triennale-ricerca-vtr-2001-2003`.

Table 1—Summary statistics.

Variable	N	Mean	S.D.	P10	P90
Panel A: New hires career data					
Has worked abroad (Y/N)	37,899	0.10	0.30	0	0
Number of years w. a.	3,701	4.81	3.57	1	12
Gap (years) between moving and uni. job	3,607	2.45	4.04	0	8
If gap>0, has worked in research (Y/N)	1,633	0.89	0.31	0	1
Eligible at hire	3,700	0.56	0.50	0	1
Panel B: All faculty output data					
Public university (Y/N)	982,255	0.96	0.20	1	1
Academic rank	982,255	3.84	0.79	3	5
Fractional citations	982,255	17.13	69.19	0	46
Fractional output	982,255	1.56	2.30	0	4
IHS of fractional citations	982,255	2.77	3.00	0	7
IHS of fractional output	982,255	1.54	1.16	0	3
Any citations	982,255	0.49	0.50	0	1
Any output	982,255	0.78	0.41	0	1
Panel C: Ever-treated faculty output data					
Public university (Y/N)	176,363	0.96	0.19	1	1
Academic rank	176,363	3.87	0.78	3	5
Fractional citations	176,363	22.85	74.11	0	58
Fractional output	176,363	1.45	2.03	0	4
IHS of fractional citations	176,363	3.52	3.01	0	7
IHS of fractional output	176,363	1.55	1.09	0	3
Any citations	176,363	0.62	0.49	0	1
Any output	176,363	0.83	0.38	0	1

Note: In panel A, the number of years worked abroad is top-coded at 12.5 for all hires who have worked abroad strictly more than 10 years. The gap measures the number of years between someone moving to Italy from abroad and starting a university job in my panel. In panels B and C, academic rank equals 3 for assistant, 4 for associate and 4 for full professors; fractional output is defined dividing each item by the number of authors; the scale factors used for all IHS transformations are reported in the Appendix, Table B5.
Source: Career data was collected manually. Scientific output data is from the Institutional Repositories (IRs) of individual universities. Additional details on the data collection process can be found in the Appendix, Section B.

B. Main outcome variable

My preferred metric of productivity is citation-weighted scientific output. Specifically, I focus on the *Fractional citations* measure. For a given author i and year

t, this metric aggregates the author-adjusted citations ever accrued by works published by i in year t. Importantly, this measure is based on lifetime citations recorded as of summer 2022, regardless of the year of publication. Thus, it weights output published in year t by its total eventual citations, mitigating the confounding effects of evolving citation practices (a risk in metrics based on citation flows). This approach more accurately reflects the 'revealed' importance of a contribution. Details on adjustments for newer publications with shorter citation histories are provided in the Appendix, Section B.3.

The distribution of citations is right-skewed, a feature that is often dealt with by taking the natural logarithm. However, my data contains many zeros, which make the log transformation unsuitable. Note that these zeros mostly arise from individuals that do not publish (cited) material every year, rather than from fully inactive scholars, blurring the distinction between extensive and intensive margin.[12] A common solution is to use the inverse hyperbolic sine (IHS) transformation, which is well-defined at zero.[13] Recent contributions have highlighted limitations of the IHS transformation (Aihounton and Henningsen, 2021; Chen and Roth, 2023). Of the solutions proposed in the literature, I opt for an explicit calibration of the relative importance of intensive and extensive margins; to do so, I find the optimal scale factors for my main variables using Aihounton and Henningsen (2021)'s R^2 criterion—see Table B5 in the Appendix.

C. Academic fields

In the available data departments are not consistently identified; however, granular academic fields are. This categorization is done by the MIUR and is homogeneously reported in all faculty contracts. Ministry-defined fields inform both hiring and promotions—when assessing candidates, by far the greatest weight

[12]In the main regression sample, 38.4 percent of person-year observations have zero citation-weighted output, but only 10.4 percent of individuals never have positive citations. If we consider simple output counts, 17.5 percent of person-year observations are zeros, and less than three percent of individuals appear to be completely inactive. Hence, an extensive margin analysis would not be sensible here.

[13]The IHS function is $f(x, \theta) = \ln\left(x \cdot \theta + \sqrt{x^2 \cdot \theta^2 + 1}\right)$, where θ is the scale factor.

is assigned to their scientific activity in the field to which the position formally belongs. This classification of academic disciplines is available at several nested levels of aggregation. In my analysis I focus in particular on 14 macro-fields (*Aree disciplinari*) and 190 granular fields (*Settori concorsuali*).

The cells defined by the intersection of universities and granular fields are the 'groups' I use throughout my analysis. These cells follow more relevant disciplinary boundaries than large department woulds. In particular for the estimation of productivity effects, the subject matter of one's research seems more important than administrative structures.

I use macro-fields to approximate broader 'departments' and to standardize my main regression analysis outcome, citation-weighted output. The distribution of *Fractional citations* varies notably across academic fields. Therefore, I divide the outcome by the standard deviation (SD) within each of the 14 ministry-defined macro-fields. This standardization renders the results interpretable as fractions of a SD, facilitating the aggregation of estimates from different macro-fields into a coherent average.

IV. Evidence on Migration and Selection

In this section I examine the migration and selection effects of the 'returning brains' tax-scheme. Consistently with previous findings in the literature, the evidence shows the policy to be effective in achieving its primary objective of attracting its targeted population. More significantly, the new finding that eligible hires are, on average, more productive than their counterparts strengthens the rationale for the scheme.

A. Mobility

To evaluate the reform's impact, I employ career data to simulate eligibility for the tax break. This simulation is applicable to the post-reform period to estimate unobserved take-up. In a counterfactual sense, it is also applicable to the pre-reform period, to benchmark what fraction of pre-reform hires would have been eligible if the tax-break had been in effect. The variable *Eligible at hire* is assigned a value of one if an individual is a new hire, has previously engaged in qualified research jobs abroad for at least two consecutive years, and has secured a research position within the benefit window. Conversely, *Eligible at hire* is set to zero in all other instances, thus excluding both ineligible hires (specifically, researchers have not worked abroad or not in research roles, or those with less than two years' experience) and those who might have been eligible but accepted their current faculty position after the tax benefit had lapsed.[14]

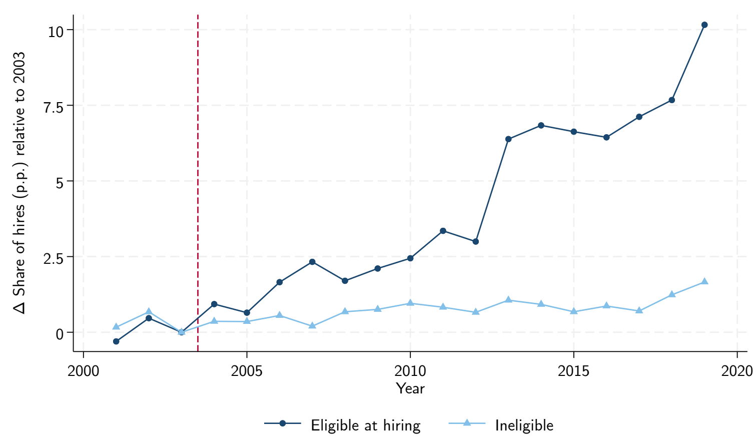


Figure 1. Hires from abroad by eligibility status (share of total)

Note: Shares are defined over annual total hires and reported as differences relative to 2003 levels. Eligibility is simulated using hand-collected career data. I show pre-reform counterfactually eligible hires to benchmark later figures.

By the end of the sample period, there is a marked increase in both the number and proportion of hires in Italian academia potentially eligible for the scheme, indicative of a significant migration effect prompted by the reform. Comparing 2001 to 2019, the count of eligible hires rose from 70 to 335. Relative to total annual hires, the fraction of eligible individuals climbed from 2 to 12.5 percent.

[14]This refers to a period of three years in Italy for those returning before 2014, or four years starting in 2015.

In contrast, the share of ineligible foreign hires, namely those moving from abroad but not qualifying for the tax break, only saw a modest increase of less than two percentage points. Considering this group as a plausible control for eligible hires, the evidence suggests that the policy accounts for the majority of the increase in the eligible entrants population.

Figure 1 displays these changes in the flows of eligible entrants in Italian academia. Flows are expressed as a share (in percentage points) of total hires for a given year, and normalized as the difference relative to 2003.

Additional findings. — First, I find that post-reform eligible entrants have longer work histories abroad, suggesting that the reform helps attract more senior scholars. Figure 2 shows the duration of their employment abroad, both before and after the reform. This chart reveals a notable shift to the right in the distribution of foreign employment duration post-reform, indicating that beneficiaries, on average, have accumulated more international experience than their pre-reform counterparts. Additionally, the absence of significant bunching around the two-year eligibility threshold is noteworthy. This lack of bunching is reassuring as it suggests limited strategic employment abroad by natives in order to qualify for the tax break. It is also consistent with the notion that academic career movements are largely influenced by the standard durations of positions and key career developments like tenure, rather than freely manipulable.

Additional findings are described here, while the respective tables and figures can be found in the Appendix, Section C.1.

As for the beneficiaries' employment spells in Italy, evidence shows appreciably low exit rates and high promotion rates. This suggests that the policy successfully motivates beneficiaries to stay longer than the duration of the benefit. Appendix Figure C6 shows the proportion of eligible hires who either exit or are promoted at various intervals post-entry. Overall, after ten years, 25.4% of eligible hires have been promoted, while only 12.7% have departed.

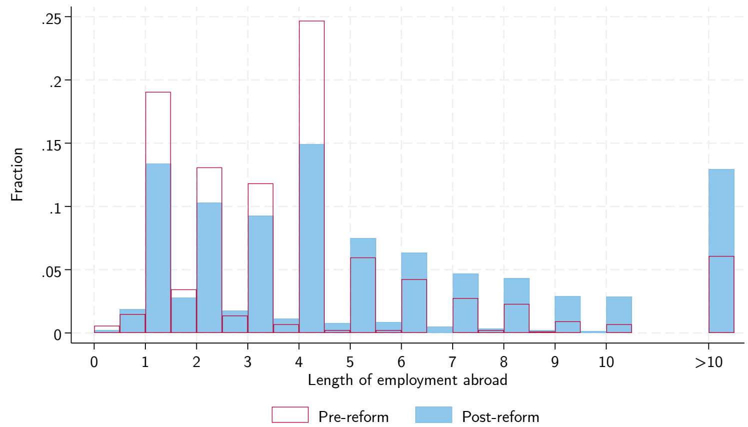


Figure 2. Years spend abroad before moving

Note: The two overlaid histograms show, separately by pre- and post-reform period, the duration of abroad employment for any hire who has worked abroad for a positive amount of time in a qualified research job. Duration is top-coded as 12.5 years for lengths strictly greater than ten. The spike around year 4 is mostly due to an imputed PhD duration for 4 years whenever this was not otherwise specified.

Appendix Table C6 presents the breakdown of eligibility for the tax-break by macro-field and time period. This table reveals notable initial disparities in internationalization levels across fields. Furthermore, it shows that, although the proportion of eligible hires rose in every field, the extent of this increase varied significantly. Mathematics, Physics and Economics were the most internationalized fields both before and after the reform. Table C7 illustrates the beneficiaries' countries of origin. The vast majority moved from other European countries, but a notable one fifth of beneficiaries moved from the United States.

Lastly, I consider suggestive evidence on possible displacement effects on locally trained researchers. Trends from high-quality surveys from the Italian National Statistical Office (ISTAT) are consistent with displacement effects induced by the greater inflow of eligible hires from abroad, and with the broader period of constrained demand from public universities. While the overall employment rate was high and fairly constant, the fraction of locally-trained PhDs working in Italy in a research-related job declined markedly (-15 percentage points between the

2004 and 2014 cohorts). This decline is entirely driven by a reduction in the share that works in Italian academia in particular (-17 percentage points). See Table C8 for more details.

B. Productivity

To determine if the policy succeeded in attracting highly productive academics, I analyze their scientific output up to the time of hire, comparing eligible individuals with all other hires. This involves testing the hypothesis that the target group exhibits higher productivity through the following equation:

$$OutputAtHire_i = \alpha + \beta Eligible_i + \phi_t + \rho_r + \varepsilon_i \tag{1}$$

Here, i represents an individual hire entering in year t with rank r. $Eligible_i$ is a binary variable set to one for hires eligible for the tax break (and zero for both native hires and ineligible individuals), ϕ_t are year-of-entry fixed effects (FEs), ρ_r are academic rank, FEs and errors ε_i allowed to be heteroskedastic. A positive and significant value of β would suggest that eligible hires display higher average productivity, even when accounting for compositional effects at the temporal, field and rank level.

The results in Table 2 consistently indicate that eligible hires are more productive than other new entrants across various output metrics. Column (1) uses total citation-weighted productivity up to year t, without accounting for the number of authors or the years of activity.[15] Column (2) adjusts this metric for the number of authors per publication,[16] and column (3) further divides by the number of years a scholar has been active by time, to account for mechanically higher output of more senior hires. Column (4) focuses on author and time-adjusted counts of

[15]As all my main outcomes, this is IHS-transformed. In the Appendix, Table C9 shows that the results in columns (1)-(3) do not rely on the IHS transformation.

[16]More precisely, if item (e.g. article) j by author i published in year t is coauthored by N_j people (including i), then citations garnered by item j are given weight $1/N_j$ when computing i's citation-weighted output in year t.

Table 2—Hire's productivity differences at the time of entry

	Citations			Papers
	IHS	Author-adj.	Author & Time-adj.	Author & Time-adj.
	(1)	(2)	(3)	(4)
Eligible	1.42**	1.98***	0.36***	-0.03*
	(0.66)	(0.72)	(0.05)	(0.02)
Observations	25,565	25,565	25,565	25,565
R^2	0.45	0.46	0.46	0.08
Controls	Y	Y	Y	Y
D.V. Mean	21.99	27.34	2.32	0.58

Note: *** p<0.01, ** p<0.05, * p<0.1, robust standard errors in parentheses. Controls: flexible time trend by macro-field, academic rank FEs. Column (1) shows IHS-transformed total citations garnered by all output published by the time of hire. Column (2) divides citations by the number of authors of each item, and then IHS-transforms; Column (3) further divides pre-hire citation counts by the number of years a hire has been active before entry, where activity is defined by the publication record; Column (4) shows results for the (untransformed) number of papers published before entry in the faculty panel, adjusting for number of authors and years active.

published journal articles, to discern whether the observed differences stem from a higher number of publications or from the greater scientific impact of eligible hires' work. The findings from column (4) suggest that the higher productivity is due to the greater impact of the publications by eligible hires.

Distribution of permanent productivity. — To move beyond on-average characterizations, and later for defining treatment intensity, I estimate individual permanent productivity for all scholars in my sample. In order to net out any peer effects, I follow Mas and Moretti (2009) in estimating individual productivity while controlling for peer composition. To achieve this, I estimate the following equation:

$$y_{ict} = \theta_i + M'\gamma_{Pi} + \delta_c + \tau_{ft} + \varepsilon_{ict}. \tag{2}$$

Here, i indexes the individual, c the group (university-granular field cell), t the calendar year, f is the academic macro-field. The vector γ_{Pi} contains every possible combination of peer composition. Thus, M includes parameters that net out permanent individual productivity from any peer influence. δ_c are group FEs, τ_{ft} macro-field by calendar time fixed effects. Finally, individual fixed effects θ_i give the estimated permanent productivity. y_{ict} is citation-weighted productivity as described in Section III.

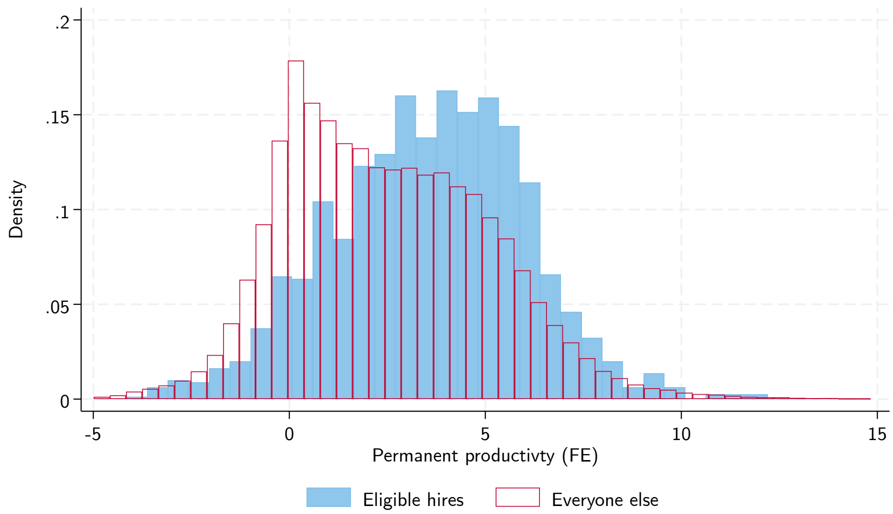


Figure 3. Distribution of permanent individual productivity by eligibility status

Note: The two overlaid histograms show the distribution of individual fixed effects from regression 2, θ_i, which can be interpreted as individual permanent productivity net of any peer effects. The 'Everyone else' population includes both incumbents, native and ineligible hires.

Figure 3 displays histograms of the estimated individual productivity θ_i, contrasting eligible hires with all other academics in my sample, namely native incumbents and all other entrants. Eligible hires are clearly shifted to the right, displaying an average θ_i that is 43 percent higher than the rest of the population. However, there is considerable overlap between the two distributions. To the extent that there is variation within a department and hires do not perfectly sort across departments, there is ample room for eligible hires' productivity to exceed, align with or—to a lesser extent—fall below the average productivity within a department.

V. Empirical strategy for estimating productivity effects

A. *Motivating and defining treatment intensity*

Arguably, selection concerns threatening identification are more pronounced when comparing treated units to fully untreated ones—for instance, departments that make an eligible hires to those that do not. Therefore, I estimate the productivity effects of eligible hires by comparing their differential effects across one key heterogeneity: the gap between their own productivity and that of their news peers. Indeed, productivity effects of eligible hires are expected to depend on this gap (Mas and Moretti, 2009; Cornelissen, Dustmann and Schönberg, 2017). A larger productivity difference between a hire and their new peers should produce a larger impact on peer productivity, if any.

Thus, I define a measure of treatment intensity at the event level that accounts for the wide distribution of scholarly productivity shown in Figure 3. An event is defined as the first instance a group hires an academic eligible for the tax scheme. Let $\theta_{h(c)}$ be permanent productivity, net of any peer effects, of the new hire $h(c)$ who joins group c; let $\overline{\theta}_c$ be the average permanent productivity of the incumbents in group c at the time of h's arrival. Then, treatment intensity for the event affecting group c is:

$$TI_c = \theta_{h(c)} - \overline{\theta}_c \tag{3}$$

Recall that θs are expressed in terms of standard deviations of the main outcome, *Fractional citations*.[17]

Events are then categorized based on the terciles of TI. The top tercile encompasses events where the new hire's productivity substantially exceeds their peers' average (gap greater than 1/5 of a SD). The bottom two terciles encompass events

[17]SDs are computed separately for each one of the 14 academic areas defined by the ministry. This standardization accounts for field-specific variations in citation practices and levels.

where hires are close to the group average. An important benefit of this treatment intensity definition is that is based on an estimate of individual *permanent* productivity, rather than on observable which departments might select on. This makes the groups in the different terciles more comparable.[18]

To see that this distinction matters, I compare treated groups to untreated ones—with and without distinguishing events by their treatment intensity (TI). This analysis is also interesting per se as it gives an estimate of the average effect of any eligible hire on the productivity of the receiving group.

I estimate the effect of any eligible hire on average group productivity through a simple Difference-in-Differences approach. In particular, I estimate this regression equation:

$$y_{ct} = \alpha_c + \varphi_t + X'_{ct}\Gamma + \beta Treat_c Post_{ct} + \varepsilon_{ct} \tag{4}$$

where c indexes the group, t the year, α_c are group fixed effects (FEs), φ_t are year FEs; $Treat_c Post_{ct}$ takes value one only after a group is treated, if ever. X_{ct} contains more flexible time trends, either by field or university, or empty. The outcome y_{ct} is average group productivity (including the new hire) which measures the *Total Effect* further explored in my main analysis. Errors are clustered at the group level.

I obtain effects by treatment intensity by splitting the quantity of interest for events within and without the top treatment intensity tercile (TIT). That is, I estimate:

$$\begin{aligned} y_{ct} = \alpha_c + \varphi_t + X'_{ct}\Gamma + \beta_1 Treat_c Post_{ct} \mathbb{1}[TIT_c = 3] + \\ \beta_2 Treat_c Post_{ct} \mathbb{1}[TIT_c < 3] + \varepsilon_{ct} \end{aligned} \tag{5}$$

[18]In the Appendix, Figure C8 shows the distribution of average group productivity $\overline{\theta}_c$ by tercile of TI, which indicates that groups across the full support of $\overline{\theta}_c$ are at risk of being in any of the three terciles.

The results reported in Table 3 show small effects on average productivity overall (Panel A), but great heterogeneity by treatment intensity. In the simplest specification (column 1), the overall effect of any hire is about 4 percent of a standard deviation (SD). However, the effect is 17 percent of a SD and strongly significant for hires with the largest productivity gaps relative to their new peers. Results are very similar if one further controls for field-specific (column 2) or university-specific (column 3) time trends.

A causal interpretation of these results requires us to believe that the parallel trend assumption between treated and untreated groups holds. Figure 4 shows the dynamic equivalent of Equation 4. The event study suggests that trends are indeed comparable before the event.

However, to address concerns about selection into treatment and the timing of confounding shocks, and to further breakdown the drivers of the overall effect, I adopt a more robust empirical strategy described next.

Table 3—Effects on average group productivity—Treated *vs* Untreated comparison

	Dep. Var.: Average productivity (SDs)		
	(1)	(2)	(3)
Panel A: Overall effect			
$Treat \times Post$	0.044	0.044	0.033
	(0.012)	(0.012)	(0.012)
	$[0.020, 0.068]$	$[0.021, 0.068]$	$[0.008, 0.057]$
Panel B: Effect by treatment intensity			
$Treat \times Post \times \mathbb{1}[TIT = 3]$	0.178	0.173	0.162
	(0.021)	(0.020)	(0.021)
	$[0.137, 0.219]$	$[0.133, 0.213]$	$[0.121, 0.203]$
$Treat \times Post \times \mathbb{1}[TIT < 3]$	-0.028	-0.024	-0.037
	(0.015)	(0.014)	(0.015)
	$[-0.057, 0.001]$	$[-0.052, 0.004]$	$[-0.066, -0.008]$
N	149,198	149,198	149,198
Groups	8,054	8,054	8,054
Year FE	X		
Year x Field FE		X	
Year x University FE			X

Note: The outcome is average citation-weighted output for all faculty in the group, including the new eligible hire. This is the same as the outcome for the *Total Effect* discussed below. Standard errors are in parentheses. Bounds are 95 percent confidence interval extrema and reported in the square brackets. The control is composed by all untreated groups. Neither treatment nor control are restricted by university type (public or private) or event time.

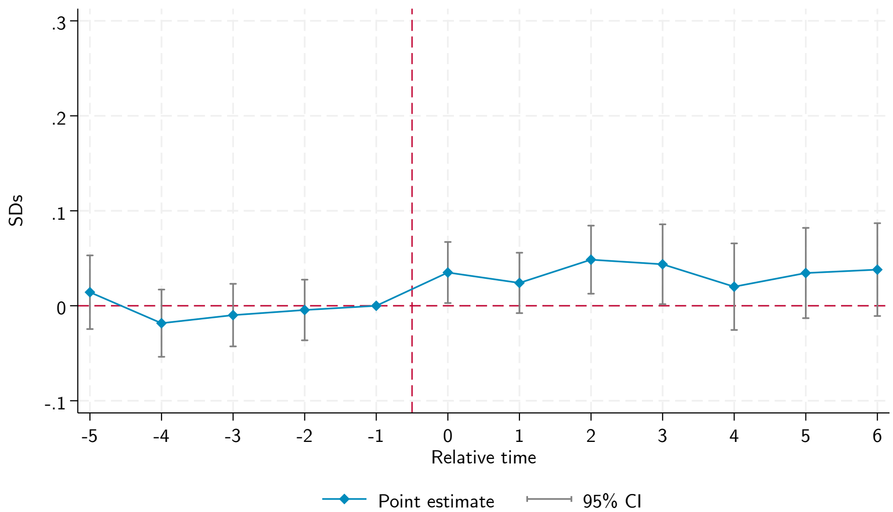


Figure 4. Effects on average group productivity—Treated *vs* Untreated comparison

Note: The outcome is average citation-weighted output for all faculty in the group, including the new eligible hire. This is the same as the outcome for the *Total Effect* discussed below. The control is composed by all untreated groups. Neither treatment nor control are restricted by university type (public or private) or event time.

B. *Main empirical strategy*

My main empirical strategy differs in two important ways from the simple comparison of treated *versus* untreated groups. First, I restrict the analysis to groups within public universities that see an eligible hire up to 2015. This restriction ensures we consider a time and population that was subject to demand constraints in the academic labor market, and could not easily manipulate the timing of a hire. Second, within ever-treated groups, I compare units in the top treatment intensity tercile to those in the bottom two terciles. Thus, I aim to estimate the productivity effects of eligible hires by contrasting changes following eligible hires with productivity above their group average, to changes in departments where $\theta_{h(c)}$ and $\overline{\theta}_c$ are close.

To do so, I estimate the following dynamic Difference-in-Differences model:

$$(6)\quad y_{ct} = \alpha_c + \varphi_t + X_{ct}'\Gamma + \sum_{m=-5,m\neq-1}^{6} \gamma_m E_{cm} Treat_c + \sum_{m=-5}^{6} \kappa_m E_{cm} + \bar{\gamma} + \bar{\kappa} + \varepsilon_{ct}$$

where $E_{cm} = \mathbb{1}[m = t - EventTime_c]$ is an indicator for group c being m periods away from the event time; t indexes calendar time; α_c are group FEs, φ_t are year FEs; $Treat_c = \mathbb{1}[TIT_c = 3]$ is an indicator for group c being in the top tercile of *TI*. The outcome y_{ct} is either average productivity, average productivity excluding eligible hires (leave-out average) or average individual *permanent* productivity (again excluding eligible hires). Respectively, these outcomes measure the *Total*, *Indirect* and *Composition Effects*. X_{ct} is a more flexible calendar time trend by some group characteristic. It is empty in the main results, but varied for robustness in the Appendix.[19] $\bar{\gamma}$ and $\bar{\kappa}$ are relative time bins for observations

[19] In the Appendix, Section D.1 shows results for X_{ct} including year by field FE, or interacted with terciles of: pre-reform funding per person, pre-period extensive margin in citations, pre-period average permanent productivity.

before $m = -5$ and after $m = 6$. Regressions are weighted by average group size in the preceding five periods to limit the influence of outliers. Error terms are clustered at the group level. γ_m are the coefficients of interest, capturing the effect of a higher productivity hire.

To capture average post-period effects, I also consider the static version of Equation 6:

$$(7) \qquad y_{ct} = \alpha_c + \varphi_t + X'_{ct}\Gamma + \gamma Treat_c Post_{ct} + \sum_{m=-5}^{6} \kappa_m E_{cm} + \bar{\gamma} + \bar{\kappa} + \varepsilon_{ct}$$

where $Post_{ct} = \mathbb{1}[t \geq EventTime_c]$.

Causal interpretation of my results relies on three key assumptions. First, parallel trends in baseline outcomes. Second, no anticipatory behavior prior to treatment. Third, treatment effect homogeneity.[20] My analysis meets, or properly relaxes, all these requirements. In the context of research production, it is improbable to see output responses before inputs change, mitigating concerns about pre-treatment effects. For peer effects more generally, anticipatory effects also seem unlikely, as direct interactions are typically crucial for any influence to manifest. Nonetheless, the insignificance of pre-trends tests provides reassurance that subjects are not preemptively adjusting to future hires. Treatment effect heterogeneity is a known challenge, with extensive literature indicating that standard Two-Way Fixed Effects (TWFE) models can incur significant bias when this assumption is violated (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; de Chaisemartin and D'Haultfœuille, 2022; Borusyak, Jaravel and Spiess, 2023). Since it most closely matches my application, I use Sun and Abraham (2021)'s unbiased estimator throughout my analysis, thus allowing for heterogeneity by treatment cohort.

Arguably, the parallel trends assumption is more credible when comparing groups that all select into making an eligible hire, but differ in the size of the

[20]See Sun and Abraham (2021) for a formal review of standard event study assumptions.

permanent productivity gap between hire and peers. First, by comparing ever-treated units, I account for selection effects that could otherwise invalidate parallel trends. Second, comparing events that happen in the same relative time, but differ in their TI, differences out any effects of hiring *per se* and captures the difference induced by the higher relative productivity of hires in the treatment group.[21]

A remaining threat to identification are contemporaneous shocks that determine an eligible hire and affect peer productivity—at the same time and differentially for top treatment intensity events. For instance, one might worry that a new department chair simultaneously makes top hires and boosts productivity.

However, the institutional context at hand greatly mitigates these concerns. First, Italian public universities between 2003 and 2015 were subject to a plethora of hiring constraints, from budget cuts to turnover rules and hiring freezes, which capped demand and significantly reduced a department's ability to perfectly manipulate the timing of new hires. Therefore, I restrict my main analysis to this period. Section A.3 in the Appendix details these constraints with selected legislative references. Second, Italian public universities are governed by perfectly homogeneous national laws and regulations, set pay scales and nation-wide competitive examinations for qualifications and promotions—all of which severely limits the extent to which individual department heads or deans can affect scientific production differentially. Third, typical university statutes provide for very collegial management of departmental activities, with chairs mainly tasked with coordination and representing the department in broader university processes.[22]

Nevertheless, I turn to my secondary strategy for an empirical design that controls for department-level simultaneous shocks.

[21] For instance, any new hire may have other effects like disruption or correlate with temporary changes in productivity following a departure of incumbent faculty. We want to net out these other shocks and only identify the effect of a greater productivity hire.

[22] For instance, see articles 26 and 27 from the statute of Bari State University: `https://www.uniba.it/it/ateneo/bollettino-ufficiale/Statuto%20Universita%20degli%20Studi%20di%20Bari%20Aldo%20Moro%20-%20DR%203235-2021.pdf`. Working link as of July 2024.

C. Secondary empirical strategy

In my secondary strategy, I focus on top TIT events and compare treated groups to the broader departments to which they belong, in order to net out any department-level simultaneous shocks. Namely, I contrast the change in productivity in the specific academic fields affected by new hires to the change in their departments at large, conditional on the event being in the top tercile of TI. Recall that in my data 'groups' correspond to granular scholarly fields and 'departments' to coarser disciplinary areas. In a university there can be as many as 190 'groups' across 14 'departments'.

Thus, for this analysis event time is defined at the department level, with treatment units being the specific groups that receive the eligible hire and control units being other groups in the same department that do not see such a hire. Groups and department that receive other types of hires and excluded. The previous restriction to public universities treated before 2015 is kept.

To operationalize this alternative approach for group-level outcomes, I estimate the following Difference-in-Differences model:

$$y_{cdt} = \alpha_c + \varphi_t + X'_{ct}\Gamma + \gamma Post_{dt}Treat_c + \sum_{m=-5}^{6} \kappa_m E_{dm} + \bar{\gamma} + \bar{\kappa} + \varepsilon_{cdt} \tag{8}$$

which is analogous to Equation 7m except for the different sample definition described above and event time now being defined at the department level as indexed by d. Once again, γ is the coefficient of interest.

Notice that, while addressing concerns about simultaneous shocks like changing department chairs, this approach also has drawbacks. First, it restricts the range of events considered as it now defines event time based on the first eligible hire at the department level. Because several departments experience distinct events in different groups and years, this reduces the number of events and skews the sample to earlier cohorts. Second, it compares treated and untreated groups within the same department, which may raise selection concerns and the possibility of confounding spillovers. Nonetheless, consistent findings across both approaches bolster the credibility of our inferences despite their individual limitations.

D. Individual-level spillovers

Finally, I adapt my main empirical strategy to the level of individual scholars to estimate individual-level productivity effects. I focus on own productivity, fix incumbents affected at $t = 0$ by the new hire and control for their individual FE. Thus, these results can be interpreted as *Spillover Effects*—the individual-level response of a new hire's peers.

The specification is analogous to 6, but also includes person FEs. Then:

$$y_{it} = \zeta_i + \alpha_c + \varphi_t + X_{ct}'\Gamma + \sum_{m=-5,m\neq-1}^{6} \gamma_m E_{im} Treat_{ic(0)} + \sum_{m=-5}^{6} \kappa_m E_{im} + \bar{\gamma} + \bar{\kappa} + \varepsilon_{it} \tag{9}$$

where y_{it} is individual productivity and $Treat_{ic(0)} = \mathbb{1}[TIT_{ic(0)} = 3]$ is an indicator for the group in which i is in $m = 0$ being in the top tercile of treatment intensity. ζ_i are individual FEs, α_c are current-group FEs (individuals may move). Taking advantage of the increased sample size, X_{ct} now includes year by field FEs by default. It is varied in the Appendix for robustness.[23] Errors are clustered at the individual level. γ_m are the coefficients of interest, capturing the effect of a higher productivity hire.

This equation is similarly adapted to a static setting as in 7 to estimate average post-period effects.

[23]In the Appendix, Section D.3 shows results for X_{ct} being omitted and when including flexible time trends by terciles of: pre-period extensive margin in citations, pre-period average permanent productivity.

VI. Group-level results

This section presents group-level event study results. First, I detail the findings from the primary empirical approach focused on the impacts of top hires. Next, I investigate effect heterogeneity in the main result. Lastly, I present estimates from the secondary strategy.

A. Main strategy

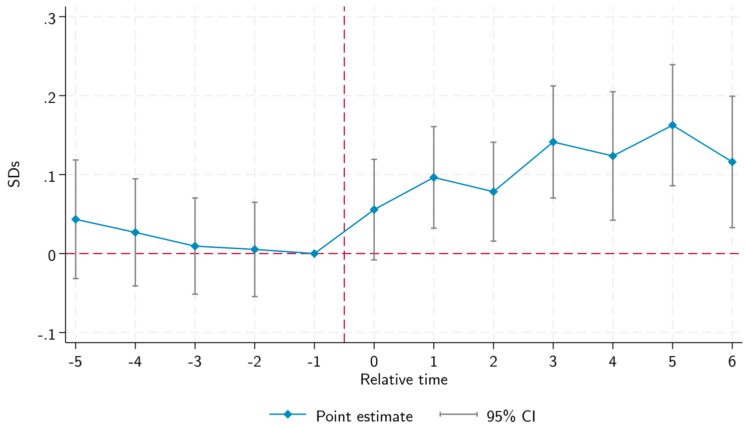
The results in Figure 5 and Table 4 reveal substantial productivity gains from top eligible hires. The *Total Effect* is approximately 0.14 standard deviations (SDs) after three years—the length of a typical pre-tenure contract—and 0.10 SDs on average in the post-period. The effect peaks after five years. The *Indirect Effect*—enhanced productivity of other scholars in the affected group—is around 0.08 SDs after three years. However, because it builds up more gradually and is less precisely estimated then the other effects, the *Indirect Effect* post-period average of 0.04 SDs that is marginally insignificant.

A large fraction of the estimated increase in group-level scientific productivity is explained by a a notable increase in the average permanent productivity of non-eligible faculty (*Composition Effect*). This effect is more precisely estimated because it captures individuals' lifetime permanent productivity and is therefore less noisy. After three years, I estimate a *Composition Effect* of 0.05 SDs, about two-thirds of the overall *Indirect Effect*, indicating that higher-productivity local researchers sort into groups with top eligible hires. Average post-period magnitudes suggest that talent reallocation, both of eligible hires moving from abroad and local researchers, rather than individual-level spillovers explain overall group-level productivity gains.

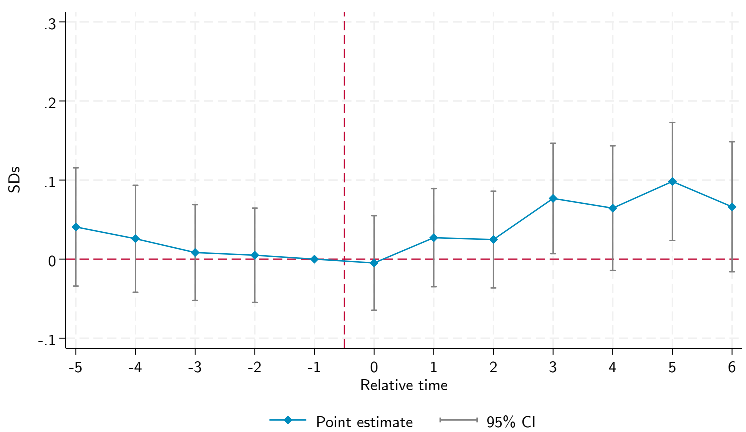
Table 4—The effects of top eligible hires on group productivity—Summary

Effect:	Total	Indirect	Composition
	(1)	(2)	(3)
Panel A: Average post-treatment effect			
$Treat \times Post$	0.095	0.036	0.043
	(0.024)	(0.023)	(0.011)
	$[0.049, 0.142]$	$[-0.008, 0.081]$	$[0.021, 0.065]$
Panel B: Effect after 3 years			
$Treat \times \mathbb{1}[m = 3]$	0.141	0.077	0.052
	(0.036)	(0.036)	(0.015)
	$[0.070, 0.212]$	$[0.007, 0.147]$	$[0.023, 0.080]$
N	10,482	10,482	10,477
Groups	515	515	515

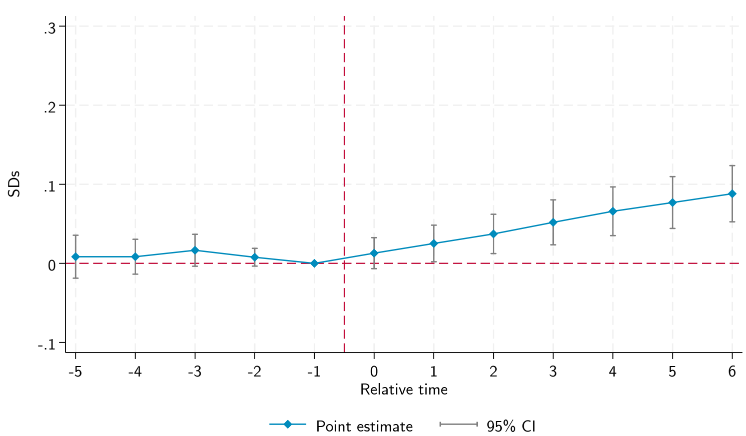
Note: The outcome for *Total Effect* is average citation-weighted output for all faculty in the group, including the new eligible hire. The outcome for *Indirect Effect* is average citation-weighted output for all faculty, except eligible hires. The outcome for the *Composition Effect* is average permanent productivity (FE from equation 2) for all non-eligible faculty in the group. Standard errors are in parentheses. Bounds are 95 percent confidence interval extrema and reported in the square brackets.



(a) Total effect



(b) Indirect effect



(c) Composition effect

FIGURE 5. THE IMPACT OF TOP ELIGIBLE HIRES ON GROUP PRODUCTIVITY

Note: The outcome in Panel (a) is average citation-weighted output for all faculty in the group, including the new eligible hire. The outcome in Panel (b) is average citation-weighted output for all faculty, except eligible hires. The outcome in Panel (c) is average permanent productivity (FE from equation 2) for all non-eligible faculty in the group.

B. Heterogeneity

To further explore which cuts of the data are contributing to positive indirect effects of eligible hires, I categorize groups based on terciles of key characteristics, measured strictly before the event date. These characteristics include: (*i*) the proportion of faculty with any cited output in $t = -1$; (*ii*) $\overline{\theta}_c$, average permanent productivity of incumbents in $t = -1$; (*iii*) group size in the preceding five years; (*iv*) funding levels per person from 2001-2003; (*v*) international mobility time per person from 2001-2003; (*vi*) government-assessed research quality score from 2001-2003. Values for variables (*iv*) to (*vi*) are imputed from department-level reports from the 2001-2003 VTR.

Then, I estimate the static Difference-in-Differences equivalent of Equation 6, adding indicators for each tercile of a given characteristic H. The corresponding regression equation is:

$$(10)\quad y_{ct} = \alpha_c + \varphi_{ft} + X'_{ct}\Gamma + \sum_{h=1}^{3} \gamma_h Post_{ct} \mathbb{1}[TIT_c = 3]\mathbb{1}[HT_c = h] + \bar{\gamma} + \bar{\kappa} + \varepsilon_{ct}$$

where $\mathbb{1}[HT_c = h]$ is an indicator for group c being in tercile h of characteristic H.

Figure 6 reports the average post-period indirect effect across these group characteristics, γ_h. Several points are noteworthy. The overall effect is not driven by underperforming groups expected to improve regardless—as would be the case if results were driven by mean reversion. The effect is concentrated in groups that are in the middle of productivity distribution, as indicated by the extensive margin of citations and $\overline{\theta}_c$, and by those with intermediate levels of pre-2004 funding. Moreover, results are homogeneous across the size and research quality score (VTR) dimensions. Interestingly, the indirect effect is more pronounced in groups that were highly internationalized before 2004, suggesting that these departments were already well-positioned to disproportionately benefit from international collaboration and human capital.

The estimates, while not precise enough for detailed field comparisons, are also broken down by broad scientific areas in Appendix Figure D13. This shows more pronounced effects in the social sciences and humanities.

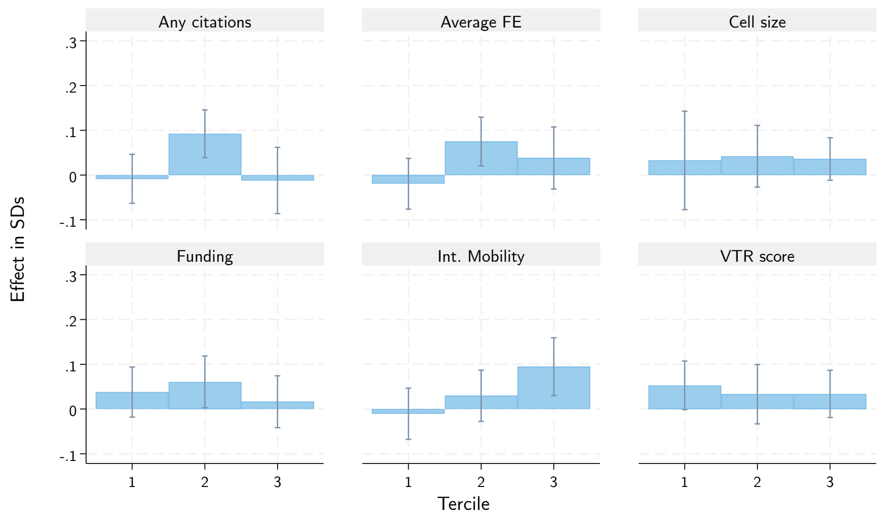


Figure 6. Indirect Effect—Heterogeneity by pre-period covariates

Note: The outcome for *Indirect Effect* is average citation-weighted output for all faculty, except eligible hires. From the left to right and from top to bottom, the sources of heterogeneity (by tercile) are: (*i*) the proportion of faculty with any cited output in $t = -1$; (*ii*) the average permanent productivity in the cell in $t = -1$, denoted as $\overline{\theta}_c$; (*iii*) group size in the preceding five years; (*iv*) funding levels per person from 2001-2003; (*v*) international mobility time per person from 2001-2003; (*vi*) government-assessed research quality score (VTR) from 2001-2003. Confidence intervals are at the 95% level.

C. *Secondary strategy*

My secondary strategy confirms the general takeaways of the main approach. Table 5 shows positive and significant *Total* and *Composition Effects*. Magnitudes are overall smaller, which can be attributed both to different sample restrictions and within-department spillovers across groups.

All in all, productivity is going up differentially in the treated groups, on top of any department-level shift.

Table 5—The effects of top eligible hires on group productivity—Alternative strategy

Effect:	Total (1)	Indirect (2)	Composition (3)
Panel A: Average post-treatment effect			
$Treat \times Post$	0.051	-0.006	0.029
	(0.020)	(0.019)	(0.010)
	$[0.011, 0.090]$	$[-0.043, 0.030]$	$[0.009, 0.048]$
Panel B: Effect after 3 years			
$Treat \times \mathbb{1}[m = 3]$	0.071	0.005	0.035
	(0.035)	(0.033)	(0.015)
	$[0.003, 0.138]$	$[-0.061, 0.070]$	$[0.007, 0.064]$
N	22,161	22,161	22,143
Groups	1,095	1,095	1,095

Note: The outcome for *Total Effect* is average citation-weighted output for all faculty in the group, including the new eligible hire. The outcome for *Indirect Effect* is average citation-weighted output for all faculty, except eligible hires. The outcome for the *Composition Effect* is average permanent productivity (FE from equation 2) for all non-eligible faculty in the group. Average post-period effects are obtained through the static Difference-in-Difference equivalent of specification (4). Standard errors are in parentheses. Bounds are 95 percent confidence interval extrema and reported in the square brackets.

VII. Individual-level results

This section presents individual-level event study results. First, I report results from the main empirical approach, focusing on the effects of top hires. Next, I test effect heterogeneity in the main result.

The findings in Figure 7 and Table 6 indicate a negligible and insignificant average *Spillover Effect*. The post-period average is almost exactly zero SDs (about one tenth of one percent of a SD), with fairly tight bounds around three percent of a SD. This null result extends to the sheer number of published output (i.e. unweighted by citations), the probability of that output ever garnering a positive citation count, and the probability that one's output is in the top decile of the citation distribution in a given year.

Mirroring group-level analysis, I test whether findings are heterogeneous across several cuts of the data. Results are reported in Figure 8. As before, heterogeneity analysis is conducted estimating the static difference-in-difference equivalent of Equation 9, adding indicators for each tercile of a given characteristic H to the diff-in-diff interaction. The resulting regression equation is then:

$$(11)\quad y_{it} = \alpha_i + \zeta_c + \varphi_t + X'_{ct}\Gamma + \sum_{h=1}^{3} \gamma_h Post_{it} Treat_{ic(0)} \mathbb{1}[HT_{ic(0)} = h] + \bar{\gamma} + \bar{\kappa} + \varepsilon_{it}$$

where $\mathbb{1}[HT_{ic(0)} = h]$ is an indicator for initial group c in which i is at time 0 being in tercile h of characteristic H.

The results are remarkably consistent in showing small and insignificant effects in each dimension of heterogeneity.

Figure 9 categorizes the data by broad scientific areas and displays the average post-period effect in each. The result shows that the pattern of small and insignificant average spillovers extends to all main disciplines.

Table 6—The spillover effect of top eligible hires on individuals—Summary

Dep. Var.:	Productivity	Output count	Any citations	Top decile citations
Unit:	SDs	SDs	Probability	Probability
	(1)	(2)	(3)	(4)
$Treat \times Post$	-0.001	0.001	-0.001	0.000
	(0.016)	(0.015)	(0.007)	(0.006)
	$[-0.032, 0.029]$	$[-0.029, 0.030]$	$[-0.015, 0.012]$	$[-0.011, 0.011]$
N	108,421	108,421	108,421	108,421
Individuals	6,456	6,456	6,456	6,456

Note: The outcome is own citation-weighted output. Standard errors are in parentheses. Bounds are 95 percent confidence interval extrema and reported in the square brackets.

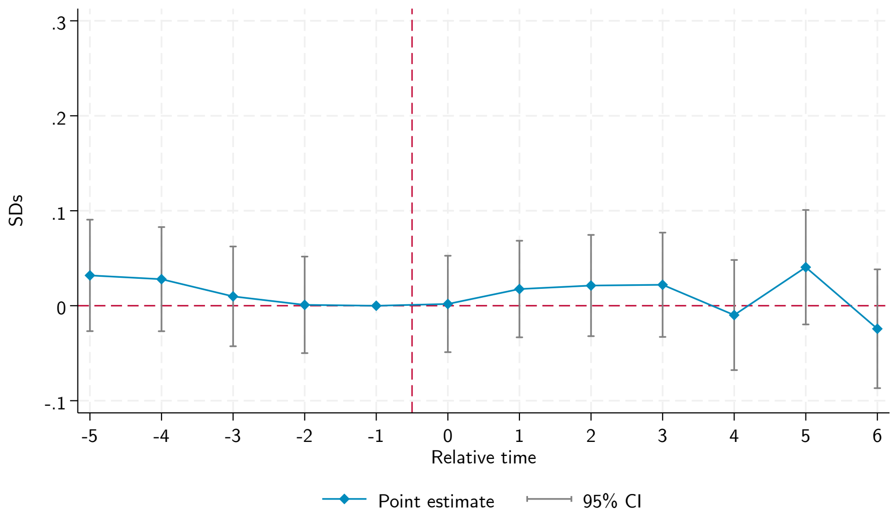


FIGURE 7. THE SPILLOVER EFFECT OF TOP ELIGIBLE HIRES ON INDIVIDUALS

Note: The outcome is own citation-weighted output.

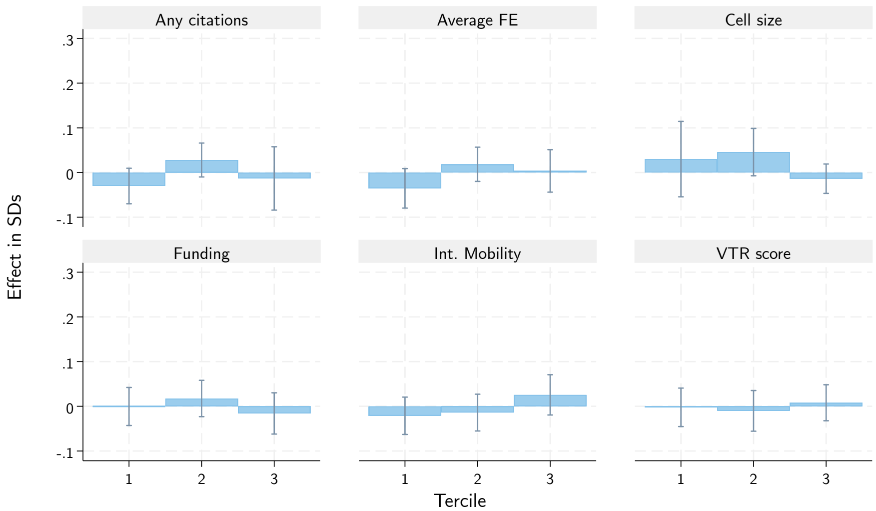


Figure 8. Spillover effect—Heterogeneity by pre-period covariates

Note: The outcome is own citation-weighted output. From the left to right and from top to bottom, the sources of heterogeneity (by tercile) are: (*i*) the proportion of faculty with any cited output in $t = -1$; (*ii*) the average permanent productivity in the cell in $t = -1$, denoted as $\overline{\theta}_c$; (*iii*) group size in the preceding five years; (*iv*) funding levels per person from 2001-2003; (*v*) international mobility time per person from 2001-2003; (*vi*) government-assessed research quality score (VTR) from 2001-2003. Confidence intervals are at the 95% level.

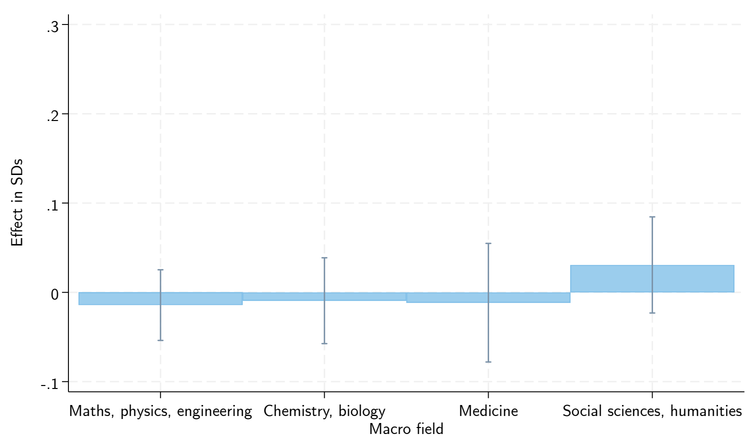


Figure 9. Spillover effect—Heterogeneity by broad scientific areas

Note: The outcome is own citation-weighted output. Confidence intervals are at the 95% level.

VIII. Conclusion

This study delves into the effectiveness of Italy's 'rientro dei cervelli' (returning brains) preferential tax scheme, which was designed to attract highly skilled researchers and university professors back to the country. By leveraging a comprehensive dataset encompassing 90 percent of Italian faculty from 2000 to 2020, the analysis provides robust evidence on both the migration and productivity impacts of the policy. The findings indicate that the tax incentive successfully increased the influx of eligible hires, elevating their share from a modest 2 percent to a substantial 12.5 percent of total hires. Crucially, these beneficiaries exhibit significantly higher productivity levels, with an average permanent productivity that surpasses their peers by approximately 43 percent.

Employing a dynamic Difference-in-Differences framework, I find that the arrival of top eligible hires leads to a notable increase in the average productivity of their respective academic groups by approximately 0.14 SDs. This enhancement is attributed to both the direct contributions of the new hires and indirect effects stemming mostly from the subsequent sorting of other high-productivity local researchers. However, these sizeable composition changes go hand in hand with negligible individual-level responses. At the individual level, the analysis demonstrates that spillover effects on incumbent researchers are minimal and statistically insignificant, with estimated changes not exceeding 0.03 SDs in either direction.

All in all, the results suggest that while the tax incentive effectively reallocates talent, it does not foster substantial improvements in the productivity of existing faculty members. Consequently, the primary benefits of the policy are the new hires' direct contributions within the departments that successfully attract highly productive researchers, without generating widespread productivity gains.

From a policy perspective, a mixed picture of the effectiveness of the 'returning brains' scheme emerges. On one hand, it effectively mitigates brain drain by attracting a significant number of highly skilled researchers. On the other hand, the policy entails substantial fiscal costs, introduces distortions in the international academic labor market and fosters tax competition across jurisdictions. Moreover, the limited evidence of individual-level spillovers indicates that the broader benefits of the policy are not nearly as sizeable as the direct gains experienced by the treated departments.

REFERENCES

Agrawal, David R., and Dirk Foremny. 2019. "Relocation of the Rich: Migration in Response to Top Tax Rate Changes from Spanish Reforms." *The Review of Economics and Statistics*, 101(2): 214–232.

Aihounton, Ghislain B D, and Arne Henningsen. 2021. "Units of measurement and the inverse hyperbolic sine transformation." *The Econometrics Journal*, 24(2): 334–351.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva. 2016. "Taxation and the International Mobility of Inventors." *The American Economic Review*, 106(10): 2930–2981. Publisher: American Economic Association.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2022. "Immigration and Entrepreneurship in the United States." *American Economic Review: Insights*, 4(1): 71–88.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *The Quarterly Journal of Economics*, 125(2): 549–589. Publisher: Oxford University Press.

Bassetto, Jacopo, and Giuseppe Ippedico. 2023. "Can Tax Incentives Bring Brains Back? Returnees Tax Schemes and High-Skilled Migration in Italy | Publications | CESifo." *CESifo Working Paper No. 10271.*

Bernstein, Shai, Rebecca Diamond, Abhisit Jiranaphawiboon, Timothy McQuade, and Beatriz Pousada. 2022. "The Contribution of High-Skilled Immigrants to Innovation in the United States."

Borjas, George J., and Kirk B. Doran. 2012. "The Collapse of the Soviet Union and the Productivity of American Mathematicians*." *The Quarterly Journal of Economics*, 127(3): 1143–1203.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2023. "Revisiting Event Study Designs: Robust and Efficient Estimation." arXiv:2108.12419 [econ].

Burchardi, Konrad B., Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J. Terry. 2020. "Immigration, Innovation, and Growth." *NBER Working Paper No. 27075.*

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

Chen, Jiafeng, and Jonathan Roth. 2023. "Logs with Zeros? Some Problems and Solutions*." *The Quarterly Journal of Economics.*

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg. 2017. "Peer Effects in the Workplace." *The American Economic Review*, 107(2): 425–456. Publisher: American Economic Association.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2022. "Difference-in-Differences Estimators of Intertemporal Treatment Effects." *NBER Working Paper No. 29873.*

Doran, Kirk, Alexander Gelber, and Adam Isen. 2022. "The Effects of High-Skilled Immigration Policy on Firms: Evidence from Visa Lotteries." *Journal of Political Economy*, 130(10): 2501–2533. Publisher: The University of Chicago Press.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics*, 2(2): 31–56.

Kerr, Sari Pekkala, William Kerr, Çağlar Özden, and Christopher Parsons. 2016. "Global Talent Flows." *Journal of Economic Perspectives*, 30(4): 83–106.

Kerr, William R., and William F. Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention." *Journal of Labor Economics*, 28(3): 473–508. Publisher: The University of Chicago Press.

Kleven, Henrik, Camille Landais, Mathilde Muñoz, and Stefanie Stantcheva. 2020. "Taxation and Migration: Evidence and Policy Implications." *Journal of Economic Perspectives*, 34(2): 119–142.

Kleven, Henrik Jacobsen, Camille Landais, and Emmanuel Saez. 2013. "Taxation and International Migration of Superstars: Evidence from the European Football Market." *American Economic Review*, 103(5): 1892–1924.

Kleven, Henrik Jacobsen, Camille Landais, Emmanuel Saez, and Esben Schultz. 2014. "Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark *." *The Quarterly Journal of Economics*, 129(1): 333–378.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1): 112–145.

Moretti, Enrico, and Daniel J. Wilson. 2017. "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists." *American Economic Review*, 107(7): 1858–1903.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish Émigrés and US Invention." *American Economic Review*, 104(10): 3222–3255.

Moser, Petra, Sahar Parsa, and Shmuel San. 2020. "Immigration, Science, and Invention. Evidence from the Quota Acts." *Working paper.*

OECD. 2017. "Scientists on the move." OECD, Paris.

Prato, Marta. 2025. "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth*." *The Quarterly Journal of Economics*, 140(1): 165–238.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Timm, Lisa Marie, Massimo Giuliodori, and Paul Muller. 2025. "Tax Incentives for Migrants With Mid-Level Earnings: Evidence From the Netherlands." *American Economic Journal: Applied Economics.*

Waldinger, Fabian. 2012. "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany." *The Review of Economic Studies*, 79(2): 838–861. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

‘Returning brains’: tax incentives, migration and scientific productivity

Online Appendix

Pier Paolo Creanza*

June 13, 2025

Contents

*Princeton University, pcreanza@princeton.edu.

A Policy details

A.1 Legal framework

Legal reference	Year of approval	Fiscal years affected	Content	Benefit duration	Additional benefits
DL 269/2003 art. 3	2003	2004-2008	First introduces the 90% reduction in taxable income for researchers who move from abroad	3 years	None
DL 185/2008 art. 17	2008	2009-2013	Renews provision until 2013	3 years	None
DL 78/2010 art. 44	2010	2014-2015	Renews provision until 2015	3 years	None
Legge 190/2014 art. 1	2014	2015-2017	Renews provision until 2017, extends benefit duration	4 years	None
Legge 232/2016 art. 1	2016	2018-today	Makes tax scheme permanent	4 years	None
Legge 34/2019 art. 5	2019	2020-today	Extends benefit duration, introduces several additional cause for benefit extension	6 years	2 extra years if beneficiaries with one underage child buy a house in Italy; 5 years with 2 children; 7 years with 3 or more children.

Table A1: Legal references for 'returning brains' tax scheme

A.2 Policy benefits and revenue costs

Table A2 reports examples of tax benefits reaped by the people targeted by the reform. To populate the table, I consider 2007-2021 average and marginal tax rates, and the original 3-year duration of the benefit. Incomes are in 2011 EURs. Rows 6-7 are representative of the 2011 salary range for a *Ricercatore* in a public university, while rows 4-5 and 3-2 are the same for Associate and Full professors respectively. The top row considers a higher salary level plausible for private universities. Columns **Net income**, **ATR** (average tax rate) and **MTR** (marginal tax rate) describe the circumstances faced by an ineligible scholar. The last column, **Beneficiaries' taxes**, illustrates the fiscal liability for scholars who qualify for the tax break.

Table A2 shows how stark the incentives introduced by the tax scheme are. Anyone making less than 80 thousand euros a year pays no income tax at all, and this applies to virtually anyone employed in a public university. Even someone making as much as 120 thousand euros a year still faces an average tax rate around 1%. As a direct consequence of the general progressivity of Italian income taxation, the benefits induced by the reform are highly 'regressive'—that is, more senior scholars are given much sharper incentives to move and take up the scheme.

Row	Gross income	Net income	ATR	MTR	Beneficiaries' taxes
1	120,000	75,500	37%	43%	1,200
2	80,000	52,000	35%	43%	0
3	62,500	42,900	31%	41%	0
4	50,500	35,100	30%	38%	0
5	39,500	28,600	28%	38%	0
6	29,000	22,100	24%	38%	0
7	20,500	16,900	18%	27%	0

Table A2: Illustration of the fiscal incentive

As for estimating revenue cost, consider Table A3. The first column is constructed from my simulation of eligibility based on career data and accounts for the several-year duration of the benefit. The second column is populated from official statistics. Those for 2017-2019 are publicly available. Data for 2012-2016 were obtained through a formal information request to the Ministry of Economy and Finance (MEF - Dipatimento delle Finanze). The third column column uses the 2012-2019 ratio of actual to estimated beneficiaries to project 2004-2011 totals using trends in eligible faculty. The last column is populated using standard personal income tax rates (IRPEF) and using the average gross income reported in official statistics when available, and the 2012-2019 average of around €120,000 for 2004-2011.

Year	Estimated eligible faculty	Actual total beneficiaries	Projected total beneficiaries	Projected foregone **revenue** (mi EUR)
	(1)	(2)	(3)	(4)
2004	67	-	329	14.3
2005	171	-	841	36.6
2006	277	-	1,362	59.2
2007	272	-	1,337	58.2
2008	259	-	1,273	55.4
2009	200	-	983	42.8
2010	222	-	1,091	47.5
2011	228	-	1,121	48.8
2012	249	1,235	-	53.7
2013	259	1,558	-	67.8
2014	252	2,298	-	100.0
2015	288	2,038	-	88.7
2016	380	1,571	-	68.3
2017	462	1,624	-	70.6
2018	677	1,646	-	71.6
2019	851	1,765	-	76.8
			Total projected cost:	960.1

Table A3: Estimated revenue cost

A.3 Hiring limitations

The sheer volume of the provisions taken by the Italian government to the effect of severely limiting the ability of public universities to hire faculty (and staff) makes a comprehensive review a very challenging task. In the end, while making sense of the these legal changes is hard, one can discern how binding these limitations were by looking at the levels of new hires and total faculty in public universities, as depicted in Figure A1. The impact of hiring restrictions is manifest between 2003 and 2015.

Yet, it is useful to have at least a sense of the type of policies that were in place between 2003 and 2015, especially as laws and ministry regulations changed hiring limits and turnover rules sometimes more than once in the same year. Below is a selection of the main provisions affecting our period of interest.

- Legge 289/2002. Set the 2003 government budget. It blocked hiring in all public administrations, not just universities, and allowed for hardly any exceptions. It also set the objective of personnel reduction of about 1 percent in both 2004 and 2005.

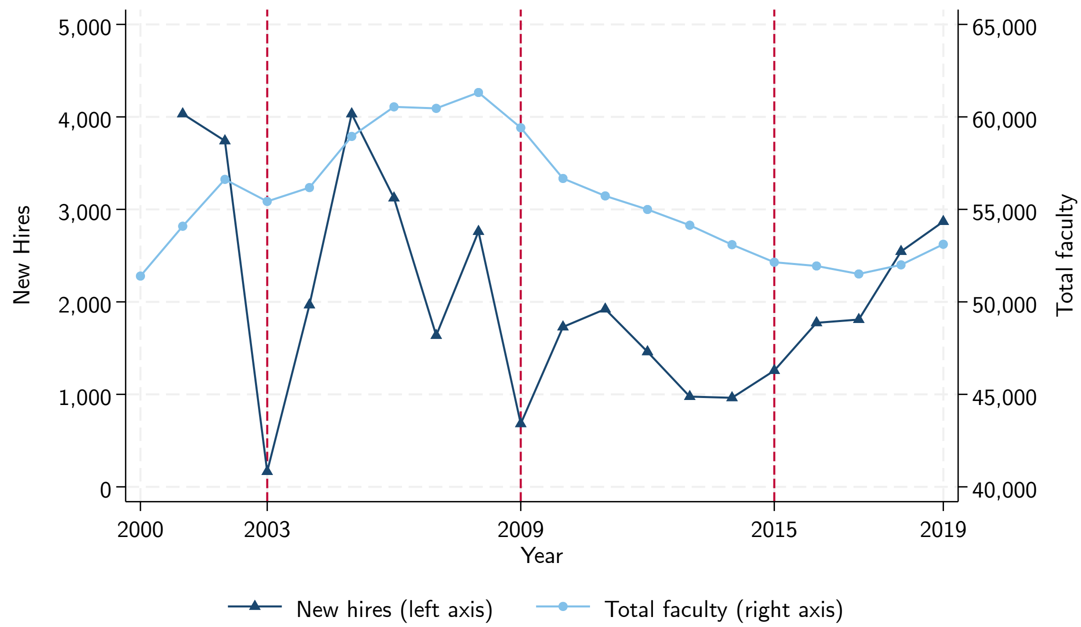


Figure A1: Stock and flow of Italian faculty.
Notes: Data are from the Faculty Rolls describd in the paper and include all Italian public universities; vertical red lines highlight the hiring freezes of 2003 and 2009 and the loosening of hiring restrictions started around 2015.

- Legge 350/2003. Set the 2004 government budget. While universities are not formally prohibited from recruiting, they can only hire individuals who have won public competitions by the end of 2003 and have not yet entered public employment, and only in fixed spending limits. Universities also cannot launch new public competitions.

- Legge 311/2004. Set the 2005 government budget. Confirms and extends budget cuts and personnel reduction of about 1 percent over the next two years.

- Legge 296/2006. Set the 2007 government budget. Introduced a university-level turnover rule limiting hires to 20% of terminations/retirements in the preceding year, both in terms of expenditure and of people.

- Legge 133/2008. Extended the 20% turnover rule for years 2009 to 2011. Mandated a 50% turnover rule for 2012.

- Legge 1/2009. Prohibited hiring for schools where personal expenditures were above 90% of ordinary government funding (FFO). Raised the turnover rule to 50% for 2009 to 2011.

- Legge 122/2010. Changed the turnover rule again to 20% until 2013 and 50% from 2014.

- Decreto-legge 216/2011. Changed the turnover rule to 50%.
- Decreto legislativo 49/2012. Introduced a more complicated categorization of universities in classes of financial health. The lowest class faced a turnover rule of 10%, the middle one of 20% and the top class of 20% plus an additional amount determined by a function of schools' balance sheets. It further added constraints on the relative share of full, associate and assistant professor and minimal requirements of outside hires.
- Decreto-legge 95/2012. Introduced a national turnover rule of 20% from 2012 to 2014, to become 50% in 2015 and 100% from 2016 (de facto capping expenditures on faculty at 2016 level thereon out).
- Decreto-legge 69/2013. Changed national turnover rules to 20% till 2013, 50% from 2014 to 2015 and 100% thereafter.

For more information see the following internet resources:

- `https://www.roars.it/breve-storia-della-normativa-dei-blocchi%-parziali-del-turnover-universitario/`
- `https://www.mur.gov.it/it/aree-tematiche/universita/programmazione%-e-finanziamenti/facolta-assunzionali`

B Data

B.1 Manual data collection on careers

In the following I report the algorithms used to find and code information on the careers for new entrants between 2001 and 2019.

B.1.1 Algorithms

Data finding algorithm

1. Search online for the NAME SURNAME UNIVERSITY (of entry, from the faculty rolls) of the individual and the keyword "CV".
2. Look for a CV. It does not need to be the most recent, but it should postdate the year of entry or at least be contemporaneous to it.
   - if successful, enter the data and proceed to the next person;
   - if unsuccessful OR if the CV does not contain all the information required, proceed to the next step.
3. Among the search results from Step 1, look for a personal page. These are likely to be on a personal URL (e.g. `www.namesurname.com` or on Google Sites, like `sites.google.com/view/namesurname`) or on a university website (e.g. `www.unito.it/faculty/namesurname`). They may also be on laboratory or project specific websites (e.g. `www.behavioral_lab.uniba.edu/fellows`) One can generally find a CV or a short bio on these pages.
   - if successful, enter the data and proceed to the next person;
   - if unsuccessful OR if the source found does not contain all the information required, proceed to the next step.
4. Look for a LinkedIn page. If a LinkedIn page exists, it will probably be among the results of your search in Step 1. Otherwise, make direct search for it.
   - if successful, enter the data and proceed to the next person;
   - if unsuccessful OR if the source found does not contain all the information required, proceed to the next step.
5. Among the search results from Step 1, look for publication indexing websites, like Google Scholar, ORCID, Academia.edu, ResearchGate.net, IEEE Xplore, and others. Often these have either a short bio or a list of affiliations. Both can be used to determine if someone has worked abroad. E.g. "Professor Car earned his PhD in Milan where he started as assistant professor as well. From 2010 he is associate professor at Bari's Polytechnic University", thus Car then can be imputed as not having worked abroad.
   - if successful, enter the data and proceed to the next person;
   - if unsuccessful OR if the source found does not contain all the information required, proceed to the next step.

6. Among the search results from Step 1, look for: (a) biographies from conferences, articles or blogposts; (b) biographies from published books (e.g. they have a book on IBS.it, click on the book page, check author information), (c) short biographies from hiring announcements and press (e.g. "Bocconi University is happy to announced that Dr XYZ will join the faculty of accounting beginning September 2011. Dr XYZ has recently earned her PhD from Seoul Private University, Korea.").
   - if successful, enter the data and proceed to the next person;
   - if unsuccessful OR if the source found does not contain all the information required, proceed to the next step.
7. Make a new search online using NAME SURNAME FIELD (e.g. physics, in our list), without the keyword "CV". In most cases, search results will be similar to the search done in Step 1.
   - if the search results contain novelties (e.g. pages not checked before), check these new sources according to the order described in Steps 2-6 (i.e. first CV, then personal pages, then LinkedIn, etc).
   - if the search results are the same or the new sources are irrelevant (e.g. newspaper articles about people with the same name), end the search for this person and proceed to the next one.

Data entry algorithm

1. Has the person worked (or studied for a PhD) in any country that is not Italy before their year of entry?
   - if no: insert "worked abroad = no" and go the next person. Leave all other variables blank.
   - if yes: proceed and only consider employment before the year of entry

   Note: if people have multiple positions that overlap, I do not count the years when they also have an Italian position: I assume they were in Italy.
2. When abroad, were they employed in academic or research positions? Note: academic/research positions include professors (associate, full, assistant, adjunct, emeritus, distinguished, centennial, etc), researchers ('assegnista di ricerca', research assistant, contract researcher, research unit / centro studi, etc.), PhD students (dottorandi, dottori di ricerca), post-doctoral fellows, research fellows, scientists. Academic jobs do not include managers, consultants, journalists, analysts, high school teachers, Master/Bachelor students. They also do not include summer schools, teachers programs, short-term/temporary teaching jobs.
   - if no: insert "worked abroad = no" and go the next person. Leave all other variables blank.
   - if yes: proceed
3. Ignore temporary positions like 'visiting', 'guest', 'exchange', 'invited' periods and all short-term jobs (e.g. teaching a 2-week summer school).
   - if you find these jobs, exclude them, and if this was the only experience abroad, insert "worked abroad = no". Go the next person and leave all other variables blank.

- if there are no such jobs, assume they were in permanent and full-time jobs and insert "worked abroad = yes". Proceed.

4. Consider now only the last period spent outside Italy before their "year of entry". Note 1: these people may have worked abroad more than once (before the year of entry). We only care about their last period abroad before their arrival in Italy. Note 2: a 'period abroad' may comprise different jobs/positions. For now, consider the whole period and not individual jobs.

5. How long has this period abroad lasted? Insert "number of years worked abroad":

    - if duration less than 2 years, even 23 months, insert the closest half-year duration less than 2 (e.g. 0.5, 1 or 1.5 years)
    - if duration is more than 2 years but less than 10, also approximate to half-years
    - if duration is more than 10 years, insert "10+"

    Note 1: it is critical for us to know if they were abroad for at least 2 years, so focus your attention to this. Note 2: for PhD students, if the duration is not specified, assume it was 4 years.

6. Have they moved directly from abroad to the entry job in our list?

    - if yes: insert "years passed since working abroad = 0"
    - if no: insert the difference (in half-years) between the end of their academic job abroad and the start of their Italian one.
      again if no: check After moving to Italy from abroad, have they worked at in research of university before the year of entry?
        - if yes: insert "other Italian job = yes"
        - if no: insert "other Italian job = no"

      Note: a non-zero difference is possible if people moved to Italy and did a non-academic job before entering the Italian university in our list. This is probably rare. Otherwise, it can be that people had a research job that is not in our list.

7. Do we have information on their last job affiliation abroad before moving to Italy?

    - if yes: insert the appropriate values (even if not all available) for "academic rank / institution / country" before moving
    - if no: leave unavailable information blank

B.1.2 Quality checks

Figure B2 shows the quality checks performed in the collection of the hires data.

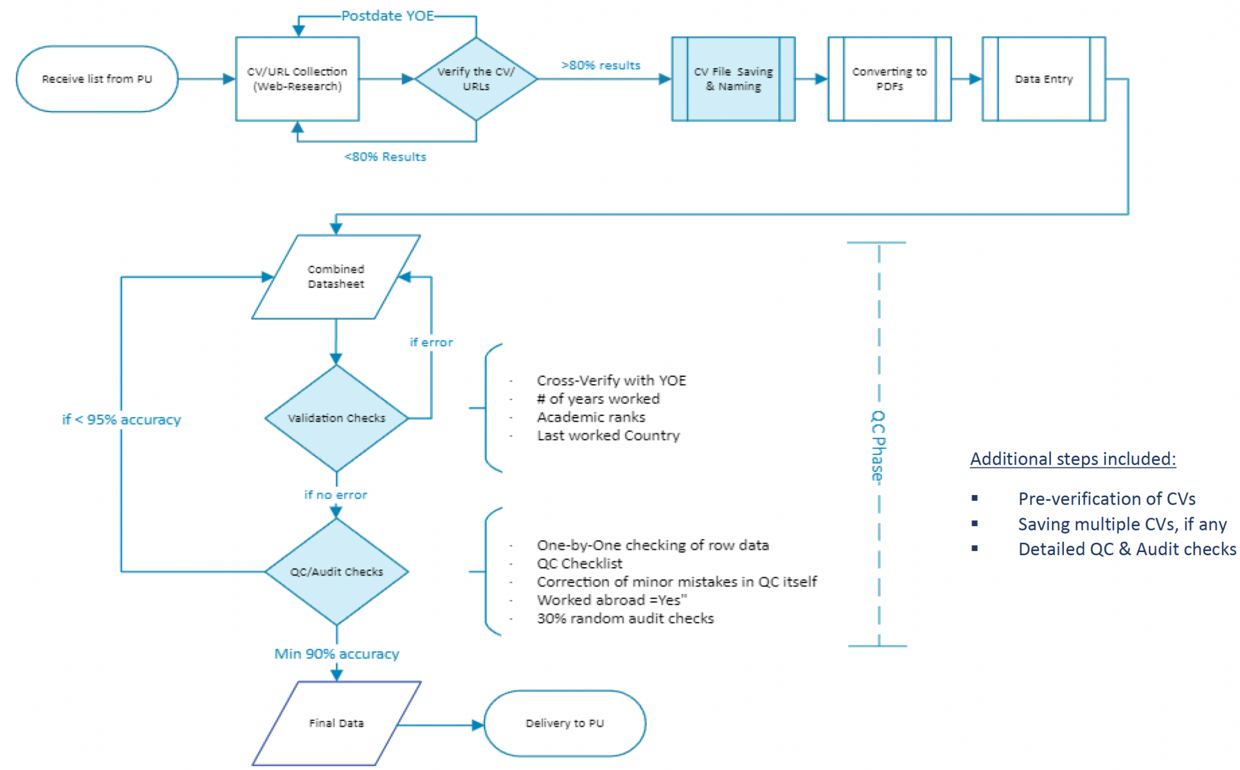


Figure B2: Data collection workflow

B.2 Scientific output data collection

I collect scientific output data from the online Institutional Repositories (IRs) maintained by Italian universities. These repositories report research output from current and ceased faculty. Their coverage its extensive as faculty and researchers are encouraged to document not only journal publications or books, but a wide range of other activity as well—including conference material, book chapters, working papers, editorships, patents, theses, etc. On the other hand, information is not consistently reported across IRs; even within an IR, not all entries will have the same level of details reported in their metadata.

However, IRs consistently report the key information I need to construct my faculty-output panel: authorship, title, type of contribution and citation counts from external indexing services. There are a handful of universities without citation counts and they're excluded from the analysis. The counts shown on the page are totals as of the time of data collection (summer 2022), and thus reflect cumulative citations up until that point.

IRs are populated by faculty and researchers individually. One might worry that coverage is thus biased because the types of faculty who upload their material (and its accuracy) may select in particular ways that correlate with productivity or tax-break eligibility. However, the Italian government uses material uploaded on IRs for periodic assessments of research quality (*Valutazione della Qualità della Ricerca*, or VQR), whose outcomes are used to allocate a large fraction of public funding to universities—public and private alike. Thus, each school has strong incentives to represent its output as completely as possible, and anecdotal evidence indeed indicates that faculty is actively required to populate their respective IR page. Linking records to external indexing services reduces the extent to which one might try to overstate their scientific productivity.

Table B4 reports faculty employment, availability of an IR and total record counts in the available IR for all Italian universities.

University	Faculty in 2019	IR available	Records in 2022
Roma La Sapienza	3,363	Yes	423,487
Bologna	2,803	Yes	259,051
Napoli Federico II	2,596	Yes	301,666
Padova	2,268	Yes	268,219
Milano	2,142	Yes	260,708
Torino	2,013	Yes	235,613
Firenze	1,667	Yes	216,248
Pisa	1,548	Yes	200,627
Palermo	1,466	Yes	121,392
Politecnico di Milano	1,430	Yes	154,370
Bari	1,404	Yes	123,164
Roma Tor Vergata	1,302	Yes	86,452
Cattolica del Sacro Cuore	1,295	Yes	98,055
Genova	1,241	Yes	153,920
Catania	1,226	Yes	126,167
Messina	1,029	Yes	118,079
Politecnico di Torino	986	Yes	119,127
Perugia	983	Yes	111,383
Campania - L. Vanvitelli	980	Yes	94,859
Salerno	973	Yes	98,146
Milano-Bicocca	972	Yes	100,725
Cagliari	965	Yes	93,471
Pavia	897	Yes	111,209
Roma Tre	864	Yes	107,813
Parma	860	Yes	108,302
Modena e Reggio Emilia	808	Yes	91,082
Della Calabria	793	Yes	63,475
Verona	751	Yes	106,326
Siena	728	Yes	75,520
Chieti-Pescara	691	Yes	68,344
Trento	669	Yes	101,372
Trieste	653	Yes	85,395
Udine	652	Yes	74,726
Ferrara	648	Yes	88,813
Salento	602	Yes	65,528
Brescia	597	Yes	67,824
Ca' Foscari Venezia	587	Yes	71,244
L'aquila	580	Yes	56,431
Sassari	576	Yes	63,089
Politecnica Delle Marche	540	Yes	57,780
Piemonte Orientale	382	Yes	44,163
Insubria	372	Yes	48,282
Bergamo	366	Yes	32,740
Foggia	347	Yes	43,910
Parthenope di Napoli	346	Yes	22,736
Urbino Carlo Bo	333	Yes	41,530
Basilicata	330	Yes	30,542
Bocconi Milano	314	Yes	24,121
Politecnico di Bari	289	Yes	28,553
Molise	287	Yes	29,901

Continues on next page...

University	Faculty in 2019	IR available	Records in 2022
Camerino	279	Yes	27,356
Macerata	275	Yes	35,057
Mediterranea di Reggio Calabria	267	Yes	23,086
Cassino e Lazio Meridionale	253	Yes	26,244
Catanzaro	240	Yes	25,111
Teramo	226	Yes	18,982
L'Orientale di Napoli	206	Yes	21,758
Sannio di Benevento	198	Yes	16,480
S. Raffaele Milano	174	Yes	39,630
Scuola Superiore Sant'Anna	154	Yes	21,429
Università Iuav di Venezia	149	Yes	22,529
Uke - Università Kore di Enna	141	Yes	12,790
Luiss Guido Carli	119	Yes	16,273
Telematica E-Campus	115	Yes	6,819
Suor Orsola Benincasa - Napoli	101	Yes	8,735
Iulm - Milano	92	Yes	9,347
Sissa - Trieste	85	Yes	12,190
Humanitas University	82	Yes	18,583
Scuola Normale Superiore di Pisa	75	Yes	21,076
Stranieri di Perugia	57	Yes	4,988
Telematica San Raffaele Roma	55	Yes	4,248
Lum Giuseppe Degennaro	40	Yes	4,133
Scuola Imt Alti Studi Lucca	39	Yes	3,774
I.U.S.S. - Pavia	33	Yes	3,151
Tuscia	320	No	-
Libera Università di Bolzano	266	No	-
Telematica Guglielmo Marconi	158	No	-
Univ. Campus Bio-Medico di Roma	146	No	-
Telematica Pegaso	115	No	-
LUMSA	97	No	-
Unicusano - Telematica Roma	93	No	-
Telematica Universitas Mercatorum	69	No	-
Roma Foro Italico	67	No	-
Link Campus	65	No	-
Telematica Internazionale Uninettuno	58	No	-
Europea di Roma	57	No	-
Stranieri di Siena	54	No	-
Valle D'aosta	49	No	-
Liuc - Castellanza	47	No	-
Telematica Unitelma Sapienza	41	No	-
Gran Sasso Science Institute	37	No	-
Studi Internazionali di Roma	33	No	-
Telematica Giustino Fortunato	33	No	-
Unicamillus	29	No	-
Scienze Gastronomiche	17	No	-
Stranieri Reggio Calabria	17	No	-
Telematica Leonardo Da Vinci	2	No	-
Telematica IUL	2	No	-
Total	**55,841**	**75 Yes / 23 No**	**5,979,449**

Table B4: List of Italian universities by size and IRs

B.3 Panel construction

Because each university maintains its own IR, but faculty co-author and move across schools, there is a large number of potential duplicates. I address this issue by removing within-person duplicate output identified in two ways: identical titles, journals and publication dates; or title strings within the same journal and publication year that start with the same character and differ by less than 15% of total characters in the longest string, in order to account for manual errors in uploading records. The final output dataset contains 5,940,440 unique records—of these, 54 percent are journal articles, four percent books, 19 percent conference material, 17 percent in-book contributions, the remaining 6 percent accounting for doctoral theses, working papers, patents and other activities.

B.3.1 Disambiguation and individual identifiers

To minimize noise and spurious matches, I opted for a rather conservative disambiguation procedure. First, I start from the full faculty rolls, which contain all employed faculty on December 31st of each year from 2000 to 2020. These rolls contain full legal name, surname, gender, academic rank (type of contract) and university and field affiliations. From the rolls I define individual identifiers from full names. Then, I drop all ids that are repeated in the same year—this steps eliminates faculty with the same name that overlap for at least one year of employment. In terms of person-year observations, it costs 68,346 observations out of the initial 1,213,197.

It is still possible that two identically named individuals have non-overlapping employment spells (e.g. 2000-2008 and 2011-2015) and be confounded as the same person. Thus, I further restrict the sample to those ids that do not have gaps in employment. This step costs an additional 15,797 person-year observations.

The resulting dataset has 88,923 unique individuals and a total of 1,129,054 observations. Notice that most regression analyses will use a slightly smaller sample. This is due to the fact that scientific output data is not available for all universities, and that a few IRs do not offer citation data. To deal with this issue, I drop all individuals that are employed for more 10% of the time in these universities with missing information, bringing the total to 83,032 individuals and 1,061,428 person-year observations between 2000 and 2020.

The faculty-output panel is obtained by merging the faculty panel with scientific output data from IRs. This match is performed exactly on authors' names. For each disambiguated individual in the faculty panel, I generate permutations of their first name(s), surname(s) and abbreviations thereof, creating a mapping between author string and unique ids. Then, author strings for all output records are matched to the unique ids. Overall, this procedure matches 88% of all records to at least one author. Then, output information is collapsed at the author-year level to produce the faculty-output panel.

The final faculty-output panel contains information on all publications, which I left-censor to records dated after 1990 (publications between 1990 and 2000 are only used to compute pre-hire productivity). Including pre-hire publications, between 2000 and 2020 the faculty-output panel has 1,263,067 observations across 77,343 disambiguated individuals.

B.3.2 New hires

The procedure to identify new hires was much less conservative, to allow for manual disambiguation (an aim I later abandoned). It defined a new hire whenever a new name-university-field combination appeared in the stock of faculty. Thus, the total number of loosely defined new hires, 41,501, certainly overstates the number of real entries in Italian academia. After accounting for duplicates and disambiguation, 36,948 individuals are recorded in the faculty panel and, after sample restrictions (universities with no IR or no citations data) 32,514 are in the faculty-output panel.

B.4 Citation and output data

B.4.1 Validation

My scientific output data comes from institutional repositories (IRs) and reflects reported totals as of mid-2022. The analysis heavily relies on citation data, and thus it is important to validate its accuracy. To do this, I employ the *InCites* tool by Clarivate, which allows me to measure the entire body of scientific output produced by scholars with an Italian affiliation that is indexed on Web of Science (WOS). Two important considerations are in order:

1. **My sample is necessarily a subset of the WOS population**: a number of universities do not maintain IRs; some do but did not include indexed citation data at the time of data collection (e.g. PoliTo) and are thus excluded from the analysis; two others only have output data starting in 2004 (i.e. UniBo, UniPa); finally, I do not cover non-university research institutions (e.g. CNR, GSSI) or private sector researchers. Further, notice that

InCites-reported location is based on affiliation, so that multiple affiliations may overstate output in a given location.

2. My sample was collected in mid-2022, whereas the WOS population information reflects totals as of January 1, 2024. Citation stocks in my data should be expected to be **lower for more recent publications** that had less time to garner citations by mid-2022.

Figure B3 compares the WOS population to my sample. To ensure comparability between the two, I only consider WOS-indexed data in my sample—namely, I omit Scopus and PubMed data, even though they further increase coverage.

Given that my analysis is mostly event study-based, sharp changes in coverage over time would be concerning. Reassuringly, Figure B3 shows this not to be the case. The differences in citation data in recent years are entirely due to the different times of data collection, as indicated by output coverage remaining constant.

B.4.2 Making citation stocks comparable over time

IR data give total cumulative citations as of mid-2022. This introduces a clear bias in that more recent publications have had less time to accumulate citations. A second bias inherent in using citation stocks is citation inflation stemming from the fast increasing number of total publications. So, as a measure of scientific impact, citation stocks can understate the importance of both old publications and more recent ones. This can be seen clearly in Figure B4, which depicts global citations stocks by year of publication, using *InCites* data collected in mid-2023. One can see how, pooling together all countries and fields, before 2010 citation stocks increase steadily, a reflection of the increasing body of published research. After 2010, stocks decline rapidly as items have had less time to collect citations. Panels (b) and (c) in Figure B3 shows similar pattern for Italian data as well.

To make my sample of Italian output comparable across time, I make a straightforward and transparent adjustment: I assume that *global average quality* of scholarship has been constant between 1990 and 2023 and adjust stocks with factors such that *global average citation stocks* would be the same across all years. To account for macro-field specific trends and levels, I separately collect output and citation data for 14 disciplinary areas that match the 14 ministry-defined macro-fields used in the analysis. Then I define rescale factors (RFs) as the ratio between average citation counts in a given year and the field-specific maximum. So that the RF for Math is one in the year that Math-related publications had the highest average citation stocks, and weakly grater than one in other years. Applying this RFs to Italian citation data makes them comparable over time, while allowing the Italian average to fall below or rise above the global one—it is a purely cross-time adjustment.

To do this, I manually collected citation stocks by macro-field and year of publication from *InCites* in mid-2023. Given the roughly one year difference between the collection of my IR data (mid-2022) and the global stocks (mid-2023), I shift RFs backward by one year starting on the peak year. The underlying assumption here is that the stock peak year (e.g. 2010) stays the same after one extra year of citations is added between 2022 and 2023, and that the overall shape of the stock distribution would also be constant (up to the rescaling induced by one extra year of citations).

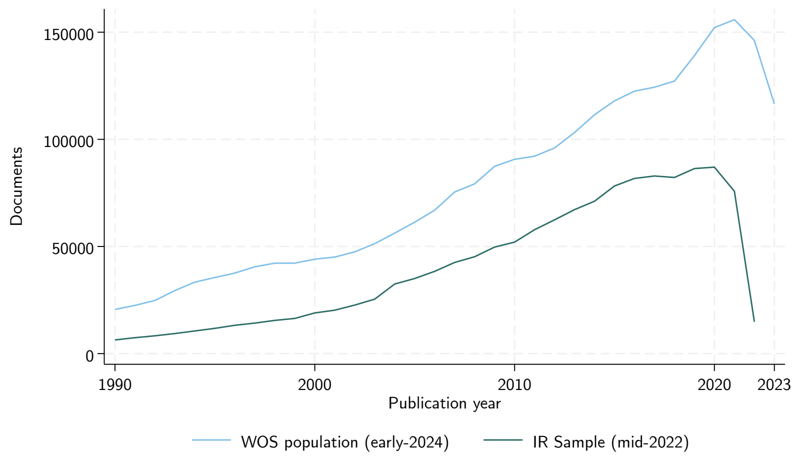
Overall, this correction based on global WoS citation stock data produces 14 field-level rescale factors (RFs) that make citation stocks perfectly comparable across time. Figure B5 illustrates the average resulting rescale factor by year of publication (averaged at the individual-level in my sample).

B.4.3 IHS transformation

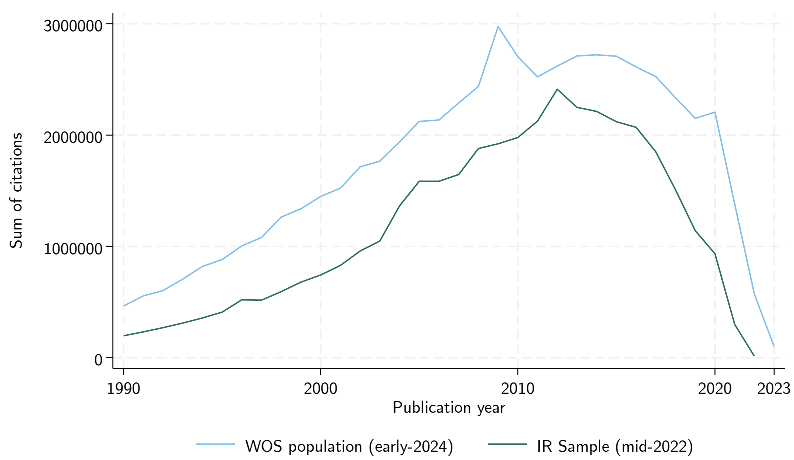
As discussed in the main body of the paper, I find the optimal scale factor for IHS transformation using the R^2 criterion proposed by Aihounton and Henningsen (2021). That is, I pick the value of θ in $f(x, \theta) = \ln\left(x \cdot \theta + \sqrt{x^2 \cdot \theta^2 + 1}\right)$ such that the R^2 is maximized in a regression of interest. My regression is a minimal individual-level event study around the year 2004, for all individuals in my sample (including not treated ones). Rather than treatment, this is just meant to capture time variation in the variable and the general structure of my regressions, and one could omit any other year. The optimal values of θ are reported in Table B5.

Table B5: Optimal IHS scale factors

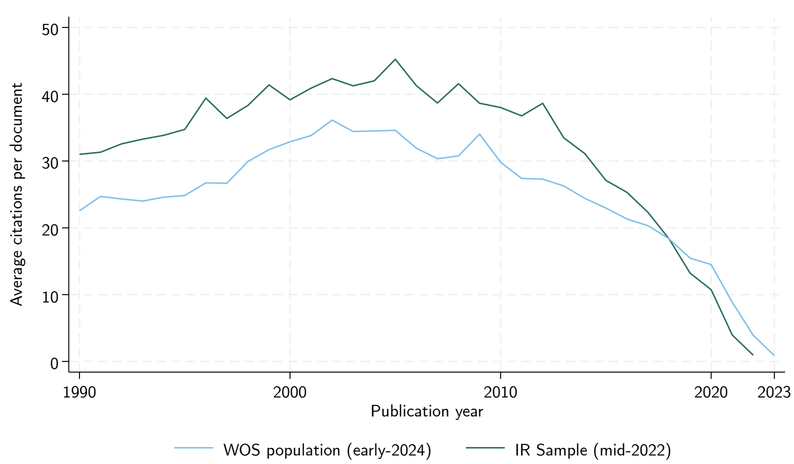
Variable	**Scale factor**
Output	0.55
Output fractional	2.75
Citations	0.55
Citations fractional	10.45



(a) Items



(b) Citation stocks



(c) Average citations per item

Figure B3: WOS scientific output validation

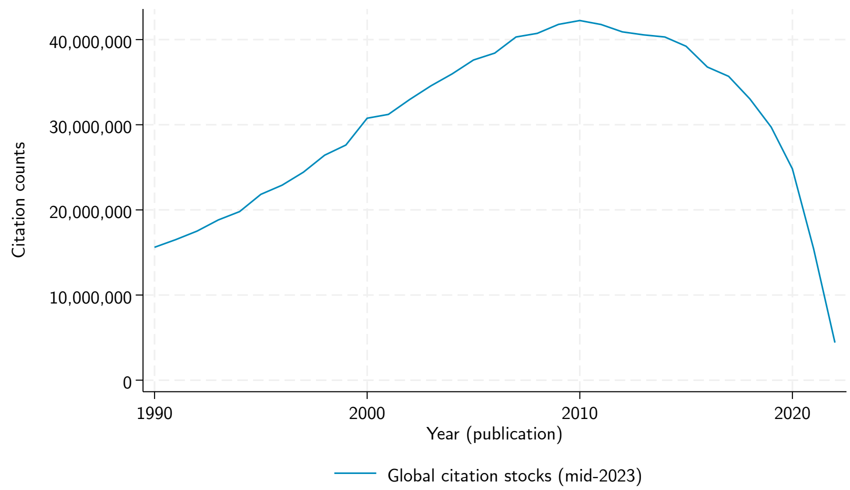


Figure B4: Global citation stocks

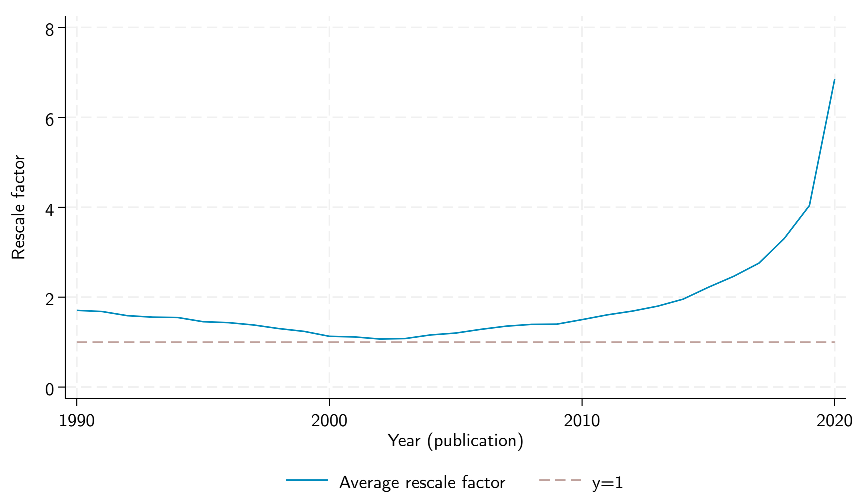


Figure B5: Average rescale factor

C Additional descriptive results

C.1 Mobility

Figure C6 shows the proportion of eligible hires who either exit or are promoted at various intervals post-entry. Interestingly, most variation clusters around the three-year mark, which corresponds to both the benefit duration in most of my sample period, but also to the standard duration of an assistant professor contract (RTD-B) that culminates in review for promotion to associate. Most exits happen at the end of the first three years, or earlier; whereas most beneficiaries who do not exit and are promoted, do so at the three-year mark (taking up the higher ranked job the following year).

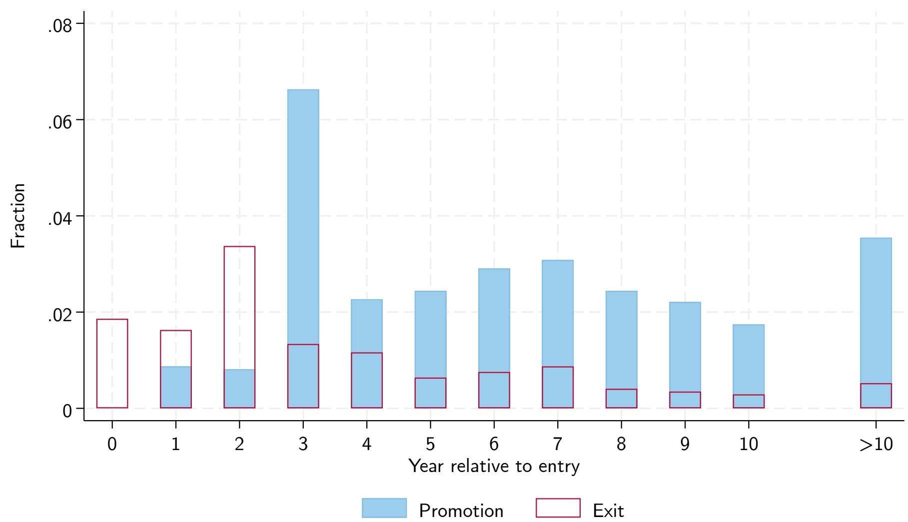


Figure C6: Exits and promotion outcomes for eligible hires

Notes: The two overlaid histograms show the share of eligible hires that either are promoted or exit, by year relative to their time of entry in the faculty panel.

Table C6 presents the breakdown of eligibility for the tax-break by macro-field and time period. Column (1) displays the simulated pre-reform eligibility, serving as a benchmark for post-reform data. Column (2) details the early post-reform phase, prior to the relaxation of hiring rules and enhancement of benefits, while column (3) covers the latter post-reform period.

Table C7 illustrates the beneficiaries' countries of origin.

Table C6: Percentage of all hires who are tax-scheme eligible—by field.

Macro-field	**2001-2003**	**2004-2014**	**2015-2019**
	(1)	(2)	(3)
Maths, computer science	6.9	8.8	22.9
Physics	6.3	12.8	22.3
Chemistry	1.7	3.9	13.7
Earth science	2.6	6.3	11.1
Biology	5.2	7.5	10.0
Medicine	0.8	3.0	3.3
Agricultural science, veterinary	0.6	1.8	6.8
Civil engineering, architecture	0.5	1.9	4.1
Industrial engineering, IT	1.1	3.4	6.6
Antiquity, literature, art history	2.4	3.5	10.2
History, philosophy, psychology	2.0	3.4	11.3
Law	0.5	0.5	2.9
Economics, statistics	6.1	11.2	21.0
Political, social sciences	2.3	3.3	12.5

Notes: Eligibility is expressed in percentage points as a fraction of all hires. Macro-fields correspond to ministry-defined *Aree disciplinari*.

Table C7: Main countries of origin for tax-scheme beneficiaries

Origin	Count	Share (p.p.)
United States	382	20.5
United Kingdom	381	20.4
Other Europe	264	14.2
Germany	218	11.7
France	162	8.7
Other Non-Europe	153	8.2
Switzerland	132	7.1
Spain	101	5.4
Netherlands	72	3.9
Total	1,865	100

C.1.1 Displacement

Because a comprehensive dataset of young researchers who earned their doctoral degrees in Italy is unavailable, I use high-quality surveys from the Italian National Statistical Office (ISTAT). These surveys are highly representative as ISTAT, through individual universities, was able to reach out to the entire population of interest and obtain an overall response rate above 70%. However, notice two limitations: the first cohort to be surveyed received its doctoral degree in 2004, and therefore we lack observations before the introduction of the 'returning brains' scheme;

moreover, six cohorts were interviewed in three survey years, and so their outcomes are measured at different points in time since their first entry in the job market. Thus, the evidence presented below is purely suggestive.

Nonetheless, Table C8 shows a pattern that is consistent with displacement effects induced by the greater inflow of eligible hires from abroad and with the broader period of constrained demand from public universities. Notice how, while the overall employment rate is high and fairly constant, the fraction of locally-trained PhDs working in Italy in a research-related job declined markedly (-15 percentage points between the 2004 and 2014 cohorts), and how this is entirely driven by a reduction in the share that works in Italian academia in particular (-17 percentage points). Interestingly, the fraction of those in tenure-track equivalent positions has declined by a similar percentage point amount, suggesting that the fraction employed in post-docs and other temporary positions like university researchers or lecturers has remained roughly constant.

Table C8: Evidence of displacement

	Outcome: **Share** (p.p.)					
Survey year:	2009		2014		2018	
PhD year:	2004	2006	2008	2010	2012	2014
	(1)	(2)	(3)	(4)	(5)	(6)
Employed	93.7	92.5	93.4	91.6	94.0	94.0
Working in Italy	88.0	85.5	88.9	86.2	85.1	83.6
and working in research	42.7	39.3	33.7	32.6	25.8	28.1
and working in academia	34.4	32.2	24.1	22.8	16.2	17.3
and in tenure-track jobs	19.2	9.8	9.7	4.5	5.8	3.5
Observations	3,928	4,886	7,888	8,434	8,172	7,885

Notes: 2009, 2014 and 2018 ISTAT surveys about employment of Italy-trained PhD holders (*Indagine sull'inserimento professionale dei dottori di ricerca*). Available online at `https://www.istat.it/it/archivio/87536`.

C.2 Selection

Table C9: Productivity differences at the time of entry—untransformed metrics

	Citations		
	Raw count	Author-adj.	Author & Time-adj.
	(1)	(2)	(3)
Eligible	1419.99***	49.45***	6.80***
	(499.64)	(13.22)	(0.97)
Observations	25,565	25,565	25,565
R^2	0.07	0.20	0.15
Controls	Y	Y	Y
D.V. Mean	1522.36	152.41	12.38

*** p<0.01, ** p<0.05, * p<0.1, robust standard errors in parentheses.
Controls: flexible time trend by macro-field, field FEs, academic rank FEs.

To assess whether the degree of positive selection of eligible hires has changed over time, consider the following event study specification:

$$Output_i = \sum_{j=-3}^{15} \gamma_j L_{jt(i)} + \sum_{j=-3, j\neq -1}^{15} \beta_j L_{jt(i)} \times Eligible_i + \phi_{i,t(i)} + \rho_i + \varepsilon_i \tag{1}$$

where $t(i)$ indexes the calendar year when individual i is hired, j indexes the leads or lags relative to the reference year 2004, $L_{jt(i)}$ is an indicator equal to 1 if the calendar year of hire $t(i)$ is at distance j from 2004, $\phi_{i,t(i)}$ represents field-specific temporal trends, ρ_i are fixed effects for academic rank. That is, I consider an event study around the reform and interact leads and lags with an indicator for eligibility in order to capture the within-year difference in scientific output by eligibility status. As outcome, I consider time- and author-adjusted citation counts.

Figure C7 shows a steady pattern in the degree of positive selection of the targeted population. Coefficients are large and mostly significant. Overall, it looks like selection stayed high after the reform, but was reduced during the economic and financial crisis of 2009-2011, and started declining around the time when the number of eligible people began increasing markedly (*ca.* 2018).

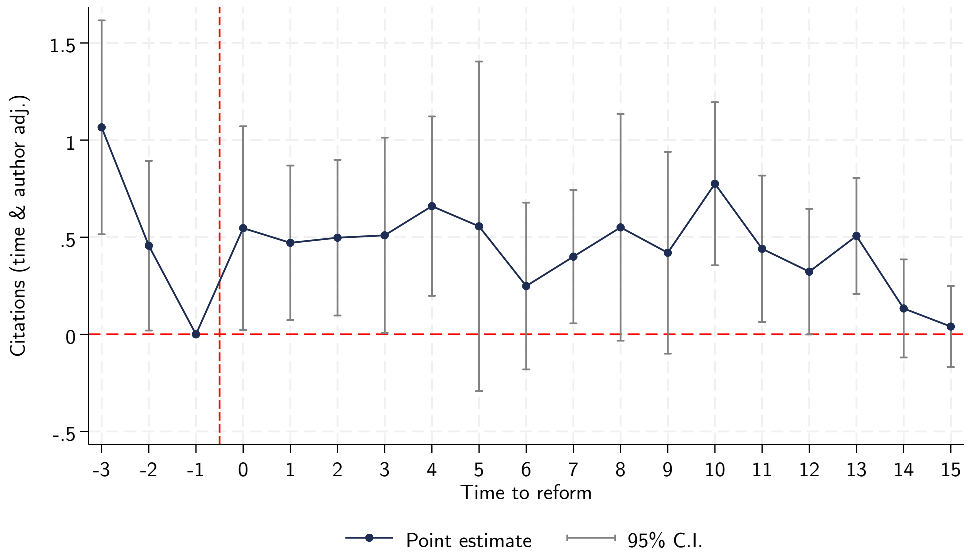


Figure C7: Strength of selection over time

C.3 Treatment Intensity

Figure C8 shows the distribution of average cell productivity $\overline{\theta}_c$ by tercile of TI.

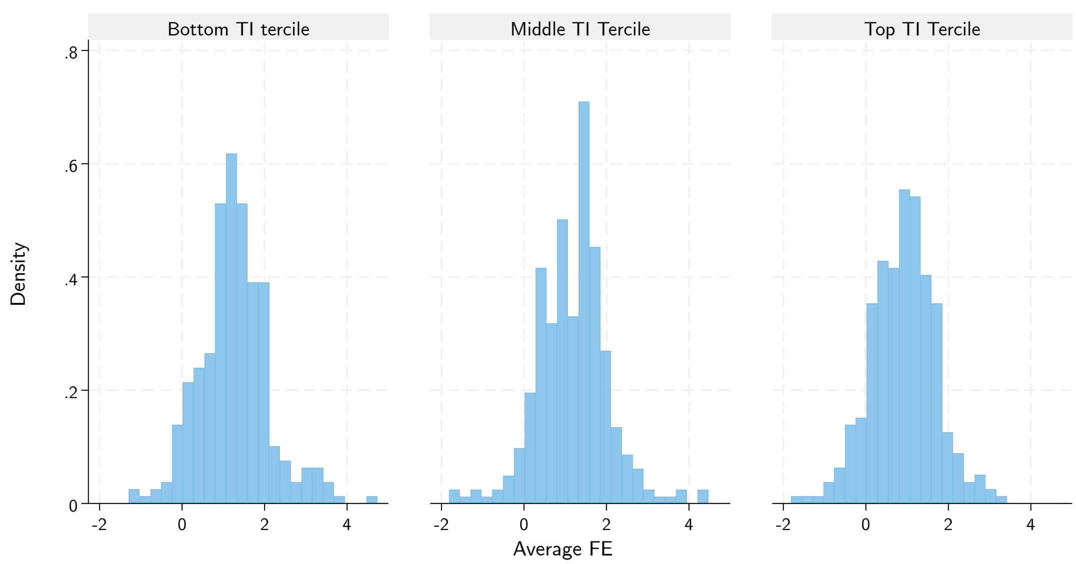


Figure C8: Average incumbent productivity by tercile of TI

Notes: The figure shows the distribution of the average permanent productivity of faculty in a given cell ($\overline{\theta}_c$), by tercile of the treatment intensity measure TI.

D Additional event study results

D.1 Group analysis robustness

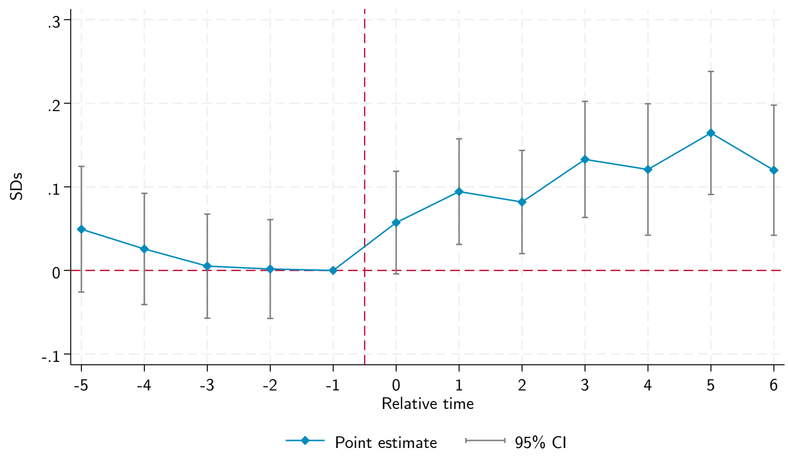
In the following I present robustness checks for different choices of X_{ct} for the main group-level results.

1. Check #1: X_{ct} is calendar-time trends by macro-field.
   See Figure D9.

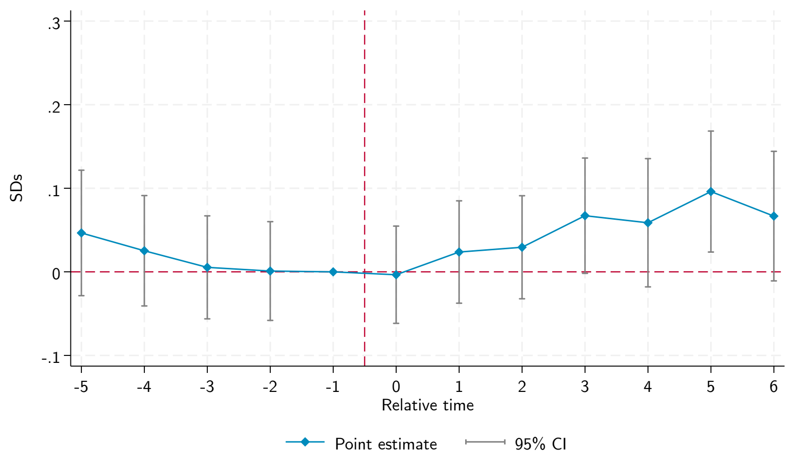
2. Check #2: X_{ct} is calendar-time trends by tercile of pre-period (2001-2003) research funding per person.
   See Figure D10.

3. Check #3: X_{ct} is calendar-time trends by tercile of pre-period ($t = -1$) fraction of faculty with any positive citations.
   See Figure D11.

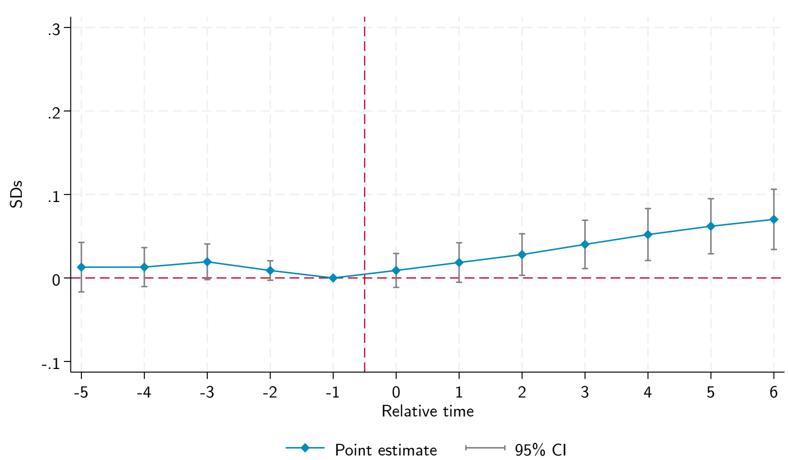
4. Check #4: X_{ct} is calendar-time trends by tercile of pre-period ($t = -1$) average permanent productivity $\overline{\theta}_c$.
   See Figure D12.



(a) Total effect

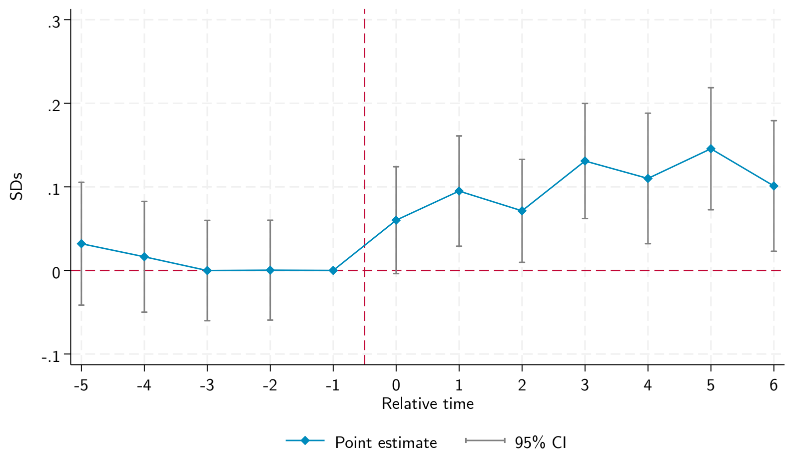


(b) Indirect effect

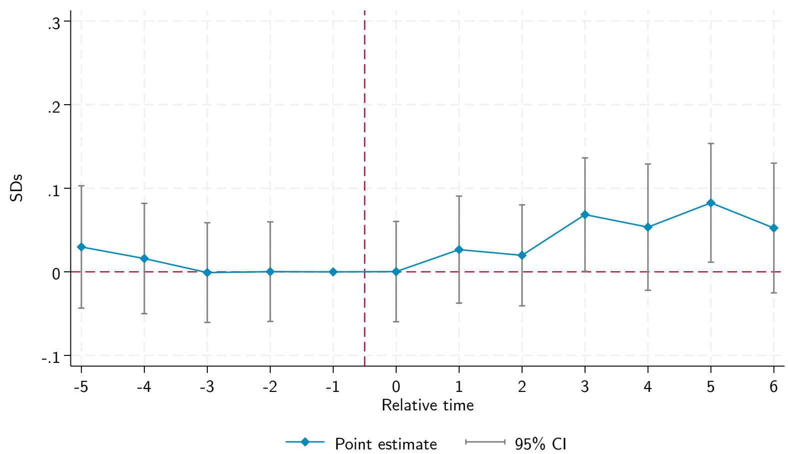


(c) Composition effect

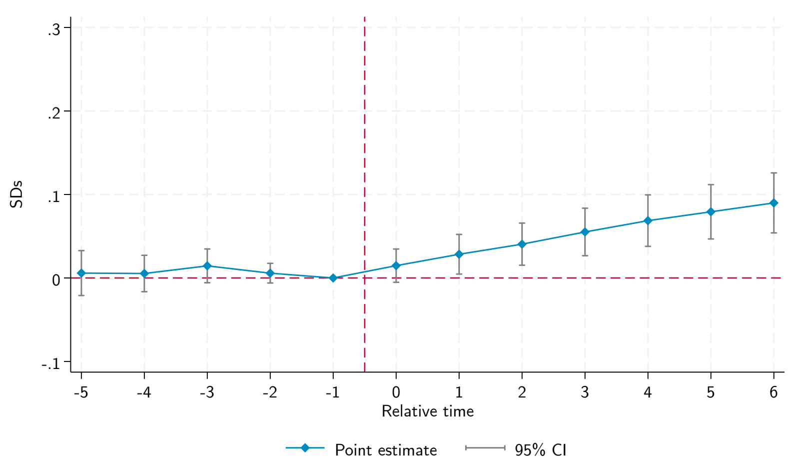
Figure D9: Group-level results, robustness check # 1



(a) Total effect

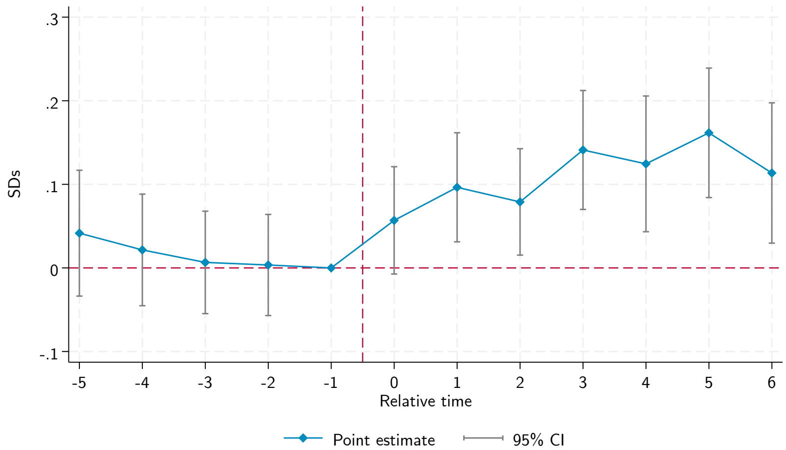


(b) Indirect effect

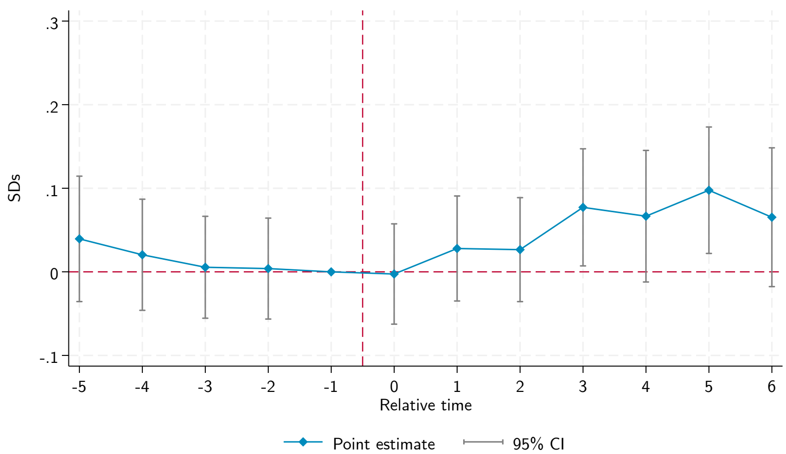


(c) Composition effect

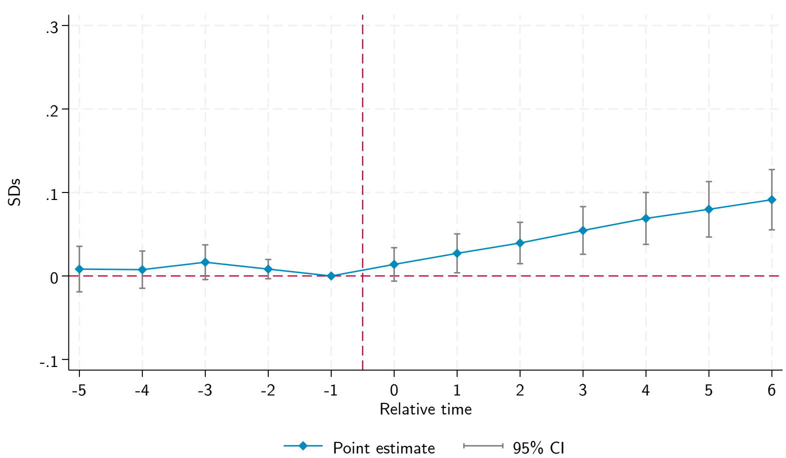
Figure D10: Group-level results, robustness check # 2



(a) Total effect

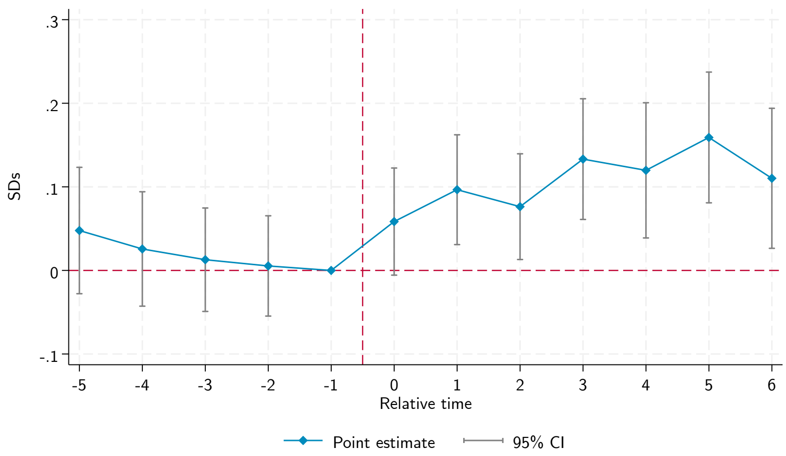


(b) Indirect effect

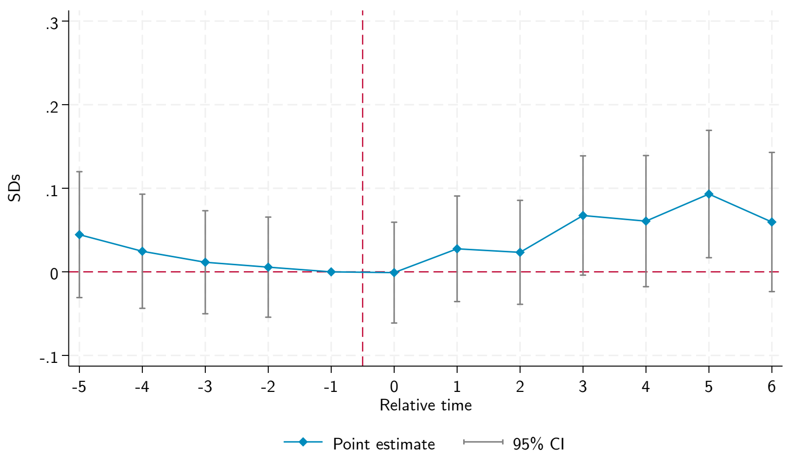


(c) Composition effect

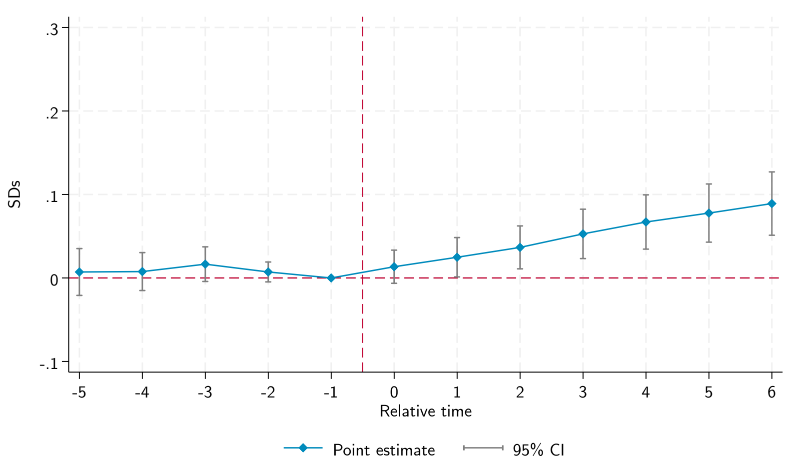
Figure D11: Group-level results, robustness check # 3



(a) Total effect



(b) Indirect effect



(c) Composition effect

Figure D12: Group-level results, robustness check # 4

D.2 Cell analysis additional results

Figure D13 breaks down the estimated cell-level effects by broad scientific area. The Figure shows more pronounced effects in Engineering, Economics, Medicine, and related fields, and smaller impacts in Mathematics, Physics, and other sciences. Humanities and other fields with low initial citation rates and fewer events yield noisier, less reliable results, and are thus not reported.

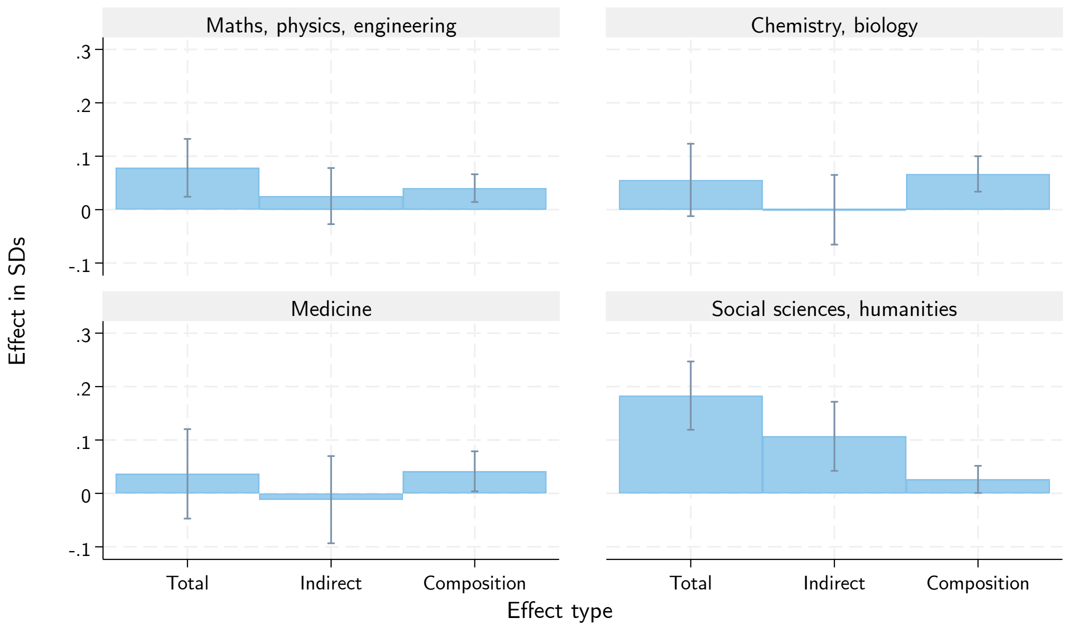


Figure D13: Indirect effect: Heterogeneity by broad scientific areas

D.3 Individual analysis robustness

At the individual-level, I conduct similar robustness checks as for cells, but using the same X_{ct} trends for terciles of the individual's cell at the time of the event ($t = 0$). Unreported (for brevity) robustness checks included a university-specific trend and including FE just for the initial cell (rather than any cell the individual moves to).

1. Check #1: X_{ct} is empty.
   See Figure D14.

2. Check #2: X_{ct} is calendar-time trends by cell-level tercile of pre-period (2001-2003) research funding per person.
   See Figure D15.

3. Check #3: X_{ct} is calendar-time trends by cell-level tercile of pre-period ($t = -1$) fraction of faculty with any positive citations.
   See Figure D16.

4. Check #4: X_{ct} is calendar-time trends by cell-level tercile of pre-period ($t = -1$) average permanent productivity $\overline{\theta}_c$.
   See Figure D17.

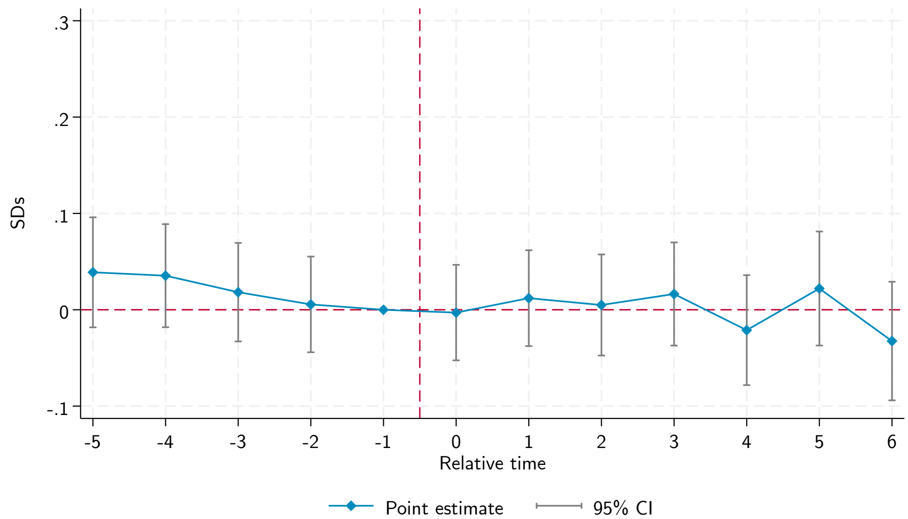


Figure D14: Individual-level results,robustness check # 1

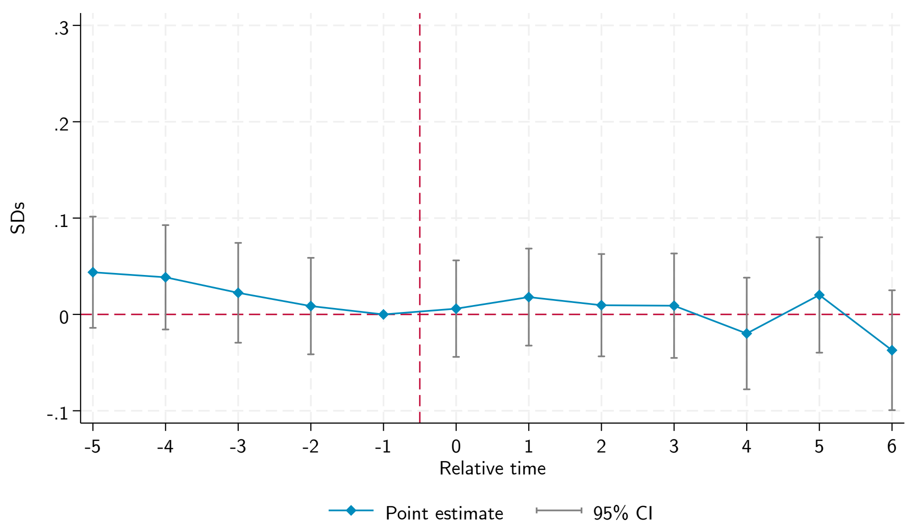


Figure D15: Individual-level results,robustness check # 2

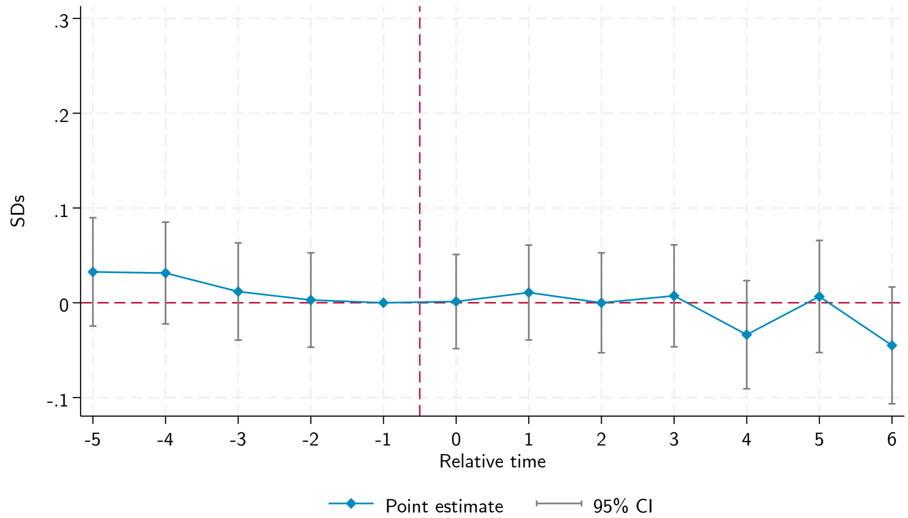


Figure D16: Individual-level results,robustness check # 3

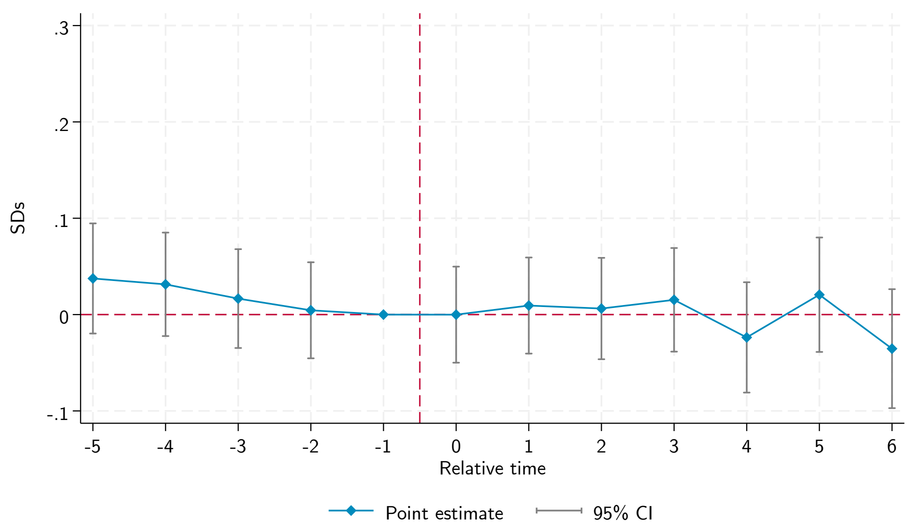


Figure D17: Individual-level results,robustness check # 4