The Journal of Technology Transfer
https://doi.org/10.1007/s10961-024-10078-6

Star help and knowledge transfer: an event study analysis of star interactions observed from acknowledgement texts

Akhil Sasidharan[1] · **John McHale**[1] · **Jason Harold**[1]

Accepted: 26 February 2024

Abstract

This paper contributes to the growing literature on the impact of connections to star scientists on the productivity of academic scientists. The existing literature generally focuses on larger economies and specific scientific fields in evaluating star-connection effects. It has rarely examined the particular channels through which stars have their effects. Using natural language processing (NLP) techniques to explore the acknowledgement texts of a broad corpus of published papers from three small open economies, we examine the effects of star help revealed by the acknowledgement texts published in articles. Using an event-study framework with matched data, we find evidence of an economically and statistically significant effect on scientist productivity in the year of acknowledgement of star help. However, there is only evidence of an enduring productivity effect if scientists maintain their acknowledgement of ties to the star over time. A similar pattern is evident across different types of acknowledgements, except for acknowledgements of star help with access to materials, which shows an enduring effect even after a single acknowledgement. The largest estimated star-help effects are found for authors in lower quartiles of the field-specific productivity distribution measured in the year before the help is acknowledged. The results are robust to using a raw-publications-based measure of scientist productivity in place of our preferred citation-weighted publications measure of productivity, suggesting that the observed productivity effect is unlikely to be due to a pure signalling effect. We discuss the implications of these findings for the design of star recruitment and integration policies.

Keywords Star scientists · Small open economies · Acknowledgements · Knowledge transfer

JEL Classification O31 · J24 · J61 · I23 · C54

✉ Akhil Sasidharan
a.sasidharan2@universityofgalway.ie

John McHale
john.mchale@universityofgalway.ie

Jason Harold
jason.harold@universityofgalway.ie

[1] Whitaker Institute for Innovation and Societal Change, University of Galway, Galway, Ireland

1 Introduction

Star scientists have long been recognised as catalysts for knowledge creation within academia. There is also a growing policy interest in smaller economies and regions in increasing their connections to academic star scientists as a means to catalyse dynamic local research clusters. Although explicit policy programmes have been implemented, there is limited understanding of the mechanisms through which star scientists have an effect on local scientific productivity. One candidate mechanism is the direct effect of help from star scientists to incumbent local scientists, even where an explicit co-authorship relationship does not exist. Existing studies have generally not focused on the impacts of star scientists in smaller economies and have tended to overlook modes of knowledge transfer beyond co-authorship. Moreover, the importance of network building, both within and across institutions, emerges as a strategy for facilitating access to star help and unlocking its potential for productivity enhancement. This research contributes insights into knowledge and technology transfer by providing new empirical evidence on the knowledge transfer mechanisms facilitated by star scientists, and offering actionable recommendations for improving academic and research management practices in small open economies. This paper addresses these gaps with new empirical evidence on knowledge transfer that occurs through a star-help mechanism in small economies, where star help is indicated by acknowledgements to various forms of star help in the acknowledgement sections of published journal articles. To broaden the evidence base and recognise that developing connections to international star scientists has been a particular policy focus for smaller open economies (SOEs), our data are drawn from academic departments in Ireland, Denmark and New Zealand. As we use the term in this paper, 'star scientist' refers to those individuals who are highly productive in terms of their publications and citations and are located chiefly at a university (Schiller & Diez, 2010; Zucker & Darby, 2006).

Scientific collaboration through formal (e.g., co-authorships) or informal means (e.g., discussions and helpful comments) connects a star with other scientists. These interactions could play an essential role in generating knowledge spillovers. This paper focuses on one particular channel of star impact and aims to investigate the effects of the star's help on their colleagues' productivity. Helpful interaction—either with co-located or non-co-located scientists—is a relevant and usually neglected channel for learning and knowledge exchange in socially interactive environments. Interactions with stars, even if they are informal, influence behaviour, research topics, collaborations, and the productivity of scientists (Bramoullé et al., 2020). More frequent and long-lasting interactions impact the extent of knowledge overlap (Nooteboom, 2000), facilitate learning processes, and enhance knowledge creation through the recombination of existing ideas between stars and scientists (Nelson & Winter, 1982; Weitzman, 1998). The empirical challenge is that evidence of help does not leave a paper trail; however, explicit acknowledgements to help from a star provide a (noisy) indicator of star interaction (Bozeman & Corley, 2004; Oettl, 2012). We broadly interpret such acknowledgements as evidence of helpful interaction. From acknowledgements in the publications, we use natural language processing (NLP) techniques to identify the star scientist names that are being acknowledged. We then draw on recent advancements in econometric techniques to estimate the productivity impacts of star scientists' help on the scientific productivity of the acknowledging authors. We also explore the effects of different kinds of star help by categorising acknowledgements into five types: Conceptual, Technical, Material, Funds & Support, and Other Types.

The basic idea behind the empirical approach we take in the paper is that help (in various forms) from star scientists can raise the productivity of the assisted scientists and that an indicator of the presence and form of such help can be gleaned from the acknowledgement texts of publications. Our core measure of productivity is (forward) citation-weighted publications in a given year. However, a key empirical challenge is that an observed correlation between acknowledgements and productivity might not reflect a causal productivity effect. An alternative possibility is that an acknowledgement leads to a signalling effect, whereby a publication with a star acknowledgement is more likely to be cited due to a perceived endorsement from a star. To distinguish between the productivity and signalling hypotheses, we repeat the analysis using a measure of raw publications as our dependent variable. Assuming blind refereeing, where referees do not have sight of the acknowledgement texts, raw publications should not be contaminated by a signalling channel. We can, therefore, attain some confidence that we are observing a true productivity effect if the results are robust to switching to a citation-insensitive dependent variable. We find similar qualitative dynamic patterns across the two dependent variables.

Our data consists of all Scopus-recorded academic publications by authors affiliated with a university in the three SOEs across 27 subject fields for the period 1990–2017. Scopus provides acknowledgement data under the 'Funding Text' heading. Using this pool of acknowledgement data, we identify star names and their type of helpful interaction for each publication. We classify the type of acknowledgements based on the keywords often used to acknowledge the star's contribution to the publication. This help by the star can take many forms: an author might acknowledge the star for helpful discussions throughout research; there may be an acknowledgement to comments received during a seminar or a conference; there may be acknowledgements to various forms of technical assistance, especially in the fields of biochemistry, pharmaceuticals, and medicine; or there could be theoretical or mathematical help received during the analysis stages of research. It is important to note that these types of help are not formally captured as a recognised co-authorship. The co-authorship channel has been a major focus of the literature exploring the effects of star connections on productivity (see, e.g., Oettl, 2012). In this paper, we explore the effects of star help where a formal co-authorship relationship does not exist. Our paper thus extends the literature by exploring the productivity effects of star help even where the star does not become a co-author. We investigate both the immediate effects of observations of acknowledgements of star help and any enduring effects over time using a dynamic event-study approach.

Using this event-study methodology, where the event is the acknowledgement by the author to a star in the publication, we test a number of hypotheses relating to the effect of star help on the acknowledging scientist's productivity. We employ a coarsened exact matching (CEM) of star-acknowledging and non-star-acknowledging scientists to create a similar treatment and control group. The findings support the hypothesis of a positive productivity effect and find an interesting dynamic pattern of productivity effects following an acknowledgement to a star. First, the results show a positive star-help effect on the authors who acknowledge a star, with the impact on productivity evident in the year of publication but largely disappearing in the subsequent years. Second, the results indicate that an author who continues to acknowledge a star in the years after the initial interaction has a higher and more persistent productivity effect. The results are robust to using raw publications as the dependent variable. In considering the types of acknowledgements, we find a significant positive but transitory effect on author productivity in the year of acknowledgement across all types, except materials where the effect is shown to be more persistent. We speculate that materials-related acknowledgements may reflect more enduring relationships that are not repeatedly acknowledged in publications over time. Finally, we find that the estimated productivity impact of star help is greater for assisted authors that are in the lower quartiles of the relevant field-specific productivity distribution in the year before the (initial) acknowledgement occurs, suggesting that less productive authors get greater benefit from a star's help.

This paper contributes to the literature in several ways. First, the paper investigates the impacts of star help on peer productivity. Prior research mainly focuses on the star scientist's proximity with limited consideration of the precise mechanisms involved (Agrawal et al., 2017; Azoulay & Zivin, 2005). Second, in contrast to the existing literature, the paper focuses on the effects of star help in small open economies. This fills a significant gap in the literature, given the importance that policies to improve connections to star scientists (including but not limited to star recruitment policies) play in the policy mix of smaller national and regional economies. Third, the paper identifies non-co-authorship-based star connections using acknowledgements in the publication, where we hypothesise that such non-co-authorship-based interactions have an impact on a scientist's productivity. Moreover, the paper differentiates these informal interactions into Conceptual, Technical, Material, Funds & Support, and Other Types. Finally, the paper employs the growing literature on heterogeneous treatment effects accounting for staggered treatment by using Sun and Abraham's (2021) estimator. This allows us to consistently estimate the effects of interest in the potential presence of heterogeneity across both cohorts and periods.

The remainder of the paper is structured as follows. Section 2 reviews the related literature and describes the framework and hypotheses. Section 3 describes the data and the use of NLP in extracting acknowledgement data from the acknowledgement texts. Section 4 outlines the event-study-based empirical framework used to identify the causal effects of star help. Section 5 reports the results, including various robustness tests for the analysis. Finally, Sect. 6 concludes with a review of the main findings and their implications for science policy.

2 Related literature and hypotheses

2.1 Related literature

A key determinant of the growth of an economy is the accumulation and transmission of knowledge (Romer, 1986, 1990). Romer identifies two processes, human capital spillovers and R&D investments, where knowledge accumulation and spillovers happen. Human capital spillovers benefit others in the economy through increased productivity and innovation. These innovations are often a result of interaction between individuals with a considerable stock of previously accumulated knowledge. As potential key figures in this knowledge spillover process, star scientists may play a central role in the workflow structure due to their unique expertise and social status (Paruchuri, 2010). These spillover agents are often highly skilled pioneers in their fields who induce knowledge flows across organisational and regional boundaries and affect ongoing research projects inside academia and firms (Bergman & Schubert, 2005; Maier et al., 2007).

Waldinger (2012) and Azoulay et al. (2010) discuss the adverse impacts of losing ties with a star on collaborators. They find a decline in productivity and quality-adjusted publications for those affected by a star's externalities. Zucker and Darby (2006) find a positive impact of star scientists in start-up biotech firms. In these studies, researchers focus more on the impact on the productivity distribution rather than the mechanism enabling the knowledge flow. Further studies have provided evidence that the influence of a star might depend upon the collaborative strength and breadth of the star's expertise (Kehoe & Tzabbar, 2015) and have documented significant long-term effects on the performance of junior researchers (Li et al., 2019).

Although the presence of stars could catalyse the development of peers' scientific output, helpful interactions are essential for some knowledge transfer mechanisms. First, human capital externalities generate knowledge flows that must be recognised as valuable, then assimilated and recombined with previous knowledge to generate new scientific knowledge. In other words, to take advantage of stars' helpfulness, scientists have to increase their absorptive capacity and their capacity to apply knowledge to their research projects (Cohen & Levinthal, 1990). Although a certain level of previous related knowledge is required for learning to take place (Cohen & Levinthal, 1990; Torre & Rallet, 2005), recurrent helpful interactions and the resulting knowledge exchange adjust the optimal level of knowledge overlap between scientists over time (Nooteboom, 2000), which is fundamental to facilitate the transfer of knowledge (Boschma, 2005).

Second, these interactions also take advantage of spatial proximity to transmit formal and tacit knowledge forms (Bathelt et al., 2004). Scientists working in the same department go through face-to-face (F2F) communications, allowing efficient transmission of complex and tacit knowledge (Gertler, 2003; Storper & Venables, 2004). However, organised proximity, i.e., belonging to the same community of practice or sharing the same system of representations, could alleviate the requirements of being co-located for knowledge transfer and learning to occur. Information and communication technologies and temporary co-locations such as meetings and academic congresses could solve communication problems (Torre & Rallet, 2005). In the case of star help, both channels (F2F and organised proximity) are helpful for knowledge transfer that affects the productivity of incumbent scientists.

Finally, information sharing is the key to scientific progress in any spillover mechanism. Merton (1973) defines the norms of unconditional knowledge-sharing as one of the critical processes in academic life. At the same time, there is tension among academic individuals to share their valuable information due to incentives associated with the research (Blumenthal et al., 1996; Murray, 2010; Walsh & Cohen, 2008). Haeussler et al. (2014) consider two possibilities in which such information sharing happens: specific sharing, where researchers share their private work on a specific request, which in turn will get recognised for future accomplishments, and general sharing, where such information is shared publically, such as sharing unpublished data, materials, etc.

This study discusses how star help affects scientists' productivity. Through this particular mechanism, the absorptive capacity of scientists is enhanced, the optimal level of knowledge overlap is adjusted through repeated social interactions, and knowledge sharing is facilitated. However, while networks can positively affect productivity, their specific nature can also play a crucial role in determining their impact. For example, a study by Horta et al. (2010) found that academic inbreeding, or hiring PhD graduates as faculty members at the same university where they received their degree, can harm scholarly output. Of course, although understudied, informal interaction is just one example of a network-intermediated productivity effect. More broadly, networks help recombine specialised ideas and knowledge transformation (Groysberg et al., 2008; Rothaermal and Hess, 2007). Another widely studied network effect is co-authorship. Grigoriou and Rothaermel (2014), Agrawal et al. (2017), and Yadav et al. (2023) find increased productivity gains and show very high positive spillovers of co-authorship with a star. Also, the network position of the co-author is crucial for the productivity that facilitates access to nonredundant knowledge

(Mohnen, 2022). However, the impact of stars on productivity through non-co-authorship-based network channels remains comparatively understudied.

One important exception to the neglect of the helpful interaction channel is Oettl (2012). He defines a taxonomy of stars based on these social interactions: maven (highly helpful and average productivity), all-star (highly helpful and highly productive), lone wolf (highly productive and less helpful), and non-star. Following Oettl's approach, we use acknowledgement texts in publications to provide a paper trail on helpfulness activities. Other studies have also utilised the information available in acknowledgement texts. Mackintosh (1972) classifies acknowledgements based on the facilities, access to data, and help of individuals. McCain (1991) classifies acknowledgements from experimental papers in genetics into five typologies: access to research-related information, unpublished results or data, peer interactive communication, technical assistance, and manuscript preparation. Cronin (1991) (later modified in Cronin et al. (1993)) introduces six typologies for acknowledgements: access, peer interactive communication, moral support, technical support, clerical support, and financial support. These studies focus on the typology of acknowledgements, while our focus is on how different types of acknowledgements are associated with the observed productivity effect on the acknowledging author. Moreover, we build on this literature by using observed acknowledgements as not just thank-you notes but indicators of important productivity-affecting activities that are distinct from co-authorships and citations (Cronin, 1991; Paul-Hus & Desrochers, 2019).

Nevertheless, the information in the acknowledgements of academic publications should be applied with caution and with sensitivity to context. Using acknowledgements as a sign of helpful interaction with a star scientist could be either a sign of intellectual debt or a signalling effect. For example, along with co-authorship, acknowledgements could also reveal other forms of collaboration (Laudel, 2002). The content of such acknowledgements is intended to repay debt towards formal and informal collaboration. In addition, acknowledgements could reflect personal relationships and individual preferences. Hellqvist (2010) conducted a qualitative analysis of acknowledgements in sociology journals and suggested that personal style, editorial guidelines, cultural norms, and ethical principles influence the pattern of acknowledgements.

The literature also considers the various motivations for acknowledging prior work in a paper. Berg and Faria (2008), for example, argue that the scientist names chosen in acknowledgements are based on the effect they may have on readers. We, therefore, recognise that acknowledgement of star help is a noisy indicator of actual star help but assume there is a sufficient signal in the acknowledgement data to make them useful in the estimation of star-help effects on research productivity. Furthermore, the type of helpful interaction also matters to those who receive such help from star scientists. As argued by Oettl (2012), the type of helpfulness being acknowledged is also relevant to the productivity effects of these acknowledgements, and we also distinguish between different types of star help in our empirical analysis.

2.2 Framework and hypotheses

We conceptualise a star scientist as a scientist who is usually productive and connected within their network. From the point of view of a non-star scientist, we assume, following the literature cited above, that forming a connection with a star scientist changes their network position and potentially their productivity. As an example, a scientist will obtain access to knowledge through their network, with their eigenvalue centrality within their

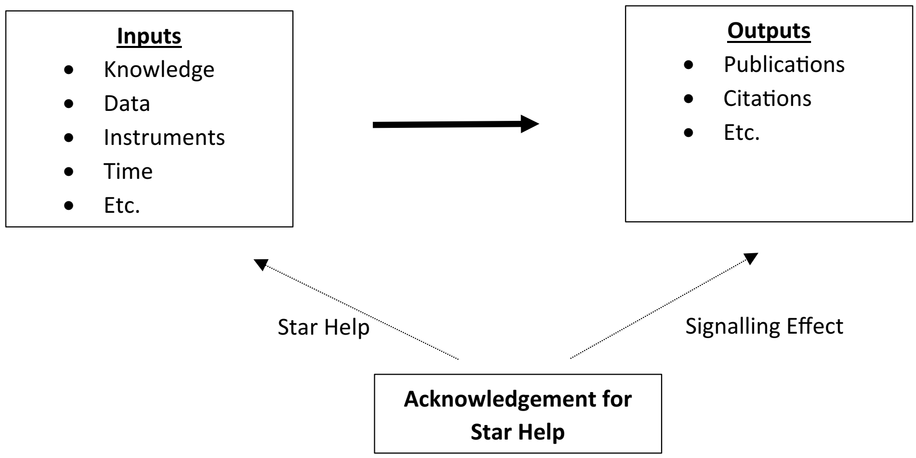

Fig. 1 Star help and observed scientific productivity

network being one commonly used measure of how easily they can access knowledge (Newman, 2018). Forming a connection with a star will change their centrality and thus their access to knowledge. As another example, forming a collaborative connection with a star can directly increase productivity and may have further indirect effects through access to the star's broader collaborative network.

This paper focuses on connections to a star that deliver potentially productivity-improving star help, separate from any direct effect through co-authorship. We think of such star help as augmenting the scientist's inputs into their production of scientific outputs (as measured by citation-weighted publications). Figure 1 captures schematically the relationship between inputs and outputs, where star help is viewed as augmenting the scientist's inputs. We also allow for the possibility that a positive effect of observed star acknowledgements could reflect a signalling effect, where the positive effect on subsequent output—particularly citations to a publication—reflects a signalling (or reputational spillover) effect from the star.

Our empirical approach is to look for evidence of the presence and type of star help in the acknowledgement text of publications. By carefully matching each scientist that is "treated" with star help to an observationally similar scientist who is not observed to receive such help, we use an event-study framework to estimate the dynamic effects of treatment.

Our first hypothesis is that an observation of acknowledgement for star help is associated with a contemporaneous increase in star productivity, with productivity measured by citation-weighted publications.

Hypothesis 1a The observation of an acknowledgement for star help is associated with an increase in citation-weighted publications in the year of acknowledgement.

In addition, an important advantage of our event-study framework is that we can observe the effects of an acknowledgement for star help in a given year on the evolution of productivity over time. This leads to our second hypothesis:

Hypothesis 2a Repeated observations of acknowledgements for star help are associated with a sustained increase over time in productivity as measured by citation-weighted publications.

A significant challenge is disentangling the input-enhancing role implied by observed star acknowledgements from signalling effects. On the assumption that acknowledgement texts are not observable to referees, raw publication data (i.e., output not weighted by subsequent citations) should not be contaminated by signalling effects. This leads to the following hypotheses based on a simple raw publications measure of output:

Hypothesis 1b The observation of an acknowledgement for star help is associated with an increase in raw publications in the year of acknowledgement; and.

Hypothesis 2b Repeated observations of acknowledgements for star help are associated with a sustained increase over time in productivity as measured by raw publications.

Finally, acknowledgements to a star scientist can be classified by type based on the keywords in each publication's acknowledgement text. These keywords assist with identifying the nature of the help between the author and star, leading to our third hypothesis that the effect of star help on an author's productivity will vary across the types of helpful interaction with the author.

Hypothesis 3 The observation of an acknowledgement for star help is associated with an increase in citation-weighted publications in the year of acknowledgement, with the magnitude of the increase potentially varying across the different types of helpful interaction.

3 Data and methodology

Our dataset consists of all the publications and their citations across the 27 subject fields identified by Scopus, where each publication contains at least one author affiliation recorded in Ireland, Denmark, or New Zealand. As discussed previously, Ireland and Denmark are chosen because they are small open economies with nationally funded star recruitment programmes, while New Zealand is chosen because it is also a small open economy with no formal star recruitment programmes. We collect variables at the publication level for the year of publication, authors, affiliations, subject field, citation count to 2019, references, abstracts, keywords, and acknowledgements. We identify authors and publications using unique Scopus identifiers. The dataset contains approximately 1.43 million publications divided over 219,582 unique authors. The unit of observation is at the author level rather than the publication level, and the dependent variable tracks the author's productivity before and after the star help revealed from the acknowledgement texts.

To help capture an author's performance over time, we restrict our analysis to the authors present in one of the three countries for at least four years. For each author who satisfies this condition, we access their catalogue of publications dating back to 1990, including those outside their affiliations in the three focal countries. The final dataset contains an unbalanced panel of 889,479 unique publications divided over 59,122 authors across the years. Using the affiliation data of the author, we determine the department (defined based on the Scopus definition of subject fields[1]). If an author publishes in more than one subject field, we assign them to the department in which they have the most publications.

3.1 Identification of a star scientist

Given that this study focuses on the impact of star scientists on their peers through channels of knowledge spillover identified by the acknowledgement text in an author's publication, the first step is to identify the star scientists in our dataset. The dataset contains the citation distribution of each author from 1990 for each subject field in any given year. A star is defined as one who is at or above the 95th percentile of scientists in the cumulative distribution of citations for the relevant subject field received for all publications up to that year. In measuring the citations to any given publication, we measure all citations up to the end of our observation period in 2019, so our publication quality measure depends on the subsequent citations to that publication.

We only identify star scientists from 1996 onward to allow enough time for the accumulation of citations, given that our publication data began in 1990. Similar to the department allocation for each author, we determine the department for each star as the department in which they had the most publications. The star scientist identified is relative to scientists in Denmark, Ireland, and New Zealand. Hence, these stars are highly productive scientists in the context of the overall distribution of scientist citations for these three countries. Through this method, we identify 981 stars from the dataset, 205 in Ireland, 548 in Denmark, and 228 in New Zealand.

3.2 Identification of star names in the acknowledgements

We identify star names and their author interaction type from the pool of available acknowledgement text. Since the study focuses on the impacts of star scientists on their peers through informal collaborations, we think of the influence of these interactions as occurring through intense discussions, critical reviews, data sharing, collaborations, and supervision. Broad forms of funding texts are available in Scopus, which is translated into acknowledgement texts. A limitation of the Scopus database is that the acknowledgement text is not available for every publication. Table 14 in the "Appendix" outlines the yearly distribution of publications and available acknowledgement texts in Scopus.

Using a Natural Language Processing (NLP) tool, we extract the names[2] of scientists from our acknowledgement text data. We use the open-source library, Spacy, an advanced

[1] The 27 subject fields are Agricultural and Biological Sciences, Arts and Humanities, Biochemistry, Business, Chemical Engineering, Chemistry, Computer Science, Decision Science, Dentistry, Earth and Planetary Sciences, Economics, Energy, Engineering, Environmental Science, Health Professions, Immunology and Microbiology, Multidisciplinary, Materials Science, Mathematics, Medicine, Neuroscience, Nursing, Pharmacology, Physics and Astronomy, Psychology, Social Sciences, and Veterinary.

[2] We provide the procedure for identifying star names and the type of star help from the acknowledgement texts in the "Appendix".

NLP tool in Python designed to process a large volume of text data. More specifically, we utilise Spacy's Named Entity Recognition (NER) feature to extract named entities from each acknowledgement text. Named entities are phrases that contain the names of persons, organisations, and locations. Spacy's NER feature has reported an accuracy of 89.8 (OntoNotes[3]) and 91.6 (CoNLL-2003 corpora[4]) percent and compares well to other packages available for NLP. The tool helps to token the words in an acknowledgement text, and we group them using a unique publication identifier. Using the NER feature, we tag the author entities in each publication's acknowledgement text, and then we match these names to those of the star scientists identified in Sect. 3.1. Overall we identify 331 stars acknowledged over 971 publications by 1815 authors.

3.3 Identification of acknowledgement types

While prior studies primarily focus on the typology of the overall acknowledgement text (Cronin, 1991; Desrochers et al., 2018; Paul-Hus et al., 2016, 2017), we focus on the types of interaction between a star and an author. Using a similar approach as Oettl (2012) we identify the type of star help based on their helpfulness. Moreover, these acknowledgement types are classified in our study based on the available keywords that determine the star help in our data. Here we define five types of acknowledgements to a star who is not a co-author: Conceptual, Technical, Materials, Funds & Support, and Other Types. Each acknowledgement type indicates the interaction between an author and a star during the research project.

Conceptual acknowledgements recognise star scientists for their intellectual contribution to the research project. For example, "We thank P. Di Vecchia, G. Grignani, C. Kristjansen, and N. Obers for useful discussions" (Harmark and Orselli, 2006) and "I would like to thank Prof. P. Sigmund for stimulating and enlightening discussions of the topic and useful comments on the manuscript" (Glazov, 1998). Using keywords, we identify this type of interaction between a star and the author. Here, keywords such as "discussions," "comments," "feedback," "critique," "advice," "suggestions," are used. Keywords that categorise help in the technical form are "technical assistance," "statistical assistance," "excellent assistance," and "expert assistance." For example, "We thank P. Rasmussen, B. Jensen, and D. Bardenfleth for expert technical assistance" (Sorensen et al., 2010). In this example, P. Rasmussen is the star, and the author acknowledges his help through their technical assistance.

Material acknowledgement shows the debt of gratitude for materials and data shared by the star for the research purpose. Keywords used to identify interactions based on this specification are "data sample," "antibodies," and "cells." For example, "Professor Klaus Bendtzen is thanked for providing antibodies for the cytokine measurements" (Theander et al., 1997). Another acknowledgement type captures the author's "funding and supports" through the star interaction. Keywords such as "grants," "funds," "support," and

[3] OntoNotes project is a collaborative effort between BBN Technologies, the University of Colorado, the University of Pennsylvania, and the University of Southern California's Information Sciences Institute. The goal of the project was to annotate a large corpus comprising various genres of text (news, conversational telephone speech, weblogs, USENET newsgroups, broadcast, and talk shows) in three languages (English, Chinese, and Arabic) with structural information (syntax and predicate argument structure) and shallow semantics (word sense linked to an ontology and coreference).

[4] CoNLL-2003 shared task: concerns language-independent named entity recognition. They concentrate on four types of named entities: persons, locations, organizations, and names of miscellaneous entities that do not belong to the previous three groups (Tjong Kim Sang, Fien De Meulder, 2003).

Table 1 Acknowledgement types that define the types of a star help

Subject Field	Conceptual	Technical	Materials	Funds & Support	Other Types	Total
Agricultural and Biological Sciences	177	105	53	29	58	422
Arts and Humanities	2	0	1	0	0	3
Biochemistry	234	113	104	37	37	525
Business	22	0	1	2	2	27
Chemical Engineering	8	0	0	0	3	11
Chemistry	51	40	14	3	24	132
Computer Science	10	0	0	4	5	19
Dentistry	5	6	6	0	7	24
Earth and Planetary Sciences	83	16	34	6	18	157
Economics	25	3	2	1	1	32
Energy	2	2	0	0	3	7
Engineering	73	11	1	3	22	110
Environmental Science	17	20	2	5	9	53
Health Professions	1	0	0	0	0	1
Immunology and Microbiology	18	28	17	4	3	70
Materials Science	13	14	12	0	13	52
Mathematics	13	2	0	1	3	19
Medicine	157	133	142	96	509	1,037
Multidisciplinary	1	4	2	0	0	7
Neuroscience	44	23	19	5	11	102
Pharmacology	30	9	1	2	3	45
Physics and Astronomy	157	26	32	11	32	258
Psychology	14	1	0	4	8	27
Social Sciences	15	5	5	5	7	37
Veterinary	1	14	3	2	1	21
Total	1,173	575	451	220	779	3,198

Distribution of acknowledgement types over the subject fields based on the interaction with a star. The table includes multiple acknowledgements from 1815 unique authors over 971 publications. The authors acknowledge a total of 331 stars from 1990–2017

"financial." are used to identify the interaction with the stars. For example, "This research was sponsored through the contestable research fund of the Waikato Management School of the University of Waikato and Professor Chris Ryan, Waikato Management School, for his continued support" (Lockyer, 2005). Finally, acknowledgements using the keywords such as; "committee," "contributions," "permission," "facilities," "director," "supervisor," "founder," "help," "dedication," are classified as "Other Types." For example, "The study was initiated by Torben Jørgensen, DMSc (PI); Knut Borch-Johnsen, DMSc (Co-PI); Troels Thomsen, Ph.D., and Hans Ibsen, DMSc" (Baumann et al., 2015). A similar approach is adopted by Paul-Hus and Desrochers (2019) to classify acknowledgement types as 'vague.'

Table 1 presents the variation across the different types of acknowledgements by 1815 authors to 331 stars over the period 1990 to 2017. Star help classified as the conceptual type constitutes 37% of the overall acknowledgements, where the subject areas: Agricultural and Biological Sciences; Biochemistry; Medicine; Physics; and, Astronomy account for 62% of conceptual acknowledgements to a star scientist. Star help based on technical acknowledgements occur primarily in the subject areas: Agricultural and Biological Sciences, Biochemistry, and Medicine, where the research in these departments can include considerable laboratory work as well as the use of complex instruments. Material acknowledgements account for 14% of the general acknowledgements and are mainly observed in the fields of Medicine and Biochemistry. Only 7% of the acknowledgements to star scientists in our sample show an acknowledgement to a star for the funding the author received.

3.4 Output measure

We use two output measures to test our hypotheses discussed in Sect. 2.2: (i) Field Normalised Total Citations (FNTC)—the sum of an individual's publication citations divided by the average citations to a publication for that subject field in that year for all countries combined; and (ii) total count of normalised raw publications of the individual author in that year. We calculate the dependent variable as follows:

$$\text{Field Normalized Total Citations}: Y_{i,t}^{FNTC} = \sum_{p_{i,t}=1}^{P_{i,t}} \frac{c_{p_{i,t}}}{\overline{c}_{s,t}};$$

$$\text{Publications}: Y_{i,t}^{P} = \frac{P_{i,t}}{\overline{P}_{s,t}};$$

where $P_{i,t}$ is the total number of publications by individual i published in year t, $c_{p_{i,t}}$ are the subsequent total citations (or "forward" citations recorded in 2019) to a publication $p_{i,t}$ that occur for individual i in year t, $\overline{c}_{s,t}$ is the average citations to a publication in the relevant subject field, s, for publications that occur in year t, and $\overline{P}_{s,t}$ is the average number of publications in subject field s.

These output measures provide the dependent variables for our regressions. Using both dependent variables, we use the author's publication data from 1996–2017 (our estimation window) to calculate the overall impact of the star's helpful interaction. In our analysis, we split the individuals into two subgroups—Treated: Authors who acknowledge the star for the helpful interaction for their publications and who did not co-author with the star before the acknowledgement; and Never-Treated: Authors who received neither helpful interaction nor a star co-authorship interaction. Section 3.5 below discusses the matching procedure to identify matched pairs of treated and control authors before the treatment event happens.

3.5 Coarsened exact matching procedure (CEM)

Matching treated units to control units in observational data helps to mitigate the confounding influence of pretreatment control variables with the primary goal of improving the balance between the treated and control groups. We, therefore, employ matching to help control for the endogeneity issue that might arise in acknowledging a star. For example, acknowledging a star can be considered random in the cases such as tips and help received from a seminar series or international conferences. At the same time, prior connections with the author also involve the star being a part of the research as an informal contributor. To address this, we create a panel data which has a control group comprising authors who never received any star exposure in terms of star's helpful interaction or co-authorship and

Table 2 Summary statistics: Control and Treated group (k-to-k matched)

Variable	Control	Treated	Diff in mean	P-value
An unbalanced panel of 1258 matched authors, with 629 in each group				
Year	2007.122	2007.099	0.024	0.944
Subject	10.906	10.906	0.065	0.672
Country	1.988	1.986	0.000	1.000
Total Career Age	20.541	20.655	-0.114	0.841
Total Career Age Bins	5.730	5.730	0.000	1.000
Cumulative Publication experience	9.068	9.049	0.019	0.959
Cumulative Publication experience Bins	2.366	2.366	0.000	1.000
Cumulative citations received per cumulative publications	44.226	44.769	-0.543	0.954
Cumulative citations received per cumulative publications Bins	2.035	2.035	0.000	1.000

Reports the t-test for the mean difference between control and treated groups one year before forming the star interaction

a treated group—a similar set of authors in terms of characteristics who received star help that is identified from the acknowledgement texts of the publications.

In our study, we ensure that the treated authors had no star interaction in terms of star co-authors before or during the event of the star's helpful interaction. To minimize the effect of confounding in our observational causal inference, we employ Coarsened Exact Matching (CEM), which is a monotonic imbalance-reducing matching method (Blackwell et al., 2009; Iacus et al., 2011; Iacus et al., 2012). CEM is a design strategy that involves matching on a set of covariates that have been "coarsened," meaning that they have fewer possible values for matching, which increases the number of matches (King et al., 2019, 2011). This technique has been shown to improve the balance of covariates between exposure groups and to guarantee balance for each covariate, only limited by the coarseness of the grouping (Fini et al., 2018). In contrast, other matching techniques, such as Propensity Score Matching (PSM), do not guarantee each variable such a balance guarantee. They may require repeated iterations to achieve balance (Fini et al., 2023). Moreover, CEM ensures balance for higher-order terms, such as interactions of covariates, while such a guarantee does not exist in a propensity score approach.

We find an author in the control group at $t-1$ year who matches similar characteristics (based on the observables) with an author from the treated group at $t-1$ who receives star help at year t. The CEM procedure allows us to define the covariates to match a categorical variable rather than a continuous variable. We identify a match based on five covariates: subject field, country, total career age, cumulative publication experience, and cumulative citations received on prior publications. Each covariate uses a categorical variable with course bins[5] (Iacus et al., 2012). Table 2 shows the difference in the mean value and

[5] We create twenty-seven bins for the subject fields and three for the countries. Eleven bins for the total career age of the scientist ranging from 4 to 50 years; less than 5 years, between 5 and 10 years, between 10 and 15 years, between 15 and 20 years, between 20 and 25 years, between 25 and 30 years, between 30 and 35 years, between 35 and 40 years, between 40 and 45 years, between 45 and 50 years, 50 years and above. Cumulative publications experience captures cumulative years since the first year of publication (here in this data from 1990 onwards). Seven bins range from 0 to 1 year, 1 to 5 years, 5 to 10 years, 10 to 15 years, 15 to 20 years, 20 to 25 years, and 25 and above. Finally, cumulative citations received per cumulative

Table 3 Dynamic star help effects for field normalised total citations

	Homogenous star help effects (1)	Heterogeneous star help effects (2)
$Staracknw_{i,t\text{-}3}$	− 0.00829	− 0.0105
	(0.0255)	(0.0256)
$Staracknw_{i,t\text{-}2}$	− 0.0226	− 0.0205
	(0.0240)	(0.0242)
$Staracknw_{i,t}$	0.247***	0.2463***
	(0.0243)	(0.0244)
$Staracknw_{i,t+1}$	0.0432*	0.0505**
	(0.0255)	(0.0256)
$Staracknw_{i,t+2}$	0.0450*	0.0581*
	(0.0263)	(0.0267)
$Staracknw_{i,t+3}$	0.0344	0.0432
	(0.0278)	(0.0285)
Constant	0.402***	0.460***
	(0.0237)	(0.00916)
R-squared	0.029	0.521
Pretest against the hypothesized trend (Roth, 2022)		
Power	0.50	0.50
Hypothesized trend	0.02	0.02
Bayes factor	0.55	0.55
Likelihood ratio	0.30	0.38
Observations	20,451	20,451
Number of authors	1,238	1,238
Author FE	YES	YES
Year FE	YES	YES

Event Window: 1997–2017 (*1996 cohort dropped for comparison only*). Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in the Eq. 1 and Eq. 3. Column 1 reports the homogenous star-help effects at the individual level. Column 2 reports the estimates using Sun and Abraham's (2021) method. Also, the 1996 cohort is dropped to compare the two methods, although it does not affect the final estimated coefficients since only ten authors are treated from 1996. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

p-value of each matching variable for both control and treated groups one year before the star interaction occurs. The unbalanced matched panel data set contains 1258 authors, with 629 authors in the treated group who acknowledge a star for the first time. Furthermore, 75

Footnote 5 (continued)

publication are categorised based on the distribution of resulted average value, which shows the citations received by the author per year based on the number of publications. Sixteen bins are created for values less than 1, between 1 to 25, 25 to 50,50 to 75, 75 to 100, 100 to 150, 150 to 200, 200 to 250, 250 to 300, 300 to 350, 350 to 400, 450 to 500, 500 to 750, 750 to 1000, and 1000 and above.

of these authors acknowledge the stars in the years after the first acknowledgement. *Similarly, for the five cohorts based on acknowledgement types: Conceptual,* Technical, Material, Funds & Support, and Other Types, we use the CEM matching procedure to identify matching pairs of control and treated author. Table 16 in the "Appendix" reports the matching statistics of matched pairs for each cohort.

4 Econometric methodology

The empirical goal of our econometric analysis is to estimate the effects of helpful interactions with star scientists on the productivity of authors over time, where we capture the interactions from the acknowledgement texts of the author's research publications. As we discussed, a star might impact the productivity of their peers through various channels, and our focus is to identify whether an author's informal non-co-authorship interactions with a star could affect their productivity measured in terms of (forward) citations to their publications (hypothesis 1) & raw publications output (hypothesis 2).

To measure these dynamic effects, we utilise an 'event-study' specification where the event is an acknowledgement to a star scientist that can occur in the past or future. Our event-study design is a staggered adoption design where units (authors) are treated at different times, and some units have never been treated. We estimate the dynamic treatment effects of the helpful interaction with a star on their peers from three years before the published acknowledgement to three years afterwards in the following event study specification:

$$\begin{aligned} lnY_{it} =& \alpha + \beta_{\leq -4} staracknw_{i,-4} + \sum_{j=-3}^{-2} \beta_j staracknw_{ij} + \sum_{j=0}^{3} \beta_j staracknw_{ij} \\ &+ \beta_{\geq 4} staracknw_{i,4} + \delta_t + \mu_i + \epsilon_{it}, \end{aligned} \tag{1}$$

where the dependent variable lnY_{it} is a measure of the citation-weighted/normalised raw publication output of author i at year t, $staracknw_{ij}$ is a binary variable equal to 1, if a star is acknowledged by the author i as of year t,j years ago, δ_t is a year fixed effect, μ_i is an author fixed effect, and ϵ_{it} is a zero mean error term. The coefficients of interest, β_j, show the proportionate effect of the helpful interaction with the star on productivity from three years before the acknowledgement to three years afterwards. We normalise the effect of the star interaction to zero for the year before the star acknowledgement, and we assume that the cumulative effect is constant at $\beta_{\leq -4}$ and $\beta_{\geq 4}$ by binning at four leads and four lags. The binning variables may not be comparable to the leads and lags of the acknowledgement binary variables in estimating the dynamic effects since they could be correlated with other excluded level variables; however, they act as essential controls in our specification (Schmidheiny and Siegloch 2019). Furthermore, standard errors are clustered at the author level and are robust to arbitrary forms of serial correlation and heteroscedasticity. McHale et al. (2022) adopt a similar econometric approach that estimates the effects of star arrival on the departments' productivity in SOEs.

An important assumption of our event study model is the generalised form of the parallel trends assumption, whereby without a star acknowledgement, the quality-adjusted output in the treatment group would have changed in the same way as it did in the non-treatment group (the authors that did not acknowledge a star scientist). The estimated lead coefficients in our specification allow us to examine indirect evidence to support the parallel trends assumption, with any observed pre-acknowledgement effects considered evidence of a failure of this assumption. On the other hand, an observed pre-trend could indicate anticipation effects of star acknowledgement on productivity; for example, if an author had prior knowledge of possible helpful interaction with a star in the future, they might change their productivity behaviour. However, in our case, this is considered to be unlikely since the author acknowledges a star for his helpfulness which could not be anticipated before the interaction, and therefore, we assume that anticipation effects are zero.

Another critical assumption in this event-study setting is that the star acknowledgement effect on author productivity is homogeneous across the timing of the acknowledgements. However, recent literature has shown that the coefficients of given leads and lags can be contaminated by the effects from other periods in the presence of heterogeneous effects across different treatment timings (Callaway and Sant'Anna, 2021; Sun & Abraham, 2021; Goodman-Bacon, 2021). In our econometric analysis, star acknowledgement has a staggered treatment timing, and heterogeneity in the effects could arise if different cohorts experience different treatment paths. Therefore, we also adopt the approach of Sun and Abraham (2021) to estimate the heterogeneous treatment effects. This approach derives the dynamic effects of star acknowledgement in a three-step estimation that is robust to treatment effect heterogeneity and calculates a weighted average of 'cohort average treatment effects on the treated' (CATT). First, we define the year in which an author i acknowledges a star scientist as e_i. Second, we estimate the weighted average of cohort effects for a given time relative to the acknowledgement event. To allow the estimated star acknowledgement effects to vary by cohort based on the year that the acknowledgement event occurs, we estimate the following equation:

$$\begin{aligned} lnY_{i,t} = \sum_e \Bigg[& \delta_{e,-4}(1\left\{E_i = e\right\}staracknw_{i,-4}) + \sum_{j=-3}^{-2} \delta_{e,j}(1\left\{E_i = e\right\}staracknw_{ij}) \\ & + \sum_{j=0}^{3} \delta_{e,j}(1\left\{E_i = e\right\}staracknw_{ij}) + \delta_{e,4}(1\left\{E_i = e\right\}staracknw_{i,4}) \Bigg] + \delta_t + \mu_i \end{aligned} \tag{2}$$

where $1\left\{E_i = e\right\}$ is an indicator variable that takes the value 1, if the individual i receives star help in the year e and 0 otherwise. $\delta_{e,j}$ is the star help effect on productivity j year after author acknowledges a star in year e. The 1996 treated cohort is dropped from the analysis since it is always treated across the observation window. A further set of weights are estimated $Pr\left\{E_i = e|E_i\epsilon[-j, T-j]\right\}$ that are equal to sample shares of each cohort for the relevant periods j. Finally, to obtain the IW estimator, we take a weighted average of the $\widehat{\delta}_{e,j}$ (or $CATT_{e,j}$) and estimate Eq. 2 with relevant weights calculated.

$$\beta_j^* = \sum_e \left[\widehat{\delta}_{e,j} Pr\left\{E_i = e|E_i\epsilon[-j, T-j]\right\}\right]. \tag{3}$$

5 Results

In this section, we present the results for our different hypotheses outlined in Sect. 2.2. We use FNTC as the dependent variable for testing hypotheses 1a and 2a and the number of raw publications for testing hypotheses 1b and 2b. Furthermore, we only use FNTC as the

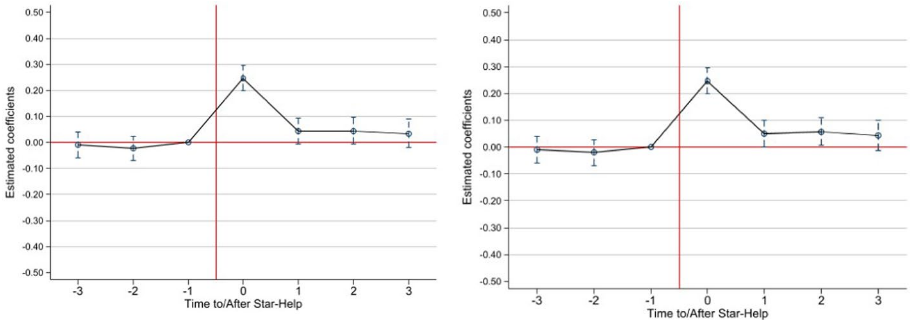

Fig. 2 Event study model with homogenous (left) vs. heterogeneous (right) star help effects at an individual level. The figure plots the dynamic effect of the star's helpful interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017 (the 1996 cohort dropped is omitted). The dependent variable is field normalised total citations

dependent variable to test hypothesis 3, where we explore the effect of star help across the types of helpful interaction identified.

5.1 Star help effect on individual output (with FNTC as dependent variable)

We examine hypotheses 1a and 2a in this section. We present the estimated results from the event study specification for both the homogenous and heterogeneous star acknowledgement effects in Table 3. The dependent variable is FNTC to the output published in year t. Column 1 of Table 3 shows the results for the homogeneous model, and the star acknowledgement effects are also shown on the graph in the left panel of Fig. 2. We find evidence that a helpful interaction with a star, as captured by a star acknowledgement, significantly affects an author's productivity supporting hypothesis 1a. A star acknowledgement is associated with an economically and statistically significant contemporaneous increase in the quality-adjusted output of the author of 24.70 log points, which translates to a 28.02% increase in output. In the years after the star acknowledgement, the estimated coefficients remain positive but decrease substantially to between 0.04 and 0.05 log points and are only statistically significant at the 10% level.

Regarding the heterogeneous model using the approach of Sun and Abraham (2021), the event study results are reported in Column 2 of Table 3 and displayed in the right panel of Fig. 2. Under the heterogeneous model, we find a similar contemporaneous effect as in the homogenous model, with a 28% increase in the quality-adjusted output of the author in the year of acknowledgement. Moreover, the results lend further support to hypothesis 1a when the differential timing of treatment is taken into account in our estimation. Furthermore, the estimated coefficients for the years following a star acknowledgement are again similar to the homogenous model (0.05 to 0.06 log points), and these coefficients are statistically significant at the 5% and 10% levels, respectively. In considering the staggered timing of treatment on authors that have a helpful interaction with a star, the homogenous results are robust to the heterogeneous specification in Eq. 3. Overall, the results suggest that the quality-adjusted output of an author increases contemporaneously when a star is acknowledged, but the effect tends to fall in subsequent years. We also dropped the 1996 (always treated) cohort from the homogenous estimation for comparison purposes, which has minimal effects on the final results. Overall, we find supporting evidence for hypothesis

Table 4 Dynamic star help effects (multiple and one-time) for field normalised total citations

	Multiple acknowledgements to stars (1)	One-Time acknowledgement to stars (2)
$Staracknw_{i,t\text{-}3}$	− 0.0891	− 0.00244
	(0.0757)	(0.0270)
$Staracknw_{i,t\text{-}2}$	− 0.000964	− 0.0262
	(0.0762)	(0.0252)
$Staracknw_{i,t}$	0.348***	0.232***
	(0.0634)	(0.0260)
$Staracknw_{i,t+1}$	0.201***	0.0269
	(0.0686)	(0.0274)
$Staracknw_{i,t+2}$	0.237***	0.0145
	(0.0736)	(0.0276)
$Staracknw_{i,t+3}$	0.239***	0.00363
	(0.0838)	(0.0289)
Constant	0.541***	0.387***
	(0.0705)	(0.0243)
R-squared	0.055	0.028
Pretest against the hypothesized trend (Roth, 2022)		
Power	0.50	0.50
Hypothesized trend	0.07	0.03
Bayes factor	0.55	0.55
Likelihood ratio	1.68	0.19
Observations	2,681	18,160
Number of authors	150	1,108
Author FE	YES	YES
Year FE	YES	YES

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in the Eq. 1. Column 1 reports the subsequent star-help effects at the individual level after the first year of star acknowledgement. Column 2 reports the first-time star acknowledgement without any subsequent interaction. Also, the 1996 cohort (always treated) is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

(1a) that there is a contemporaneous increase in the citation-weighted publications from star help.

One advantage of an event study setting is that it allows visual evidence to support or contradict the parallel trends assumption. Our results in Fig. 2 depicts that both the homogeneous and heterogeneous models show very little evidence of a pre-trend; therefore, visual inspections support the parallel trends assumption. However, Roth (2022) raises some important concerns about relying on insignificant pre-trends in the event study setting to assess the credibility of parallel trends. In particular, the test may have low power to detect

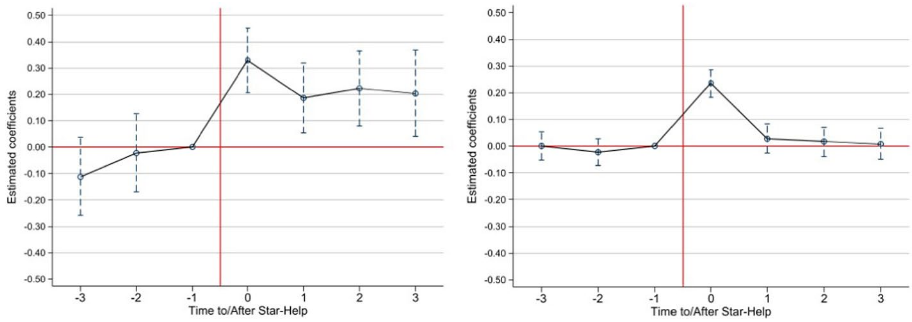

Fig. 3 Event study model with multiple (left) vs. one-time (right) star help effects at an individual level. The figure plots the dynamic effect of subsequent and first-star helpful interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

meaningful violations of parallel trends. Roth (2022) suggests a diagnostic approach to determine whether such concerns could be warranted, and we implement these diagnostics in our analysis.

Roth's (2022) technique involves constructing a hypothesised linear violation of parallel trends and then comparing the likelihood of the observed coefficients under the hypothesised trend relative to under-parallel trends. We construct a hypothesised trend with a power of 50%, which would detect a significant pre-trend 50 percent of the time. With a power of 50%, the slope estimate is found to be 0.02 for both the homogenous and heterogeneous effects models. The likelihood ratio reporting the ratio of the likelihood of the observed coefficients under the hypothesised trend relative to under-parallel trends is 0.30 and 0.38, respectively, which supports that the estimated coefficients are more likely observed under a parallel trend. The results of this diagnostic test are also shown in Table 3.

Next, we examine hypothesis 2a, which proposes that the effect of star help has a sustained increase in the quality-adjusted output of the author. Table 4, Column 1, reports the results based on the sample of the author's productivity that continues to acknowledge a star after the first point acknowledgement. After their initial interaction, the authors who maintain this relationship with a star show a substantial increase in their quality-adjusted publications. Their output is found to increase by 41.62% in the year of acknowledgement and is statistically significant at 1%. These results show the importance of maintaining the relationship with a star. The estimated coefficients decrease in the years after the first acknowledgement (20.12, 23.74, 23.94 log points) but are significant and persistent.

Additionally, the event study plot in Fig. 3 left panel shows no evidence of a pre-trend as well as a strong help effect on productivity up to 3 years after the first acknowledgement in the case where multiple acknowledgements take place. This is in contrast to an increase in productivity that is broadly limited to the year of acknowledgement in the right hand panel, which limits the sample to cases of a single acknowledgement. Overall, the evidence suggests that once-off help is associated with a time-limited increase in productivity (supporting hypothesis 1a), and that sustained help is associated with a sustained increase in productivity (supporting hypothesis 2a).

Table 5 Dynamic star help effects (overall, multiple, one-time) for raw publications

	Overall star-help effect on individual productivity (1)	Multiple star-help effect (2)	One-time star-help effect (3)
$Staracknw_{i,t-3}$	– 0.00799	– 0.00407	– 0.0113
	(0.0192)	(0.0622)	(0.0202)
$Staracknw_{i,t-2}$	– 0.0227	– 0.0280	– 0.0225
	(0.0181)	(0.0566)	(0.0191)
$Staracknw_{i,t}$	0.202***	0.251***	0.194***
	(0.0159)	(0.0469)	(0.0168)
$Staracknw_{i,t+1}$	0.0340*	0.173***	0.0118
	(0.0192)	(0.0496)	(0.0207)
$Staracknw_{i,t+2}$	0.0218	0.134**	0.00301
	(0.0207)	(0.0617)	(0.0218)
$Staracknw_{i,t+3}$	0.0220	0.110*	0.00604
	(0.0213)	(0.0581)	(0.0229)
Constant	0.447***	0.589***	0.424***
	(0.0186)	(0.0503)	(0.0199)
R-squared	0.048	0.064	0.049
Pretest against the hypothesized trend (Roth, 2022)			
Power	0.50	0.50	0.50
Hypothesized trend	0.02	0.02	0.02
Bayes factor	0.55	0.55	0.55
Likelihood ratio	0.38	0.26	0.52
Observations	20,841	2,681	18,160
Number of authors	1,258	150	1,108
Author FE	YES	YES	YES
Year FE	YES	YES	YES

Estimation Window: 1996–2017. The dependent variable is raw publications

The table reports the estimates based on the model specification in Eq. 1. Also, the 1996 cohort (always treated) is not dropped in this case.Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

5.2 Star help effect on individual output (with raw publications as dependent variable)

In our estimation above, we examine the effect of star help on the quality-adjusted productivity of authors, where we use the field-weighted citations as our dependent variable to study this effect. However, there remains a concern that the positive effect of observed star acknowledgement on citation-weighted output found above could reflect a signalling effect, whereby acknowledgement to a star scientist is used as a signal or indicator to convey information about the potential quality of the publication. For the reasons outlined in Sect. 2.2, we augment our analysis using FNTC with an analysis using field-normalised total raw publications to help disentangle the input-enhancing role implied by observed star acknowledgements from a signalling effect. Based on our assumption that the blinded manuscript submitted to a journal for peer review contains no acknowledgement texts

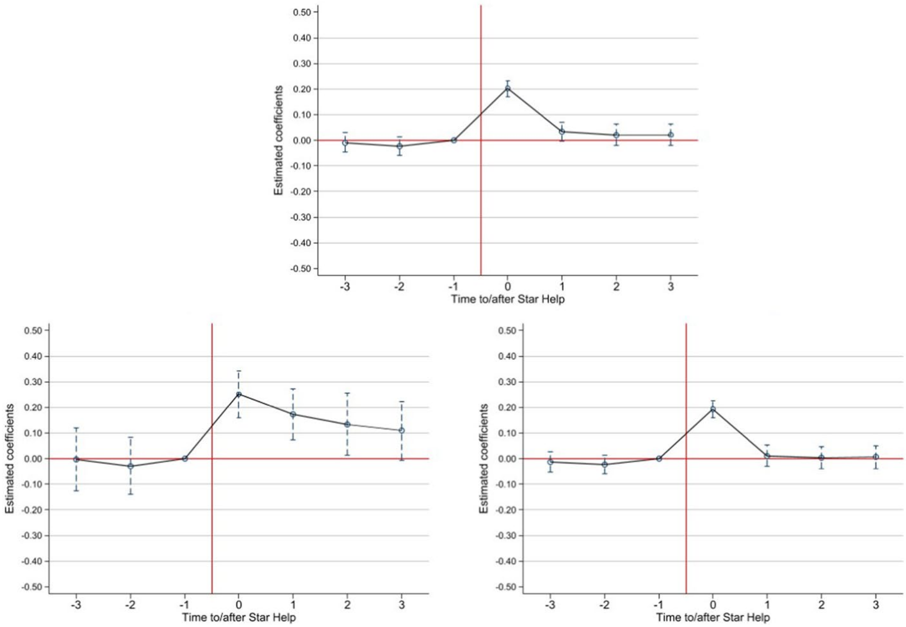

Fig. 4 Event study model with homogenous star help effects at an individual level: Overall Star Help (Row 1); Multiple Star Help (Row 2, Left); One-Time Star Help (Row 2, Right). The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is raw publications

and, thus, the decisions by the reviewers cannot be impacted by the acknowledgement to a prominent star scientist, we now present the results from testing hypothesis 1b and 2b using raw publications count.

Column 1 in Table 5 reports the estimated coefficients for the baseline regression using authors' field normalised raw publications as the dependent variable. We find that in the year of the acknowledgement event, the raw publication output of authors increased by 22.38% compared to the authors who never acknowledged any help from a star. We also find a decrease in the magnitude of the star-help effect on the raw publications output of the author. These results support hypothesis 1b, where we propose that there will be an increase in the raw publications output in the year of star help. However, similar to our findings for FNTC, we do not find any evidence of a significant effect on raw publications output in the years after the event.

For hypothesis 2b, the results based on the treated sample of authors that acknowledge a star on one occasion and those that acknowledge a star on multiple occasions are presented in Table 5, Columns 2 and 3, respectively. Regarding the authors that acknowledge the star for help in the years after the initial event, the results suggest a sustained effect on raw publications output broadly similar to our analysis with field-weighted citation as the dependent variable. An acknowledgement of star help is associated with a 28.53% increase in raw publications in the event year, and the coefficients are found to be statistically significant in the years after. This provides evidence to support our hypothesis that there is a sustained increase in the raw publications output for authors who maintain an informal collaboration with a star.

Table 6 Dynamic star help effects for field normalised total citations for five types of helpful interactions

	Conceptual (1)	Technical (2)	Materials (3)	Funds & Support (4)	Other Types (5)
$Staracknwtype_{i,t\text{-}3}$	− 0.0488	0.0121	0.0704	− 0.00899	− 0.0211
	(0.0394)	(0.0448)	(0.0756)	(0.202)	(0.0551)
$Staracknwtype_{i,t\text{-}2}$	− 0.0529	0.0106	− 0.00552	0.0200	− 0.0299
	(0.0412)	(0.0486)	(0.0589)	(0.127)	(0.0496)
$Staracknwtype_{i,t}$	0.316***	0.172***	0.183***	0.306**	0.253***
	(0.0432)	(0.0404)	(0.0626)	(0.124)	(0.0569)
$Staracknwtype_{i,t+1}$	0.0589	0.0528	0.0937	0.0822	− 0.0172
	(0.0453)	(0.0436)	(0.0703)	(0.121)	(0.0581)
$Staracknwtype_{i,t+2}$	− 0.00485	0.0224	0.191**	0.138	0.0428
	(0.0426)	(0.0434)	(0.0746)	(0.137)	(0.0661)
$Staracknwtype_{i,t+3}$	0.0627	0.0333	0.0947	0.0505	− 0.0559
	(0.0502)	(0.0495)	(0.0704)	(0.113)	(0.0616)
Constant	0.365***	0.379***	0.315***	0.310***	0.357***
	(0.0345)	(0.0388)	(0.0545)	(0.107)	(0.0578)
R-squared	0.046	0.021	0.033	0.047	0.048
Pretest against the hypothesized trend (Roth, 2022)					
Power	0.50	0.50	0.50	0.50	0.50
Hypothesized trend	0.04	0.04	0.07	0.18	0.05
Bayes factor	0.55	0.55	0.55	0.55	0.55
Likelihood ratio	1.40	0.08	0.03	0.19	0.43
Observations	7,665	4,942	3,001	947	4,152
Number of authors	468	300	176	60	248
Author FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in Eq. 4. Also, the 1996 cohort (always treated) is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

In contrast, in considering the effect of star help on the raw publications output for the authors who acknowledge a star for help in a single event, the results are aligned with our prior analysis using field-weighted citations as the dependent variable in that we do not find any evidence for a statistically significant effect in the years after the event. Figure 4 illustrates the event study plots for the two cases. Overall, by separating the input-enhancing role implied by observed star acknowledgements from a signalling effect using the raw publications measure, the results indicate that the positive effect of the observed star acknowledgement on output reflects more than just a signalling effect in both the cases of once-off acknowledgement and multiple acknowledgements.

5.3 Effects of star help through different channels

In this Section, we consider hypothesis 3, which states that the effect of star help is different in magnitude across the different types of star help identified based on the keywords contained in the acknowledgement text for the publication. We identify five types of acknowledgement to understand the effects of the helpful interaction between the author and the star: Conceptual, Technical, Material, Funds and Support, and Other Types. Our dataset is split into five cohorts, each containing treated and control authors identified from the CEM procedure. This analysis estimates the effect of a star help that happens through five different channels.

For each channel, we modify Eq. (1)

$$\begin{aligned} Y_{it} =& \alpha + \beta_{j\leq -4} staracknwtype_{i,-4} + \sum_{j=-3}^{-2} \beta_j staracknwtype_{ij} + \sum_{j=0}^{3} \beta_j staracknwtype_{ij} \\ &+ \beta_{j\geq 4} staracknwtype_{i,4} + \delta_t + \mu_i + \epsilon_{it} \end{aligned} \tag{4}$$

where $staracknwtype_{ij}$ is a binary variable that indicates if a star is acknowledged for the specific type of help in year j by an author i.

The results from the estimates are presented in Table 6. They show that a conceptual acknowledgement to a star is associated with an increase in output of 37.16%, and this is the most significant increase in output compared to all other types examined in this analysis. The help from the star through sharing their knowledge positively impacts the author's productivity. In contrast, a technical acknowledgement type is associated with the smallest effects on output, with an increase of 18.77%. This acknowledgement type is observed primarily in subject fields such as medicine and biochemistry, where the research methods involve practical examination and laboratory experiments, and stars can provide their direct expertise. Material acknowledgement to a star is associated with an increase in output of 20.08% in the year of acknowledgement and unlike the other types of help, there was a 21.05% increase in output in the year afterwards. Also, the event study plot for material acknowledgement in Fig. 5 shows a persistent productivity rise in the years after star acknowledgement, which contrasts with the absence of a sustained effect for the other types of acknowledgement. Material acknowledgements to a star are observed mainly in the fields of medicine and biochemistry, where stars can conveniently share unpublished data and loan specimens. This could suggest that the star and the author develop a more persistent connection through the materials-sharing channel than through the other acknowledgement channels.

A project funded by the star's support also impacts positively on an author's productivity, with a funding acknowledgement to a star found to be associated with a 35.80% increase in output in the year of acknowledgement. These findings suggest that a star who supports the research through financial means also has an effect on the author's productivity. The number of these acknowledgements are fewer when compared to other types of interaction channels in our data. Finally, all the other types of acknowledgement that could not be accurately classified under conceptual, technical, materials, or funds and support are categorised as 'Other Types' for the purpose of this study. In the year of these type of acknowledgement events, the author's quality-adjusted output increases by 28.79%, similar to our overall star acknowledgement effect in Sect. 5.1. These results suggest that the 'Other' category proxies for the other more concrete forms of help the stars provide.

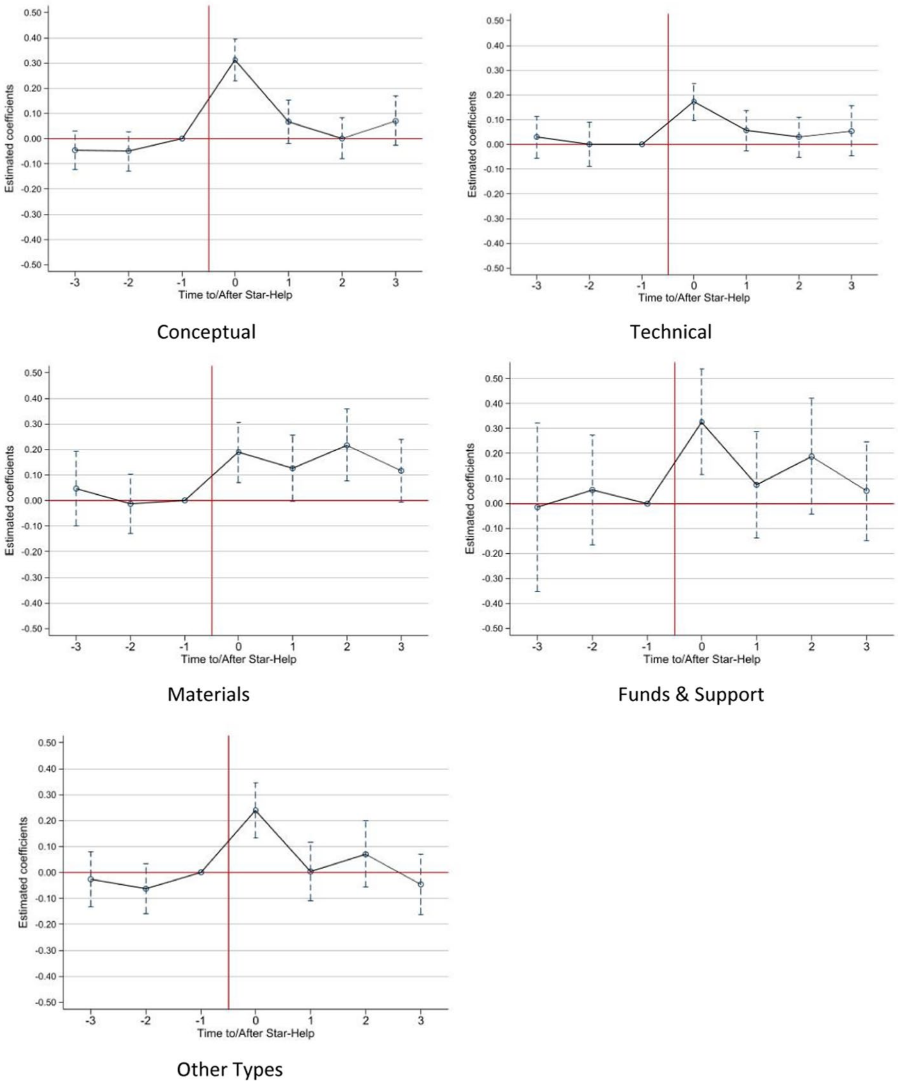

Fig. 5 Event study model with channels of help identified from the acknowledgement texts. Conceptual (row 1, left), Technical (row 1, right), Material (row 2, left), Funds & Support (row 2, right), and Other types (row 3). The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

Alternatively, the relative robustness of the results across acknowledgement types might indicate that the precise form of help is less important than the close interaction with the star. Overall, these results provide evidence to support hypothesis 3 that the effect of star

Table 7 Summary statistics: Control and Treated group (k-to-k matched) (a) a balanced panel, (b) alternate matching criteria using the availability of publication match at t = 0, (c) alternate matching criteria using acknowledgement text availability in publication at t = 0

Variable	Control	Treated	Diff in mean	P-value
(a)				
A balanced panel of 520 matched authors, with 260 in each group				
Year	2005.080	2005.060	0.027	0.961
Subject	10.219	10.219	0.000	1.000
Country	2.054	2.054	0.000	1.000
Total Career Age	30.450	30.727	− 0.277	0.632
Total Career Age Bins	7.738	7.738	0.000	1.000
Cumulative Publication experience	13.515	13.554	− 0.038	0.945
Cumulative Publication experience Bins	3.293	3.300	− 0.007	0.953
Cumulative citations received per cumulative publications	43.732	44.364	− 0.632	0.814
Cumulative citations received per cumulative publications Bins	2.246	2.246	0.000	1.000
(b)				
An unbalanced panel of 1208 matched authors, with 604 in each group				
Year	2007.331	2007.296	0.035	0.920
Subject	10.942	10.942	0.000	1.000
Country	2.013	2.013	0.000	1.000
Total Career Age	20.608	20.755	− 0.147	0.801
Total Career Age Bins	5.750	5.750	0.000	1.000
Cumulative Publication experience	9.353	9.262	0.091	0.808
Cumulative Publication experience Bins	2.414	2.414	0.000	1.000
Cumulative citations received per cumulative publications	43.560	44.257	− 0.697	0.943
Cumulative citations received per cumulative publications Bins	2.007	2.007	0.000	1.000
(c)				
An unbalanced panel of 990 matched authors, with 495 in each group				
Year	2007.911	2007.889	0.022	0.953
Subject	10.820	10.820	0.000	1.000
Country	1.974	1.974	0.000	1.000
Total Career Age	20.198	20.418	− 0.220	0.729
Total Career Age Bins	5.677	5.677	0.000	1.000
Cumulative Publication experience	9.479	9.543	− 0.065	0.877
Cumulative Publication experience Bins	2.471	2.471	0.000	1.000
Cumulative citations received per cumulative publications	34.323	34.262	0.062	0.966
Cumulative citations received per cumulative publications Bins	1.881	1.881	0.000	1.000

Reports the t-test for the mean difference between control and treated groups one year before forming the star-help relation

Table 8 Robustness test of dynamic star help effects for field normalised total citations

	Balanced panel (1)	Alternative matching using the availability of publication match in control at t = 0 (2)	Alternative matching using availability of acknowledgement text match in control at t = 0 (3)
$Staracknw_{i,t\text{-}3}$	− 0.0188	− 0.0115	− 0.00760
	(0.0421)	(0.0258)	(0.0253)
$Staracknw_{i,t\text{-}2}$	− 0.0304	− 0.0279	− 0.0225
	(0.0405)	(0.0243)	(0.0237)
$Staracknw_{i,t}$	0.240***	0.245***	0.249***
	(0.0414)	(0.0247)	(0.0247)
$Staracknw_{i,t+1}$	0.0946**	0.0527**	0.0648**
	(0.0417)	(0.0263)	(0.0274)
$Staracknw_{i,t+2}$	0.0743*	0.0493*	0.0579**
	(0.0437)	(0.0269)	(0.0280)
$Staracknw_{i,t+3}$	0.105**	0.0437	0.0462
	(0.0451)	(0.0288)	(0.0300)
Constant	0.502***	0.405***	0.389***
	(0.0245)	(0.0224)	(0.0245)
R-squared	0.024	0.029	0.034
Pretest against hypothesized trend (Roth, 2022)			
Power	0.50	0.50	0.50
Hypothesized trend	0.04	0.02	0.02
Bayes factor	0.55	0.55	0.55
Likelihood ratio	0.40	0.36	0.34
Observations	11,440	20,023	16,356
Number of authors	520	1,208	990
Author FE	YES	YES	YES
Year FE	YES	YES	YES

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in Eq. 1. Also, the 1996 cohort is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively help on the author's quality-adjusted productivity[6] is present, albeit variable across the types of help. Furthermore, we note that the estimated coefficients are statistically significant in the acknowledgement event year for each type, which can be taken as further support for hypothesis 1a.

5.4 Robustness

5.4.1 Robustness to a balanced panel

Our baseline estimation uses an unbalanced panel of 1258 authors to estimate the effects on the dependent variable. In a balanced panel, we observe units (in this case, authors) every time period, reducing the noise introduced by unit heterogeneity. To analyse whether this variation in the dataset affects our results, we use a panel of 520 authors present in the data throughout our estimation period (1996–2017). The matching statistics of these 520 authors are reported in Table 7a. We present the estimated results of the event study specification based on the balanced panel in Table 8, Column 1. Similar to the star acknowledgement associated with the unbalanced panel data, we again observe a significant contemporaneous effect (27.12%). Though the estimated coefficients in the further years fall (9.46, 7.43, 10.5 log points), they are statistically significant up to three years after the year of acknowledgement. The event study plot in Fig. 6 again shows no evidence of a pre-trend. In addition, Roth's pretest diagnostic analysis reports a likelihood ratio of 0.40, suggesting that the estimated coefficients are likely to follow a parallel trend before the treatment.

5.4.2 Robustness to alternative matching criteria

As a further robustness check, we examine the sensitivity of the results to alternative matching criteria. One concern is that authors are unlikely to publish their research work annually throughout their careers. It takes time to publish a paper following the review process and editing. As an acknowledgement can only occur in a year a paper is published, acknowledgements may be simply partly picking up the fact that a paper was published in a particular year, thus biasing our estimate of the productivity effect of an acknowledge relative to the control authors. To this end, we identify the authors in the control group that have also published at least one paper in the event year ($t = 0$) when the authors in the treated group have a publication that acknowledges a star. Table 7b reports the matching statistics of the new set of control and treated authors. Here we find 1208 matching authors that satisfy the matching criteria. In Table 8, Column 2, we present the estimated results from the event study specification. In considering the additional matching criteria, we still find a similar contemporaneous effect from star acknowledgement on output. An increase of 27.76% in the quality-adjusted output is estimated once the author gets exposure from the star, which is similar to our baseline results.

As discussed previously in Sect. 3.2, an additional concern is that the availability of acknowledgement text from the Scopus database is limited. The credibility of the star help effects can be questioned due to the comparison between a treated author who gets a match from the control group and has no acknowledgement text available in that particular year.

[6] We also check the effect of star help on the author's productivity on raw publication output (Hypothesis 2). We find a similar trend in the results for each helpful interaction with less magnitude- See "Appendix".

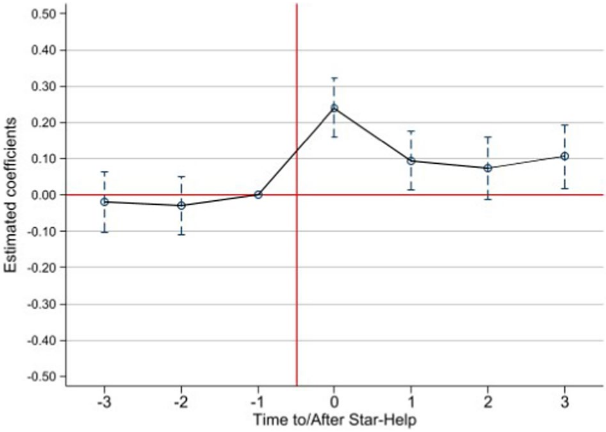

Fig. 6 Robustness test on Event study model with homogenous star help effects at an individual level for a balanced panel. The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

To check the possible implications for our results, we compare the authors in the control group that have a publication with an available acknowledgement text in the event year ($t = 0$) when the treated author acknowledges a star. Table 7c reports the matching statistics of 990 authors at year $t - 1$ with these additional criteria. The event study estimation reported in Table 8, Column 3 again shows similar results to our analysis in Sect. 5.1. We find a contemporaneous increase in the quality-adjusted output of the author by 28.27% associated with star help. Also, the plots for the estimation under both alternative matching criteria are presented in Fig. 7, and they indicate no evidence of a pre-trend. Furthermore, the likelihood ratios of observing any pre-trend are 0.36 and 0.34 from the Roth test analysis, which supports the appropriateness of the parallel trend assumption.

5.4.3 Robustness to country-specific cohorts

The publication data from Scopus comprises those authors that have published in Ireland, Denmark, and New Zealand, and the stars we identified are those individuals who are highly productive relative to their peers in these three countries. Therefore we should be able to see a similar effect of star help on the author's productivity in the year of star acknowledgement. The results reported in Table 9 and presented graphically in Fig. 8 show some differences across countries—with the largest acknowledgement effect observed for Denmark and the smallest for New Zealand—but overall, we find evidence of a significant productivity effect in the year of acknowledgement.

5.4.4 Robustness to co-authorship relations

While our baseline estimation excludes authors that have co-authored with a star prior to the star acknowledgement event from the treatment group and the control group comprises of matched authors that have never co-authored or acknowledged a star scientist at any time in the sample, there could still be the potential for endogeneity from other types of

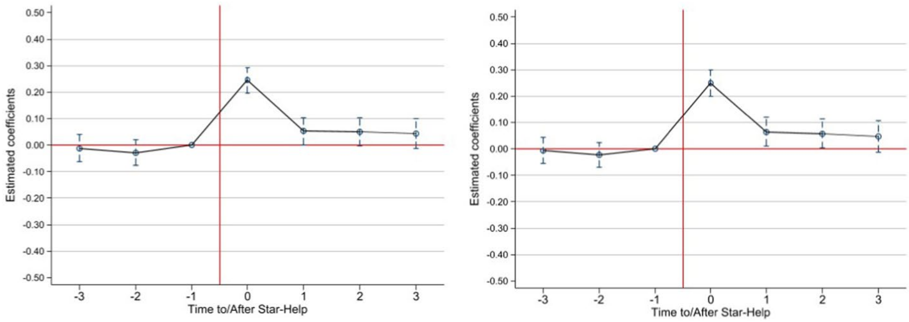

Fig. 7 Robustness test on Event study model with homogenous star help effects at the individual level using alternate matching criteria; the matched control author should have a publication in year $t = 0$ (left), and the control author should have a publication with acknowledgement text in year $t = 0$ (right), where year $t = 0$ is the year when the treatment starts. The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

social connection between the author and the star prior to the acknowledgement event and separate to direct co-authorship. For example, an author might have had an indirect social relationship with the star by co-authoring with a non-star author who may have previously co-authored with the star and thus it would be important to control for social distance in the estimation more generally.

To explore this possibility, we examine the network of co-authors for the treated authors before the acknowledgement of star help (before $t = 0$) to identify the individuals not directly connected to the star through co-authorship but sharing a mutual co-authorship connection with someone else in the second degree of separation. We identify a total of 56 out of the 629 treated authors reported in Table 2 that share a mutual co-authorship connection with the star. Then in a robustness check, we exclude these 56 treated authors and their matched control pairs from the analysis and re-estimate the baseline regression. Column 1 in Table 10 presents the estimated coefficients after their exclusion. Consistent with the baseline results in Table 3, we find that acknowledging star help is associated with a 28.40% (0.25 log points) increase in the quality-adjusted output of the author in the year of acknowledgement. In addition, Fig. 9 shows the estimated event study plot after the exclusion of the 56 treated authors and their matched pairs, and it also indicates that the baseline results are robust to their exclusion.

Another concern for our baseline results is the potential emergence of a co-authorship relationship between the author and the star post the acknowledgement event. This could have important implications for our results where the estimated increase in author output associated with the acknowledgement of star help could instead be partially attributed to the later emergence of a direct co-authorship relationship between author and star. Authors who acknowledge a star for their help could also produce more co-authored publications with the same star or indeed any star in the years after the event ($t = 0$). In a further test of the robustness of our results, we examine the effect of once-off and multiple-star help on author's output, excluding all co-authorship publications that occur between a star and author after the acknowledgement event from our estimation.[7]

[7] In an additional robustness check, we also examine whether our results are robust when we exclude the 179 matched pairs of treated and control authors from the analysis that co-author with a star after the acknowledgement event rather than just excluding the publications, and we find similar supporting evidence for the baseline results—See "Appendix".

Table 9 Robustness test of dynamic star help effects for field normalised total citations for the three countries separately

	Ireland (1)	Denmark (2)	New Zealand (3)
$Staracknw_{i,t\text{-}3}$	− 0.0881	0.0184	− 0.0254
	(0.0620)	(0.0306)	(0.0650)
$Staracknw_{i,t\text{-}2}$	− 0.00941	− 0.0322	− 0.0107
	(0.0562)	(0.0295)	(0.0575)
$Staracknw_{i,t}$	0.213***	0.271***	0.198***
	(0.0636)	(0.0288)	(0.0574)
$Staracknw_{i,t+1}$	− 0.0749	0.0782***	0.0701
	(0.0754)	(0.0287)	(0.0599)
$Staracknw_{i,t+2}$	0.0288	0.0623**	0.0117
	(0.0825)	(0.0293)	(0.0611)
$Staracknw_{i,t+3}$	− 0.0825	0.101***	− 0.0433
	(0.0712)	(0.0339)	(0.0599)
Constant	0.330***	0.418***	0.443***
	(0.0581)	(0.0283)	(0.0545)
R-squared	0.042	0.026	0.044
Pretest against the hypothesized trend (Roth, 2022)			
Power	0.50	0.50	0.50
Hypothesized trend	0.06	0.03	0.06
Bayes factor	0.55	0.55	0.55
Likelihood ratio	2.30	0.05	0.34
Observations	3,851	12,873	4,117
Number of authors	226	784	248
Author FE	YES	YES	YES
Year FE	YES	YES	YES

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in Eq. 1. Also, the 1996 cohort is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

Column 2 in Table 10 reports the results for one-time star help after excluding these publications. Again, the results are similar to the baseline and also have a similar dynamic pattern (see Table 4, Column 2 and Fig. 3 (right) for comparison). Overall, they show a star help effect of 26. 5% for the quality-adjusted productivity of the author in the year $t = 0$ as well as no evidence of a sustained productivity effect in the following years. Furthermore, Column 3 in Table 10 reports the estimated coefficients for multiple star help after excluding the post acknowledgement event star co-authorship publications from the analysis. Here, the star help effect is estimated to be a 39.1% increase in the quality-adjusted output of the author. Additionally, in the years following the initial acknowledgement event, the results show statistically significant coefficients broadly comparable with the baseline findings but with a smaller positive impact at each lag. Figure 10 also shows the event study plots of the star help effect for both once-off and multiple star help after removing these co-authored publications, and these are comparable to the baseline plots in Fig. 3. Therefore,

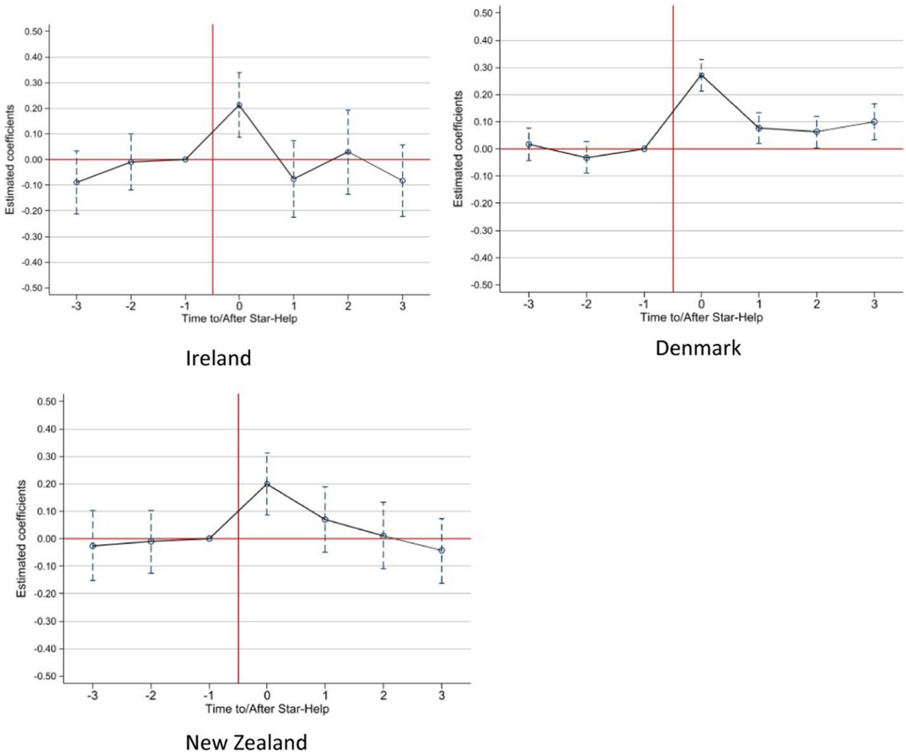

Fig. 8 Event study model with homogenous star help effects at an individual level by country (Ireland, Denmark, New Zealand). The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

our initial results are largely robust to the exclusion of co-authored publications between an author and star in the years after the acknowledgement event.

5.4.5 Robustness to single and multiple acknowledging authors

To test whether the effects of a star acknowledgement are different where the star is acknowledged multiple times, we define a dummy variable M that takes a value of 0 if the star is acknowledged just once and 1 where if the star is acknowledged more than once. This leads to a revised estimating equation:

$$\begin{aligned} lnY_{it} =& \alpha + \beta_{\leq -4} staracknw_{i,-4} + \sum_{j=-3}^{-2} \beta_j staracknw_{ij} + \sum_{j=0}^{3} \beta_j staracknw_{ij} + \beta_{\geq 4} staracknw_{i,4} \\ &+ \gamma_{\leq -4}(staracknw_{i,-4} \times M) + \sum_{j=-3}^{-2} \gamma_j \left(staracknw_{ij} \times M\right) + \sum_{j=0}^{3} \gamma_j (staracknw_{ij} \times M) \\ &+ \gamma_{\geq 4}(staracknw_{i,4} \times M)) + \delta_t + \mu_i + \epsilon_{it}. \end{aligned} \tag{5}$$

Table 10 Robustness test of dynamic Star Help Effects by excluding the authors with previous indirect relations with a star (column1) and excluding the co-authored publications with a star after the initial acknowledgement event (column 2 & 3)

	Indirect relationship with a star before the event (1)	One–time star help effect: excluding the publications (2)	Multiple star help effect: excluding the publications (3)
$Staracknw_{i,t-3}$	− 0.0122	0.00172	− 0.110
	(0.0251)	(0.0269)	(0.0752)
$Staracknw_{i,t-2}$	− 0.0245	− 0.0236	− 0.0217
	(0.0244)	(0.0253)	(0.0743)
$Staracknw_{i,t}$	0.250***	0.235***	0.330***
	(0.0250)	(0.0260)	(0.0624)
$Staracknw_{i,t+1}$	0.0589**	0.00767	0.129**
	(0.0266)	(0.0273)	(0.0634)
$Staracknw_{i,t+2}$	0.0556**	− 0.00126	0.196***
	(0.0274)	(0.0276)	(0.0707)
$Staracknw_{i,t+3}$	0.0463	− 0.00401	0.203**
	(0.0292)	(0.0284)	(0.0827)
Constant	0.402***	0.386***	0.540***
	(0.0238)	(0.0244)	(0.0692)
R-squared	0.029	0.028	0.056
Pretest against the hypothesized trend (Roth, 2022)			
Power	0.50	0.50	0.50
Hypothesized trend	0.02	0.07	0.02
Bayes factor	0.55	0.55	0.55
Likelihood ratio	0.38	2.90	0.14
Observations	19,021	18,160	2,681
Number of authors	1,146	1,108	150
Author FE	YES	YES	YES
Year FE	YES	YES	YES

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in Eq. 1. Also, the 1996 cohort (always treated) is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

At any given lead or lag, a simple test of difference in effect for the cases of single or multiple acknowledgements is then a test of the statistical significance of the relevant γ coefficient. A positive and statistically significant coefficient in period 0 (the period in which the acknowledgement occurs) would indicate that even in the period in which the acknowledgement occurs, a scientist that makes multiple acknowledgements receives a greater productivity boost than a scientist who only makes a single acknowledgement. This could reflect unobserved heterogeneity between single and multiple acknowledging scientists (that is not picked up by our controls) or indicate that the quality of initial help received is greater for multiple acknowledging scientists. However, a finding of no statistically different *initial* productivity effects of acknowledgements between the two groups

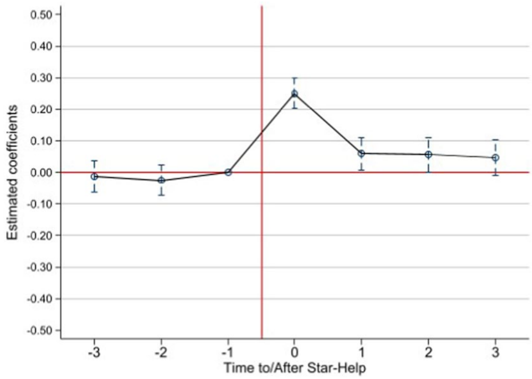

Fig. 9 Event study model with homogenous star help effects at an individual level: Excluding the authors with indirect star relationship before the acknowledgement event. The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

helps allay concerns of unobserved heterogeneity between the single and multiple citing scientists.

We report the event study estimates of the model in Table 11. In Column 1, we show the estimated beta coefficients from Eq. (5), while in Column 2, we show the estimated gamma coefficients on the interaction between the indicator variable for star help and the multiple acknowledgement dummy. Most importantly, in the year of star help ($j = 0$), we don't find any evidence of a significantly different initial productivity effect for the authors who acknowledge a star multiple times. We take this as evidence that the observation of multiple acknowledgements does not imply that the scientist is more productive and/or can make better use of star help, suggesting that the sustained productivity effect associated with multiple acknowledgements reflects a causal effect of the help rather than being due to any selection effect.

5.4.6 Robustness to the productivity of acknowledging authors

In a final robustness test, we examine for potential heterogeneous treatment effects across the distribution of the outcome variable. The effect on productivity from acknowledging star help could differ depending on the productivity of the acknowledging authors. To investigate this, we look at the cumulative FNTC of the treated authors one year before treatment, the acknowledgment to a star. This productivity indicator provides a measure of where an author is in the initial productivity distribution before treatment. Table 12 reports summary statistics for each of the four quartiles of treated authors from the distribution of cumulative FNTC at year $t-1$. We divide our overall sample of treated authors into four sub-samples based on these quartiles and then, using Eq. 1, we estimate the results for each sub-sample of treated authors and their matched pairs from the control group.

The event study estimates for each quartile are reported in Table 13 and their related event-study plots are presented in Fig. 11. The results show that the contemporaneous effect of star help on author productivity for authors in the first and second quartiles (column 1 and 2) is larger compared to the effect for authors in the third and fourth quartiles (columns 3 and 4) of the cumulative FNTC distribution. Star help is associated with a statistically significant increase of 38.13% for authors in quartile 1 and a statistically insignificant increase of 13.54% for authors in quartile 4. The event study plots also indicate that the effect is sustained in the years after the acknowledgement event

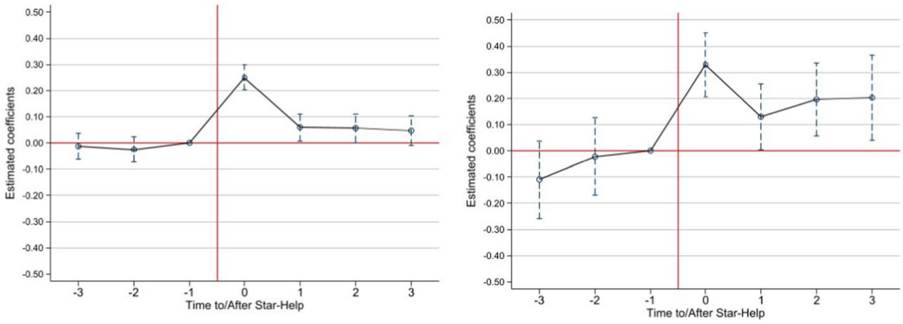

Fig. 10 Event study model with homogenous star help effects at an individual level: Excluding the future star co-authored publications for authors who acknowledge star help for: One–Time (Left); Multiple Times in the years after the first point of contact (Right). The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

Table 11 Robustness test of dynamic Star Help Effects for Field Normalised Total Citations comparing the initial productivity effects of authors who acknowledge star single and multiple times

	Overall Star Help Effects(1)		Star Help Effects;Interacted with Dummy Variable for Multiple Star Help (2)
$Staracknw_{i,t\text{-}3}$	0.000487	$M.Staracknw_{i,t\text{-}3}$	− 0.0962
	(0.0269)		(0.0812)
$Staracknw_{i,t\text{-}2}$	− 0.0243	$M.Staracknw_{i,t\text{-}2}$	0.00500
	(0.0253)		(0.0795)
$Staracknw_{i,t}$	0.237***	$M.Staracknw_{i,t}$	0.0874
	(0.0260)		(0.0688)
$Staracknw_{i,t+1}$	0.0312	$M.Staracknw_{i,t+1}$	0.132*
	(0.0274)		(0.0718)
$Staracknw_{i,t+2}$	0.0196	$M.Staracknw_{i,t+2}$	0.190**
	(0.0277)		(0.0772)
$Staracknw_{i,t+3}$	0.0129	$M.Staracknw_{i,t+3}$	0.172**
	(0.0288)		(0.0860)
Constant	0.409***		
	(0.0238)		
R-squared	0.030		
Observations	20,841		
Number of authors	1,258		
Author FE	YES		
Year FE	YES		

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in Eq. 5. Also, the 1996 cohort (always treated) is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

Table 12 Summary statistics: by quartile of the cumulative FNTC distribution for the treated group at year t-1

Quartile	N	Mean	SD	Median	Min	Max	IQR
1	2943	.544	0.312	.589	0	1.017	.58
2	2938	1.452	0.224	1.474	1.021	1.828	.364
3	2926	2.252	0.255	2.23	1.833	2.738	.445
4	2933	3.556	0.816	3.298	2.744	6.705	.959

Table 13 Robustness test of dynamic Star Help Effects for Field Normalised Total Citations comparing the low and high productivity clusters of treated authors based on the quartiles of the cumulative FNTC at t-1

	Quartile-1 (1)	Quartile-2 (2)	Quartile-3 (3)	Quartile-4 (4)
$Staracknw_{i,t-3}$	0.0246	− 0.0641	− 0.119**	0.0981
	(0.0224)	(0.0415)	(0.0515)	(0.0891)
$Staracknw_{i,t-2}$	− 0.0107	− 0.0224	− 0.0893*	0.0125
	(0.0189)	(0.0408)	(0.0510)	(0.0888)
$Staracknw_{i,t}$	0.323***	0.297***	0.184***	0.127
	(0.0270)	(0.0463)	(0.0541)	(0.0774)
$Staracknw_{i,t+1}$	0.0770***	0.0837*	0.0301	− 0.0212
	(0.0252)	(0.0431)	(0.0588)	(0.0901)
$Staracknw_{i,t+2}$	0.159***	0.0482	− 0.0246	− 0.0691
	(0.0333)	(0.0455)	(0.0574)	(0.0852)
$Staracknw_{i,t+3}$	0.110***	0.0395	− 0.0317	− 0.0144
	(0.0304)	(0.0500)	(0.0651)	(0.0922)
Constant	0.136***	0.280***	0.520***	0.772***
	(0.0252)	(0.0300)	(0.0424)	(0.0622)
R-squared	0.059	0.056	0.038	0.022
Observations	5,507	5,348	5,069	4,917
Number of authors	420	328	270	240
Author FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in Eq. 1. Also, the 1996 cohort (always treated) is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

for quartile 1 authors on average. In contrast, it is not sustained for the authors in later quartiles. In comparing the low- and high-productivity author clusters, we thus find that star help has a greater effect on authors that have low productivity before the star help is received relative to the authors that have high initial productivity. We interpret this result as indicating that lower productivity authors gain most from star help. We hope to further investigate this finding in future work, as it may have important implications for policies that can be used to support lower-productivity colleagues.

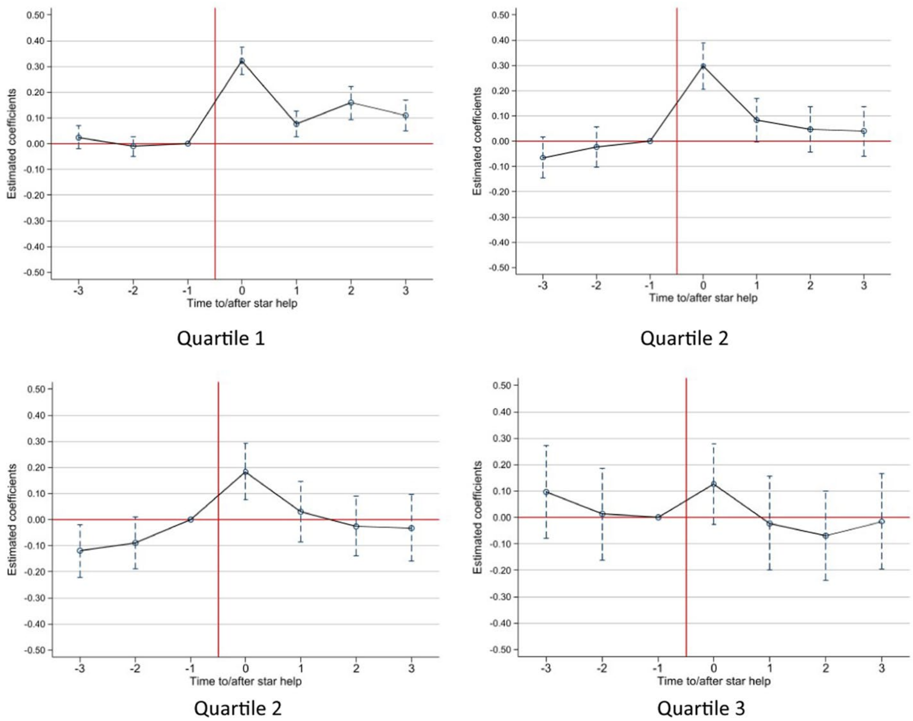

Fig. 11 Event study model with homogenous star help effects at an individual level comparing low–high productivity clusters by equal four quartiles of cumulative FNTC received for the treated authors one year before the treatment. The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

6 Summary and policy implications

Although there is a growing body of literature that estimates the effects of connections to stars on the productivity of peers, a star's role in helping their peers through non-co-author-based relationships has been largely neglected. While focusing on the performance gains of co-authoring with a helpful star, Oettl (2012) discusses the possibility of extending the research to colleagues and students who are not co-authoring with the star. Following this suggestion, this paper investigates the effects of star help on the productivity of scientists receiving the help.

To implement this empirically, we identify interactions between a star and an author identified from the acknowledgement texts in a publication. We treat these acknowledgements as an indicator of helpful interaction and examine their impacts on the author's quality-adjusted scientific output.

We analyse the output of authors who publish in three countries: Ireland, Denmark, and New Zealand. Using Natural Language Processing (NLP) techniques, we extract the names of individuals who get acknowledged for their contributions to the research. We match these individuals to previously identified stars' names relative to the peers who published in the three countries. We estimate the effect of star help by comparing scientists who acknowledge the star with a carefully matched control group of scientists who do not.

The overall effect of star help is estimated in an event study setting, where the event is the acknowledgement by an author of an identified star. We observe an increase in the quality-adjusted output of authors in the event year. Although we find evidence of a sharp fall in the productivity effect in subsequent years, the initial effect of star interaction is significant. In addition, analysis of authors who continue to acknowledge a star in the years after the initial acknowledgement shows a higher and sustained helpfulness effect on their output.

In addition, we examine the effects of different types of star help. We classify these interactions from a star based on the keywords in the acknowledgement text. We identify five interaction channels through which the knowledge transfer occurs. A disaggregated analysis of these interactions suggests that all types of help positively affect the author's output. However, authors who acknowledge a star for the conceptual help show the highest increase in the output in the year of acknowledgement. Unlike other forms of star help, authors who acknowledge a star for materials received show consistent productivity effects even without evidence of sustained acknowledgements.

A second contribution of our analysis is to examine star-help effects within an explicitly dynamic framework. Our staggered event-study findings show significant productivity effects in the year that star help is acknowledged. However, the observed effects drop sharply in the years following the acknowledgement unless that help is sustained. Moreover, these findings are robust to recent techniques that explicitly account for heterogeneity across the years in which the star help is observed to occur.

Previous work has established that the occurrence and nature of star help can be an important mediating factor in determining the effects of co-authorship relationships with a star (e.g., Oettl, 2012). Our study focuses on the productivity effects of star help in the absence of a co-author relationship. We find a robust relationship across forms of star help between the acknowledgement of help and the productivity of non-co-authoring researchers. Moreover, the effects on productivity are more significant where there is evidence of the help being sustained over time and where the researcher receiving the help is positioned in the lower quartiles of the relevant field-specific productivity distribution.

These results have important implications for the recruitment and organisational strategies of academic departments. To the extent that stars are co-located with non-star researchers, they should have more opportunities to provide help to the benefit of incumbents at the receiving department, with the beneficial effects being very impactful for less productive researchers. This points to one source of potential value from star recruitment policies in addition to the more widely studied co-authorship and recruitment-quality channels, at least to the extent that co-located stars are better positioned to provide help to their non-star colleagues (Agrawal et al., 2017). Furthermore, as the results are not specific to co-located researchers, organisational policies that help embed researchers in networks that increase the probability of interaction with stars should have productivity benefits. Where the star and non-star are co-located, formal mentoring programmes and more informal means such as regular departmental seminars, workshops and social gatherings should help initiate and develop the relationships that support the provision of star help. Star help can also be provided even where the star and non-star are located in different institutions. In this regard, it is likely beneficial for the non-stars to embed in broader networks that allow for interactions with stars, thus supporting the development of the relationships that facilitate productivity-enhancing star help. This suggests the importance of organisational support for network building, such as providing funding for conference travel and encouraging membership of cross-department societies or international funding consortia. Our results suggest that such policies could be especially beneficial for scientists in the lower parts of their field-specific productivity distributions, especially where this help can be sustained over time.

Appendix

See Table 14.

Table 14 Yearly distribution of publications with acknowledgement texts and percentage of publications in which star names are acknowledged

Year	Percentage of publications with acknowledgement texts	Percentage of stars acknowledged publications	Total
1990	5.58	1.57	10,284
1991	6.22	2.58	11,205
1992	5.75	1.15	12,051
1993	5.72	2.20	13,480
1994	5.43	1.47	15,044
1995	6.47	2.34	15,838
1996	6.29	1.23	18,053
1997	8.24	1.85	19,640
1998	9.31	1.74	20,353
1999	9.06	2.10	21,076
2000	8.75	1.02	22,512
2001	8.63	1.60	23,893
2002	8.75	1.92	25,018
2003	9.37	1.03	28,116
2004	10.29	1.23	30,702
2005	9.50	1.27	34,805
2006	9.92	1.01	36,981
2007	9.84	0.95	39,552
2008	10.28	1.09	40,251
2009	10.32	1.27	42,851
2010	10.69	0.90	45,589
2011	12.41	0.71	47,952
2012	13.42	0.71	49,451
2013	13.22	0.50	51,343
2014	13.47	0.56	52,660
2015	11.87	0.59	53,070
2016	18.81	0.65	54,015
2017	24.09	0.60	53,694
Total		971	889,479

Procedure for identifying Star names and the type of Star help from the acknowledgement texts.

Step 1: Data Collection First, we collect the acknowledgement texts from the Scopus database for all the publications in Ireland, Denmark and New Zealand from 1990 to 2017. The unique identifier for these acknowledgements is the publication ID (EID in Scopus).

Step 2: Data Organisation Once we have collected the data, we organise it in a structured format. This involves creating a dataset containing the publication ID, the acknowledgement text related to that publication, and the authors of these publications.

Step 3: Name Identification In this step, we employ Natural Language Processing (NLP) techniques using the Spacy module available in the Python programming language.

To start, the acknowledgement text is broken down into individual units called tokens (Tokenisation) using Spacy; these are typically words or phrases.
Then, using Named Entity Recognition (NER), we identify, classify and store the named entities such as PERSON, ORGANISATION, PLACES, etc.
Finally, we filter the entities tagged PERSON. Then, using the publication ID as the reference to these names, we check for any wrong identification of entities due to structural deformities in the collected acknowledgement text from the Scopus (such as commas, spaces, hyphens, and non-English names).

Step 4: Star Name Matching After extracting person names, we use fuzzy similarity to identify the exact names of scientists. *Fuzzy similarity* is a technique used in computing and is based on fuzzy logic. It involves finding strings that are approximately equal to a given pattern. We match the names identified from the acknowledgement texts with those of the scientists we have previously identified as stars.

Step 5: Verification This step involves manually verifying that the star names identified are the same as those in the acknowledgements. This is important to ensure the accuracy of the results.

Step 6: Identifying help keywords In this step, we manually check the help keywords associated with verified star names and store them. These keywords are help words or indications of star-help to the author. Furthermore, we use these help words to identify the type of helpful interaction from a star(Conceptual, Technical, Material, Funds & support and Other types), as outlined in Sect. 5.3. Table 15 shows the help words used to identify each type of acknowledgement types that defines the type of helpful interaction.

Step 7: Final Data Compilation Finally, we compile the data, which includes the star name, publication ID, author details and help keywords associated with the star scientists. This final dataset is then merged with the panel data of all authors, and dummy variables are used to tag those authors and publications that acknowledge star help (Table 16, 17, 18; Figs.12, 13).

Table 15 Identification of Help words used to describe the type of helpful interaction from a star scientists

Classification	Definition	Help words
Conceptual	Primarily thank star scientists for intellectual feedback, critique, and encouragement	Discussion, Suggestions, Reviews, Comments, Critiques
Technical	Acknowledging star scientists for their help in the technical side of the study	Excellent technical assistance, Laboratory assistance, Lab assistance, Mathematical help, Statistical advice
Materials	Mentioning the star scientists for the help with access to the materials which is required to carry out meaningful research	Supply and use of antibodies, protein cells, antibodies, Access to unpublished data
Funds & Support	To acknowledge the star scientists for the funds and support they receive throughout the time of their research	Funds, Grants, Financial help, Support
Other Types	The category to cover all the types of acknowledgments where we cannot infer or identify the proper type of helpful interaction from the star	Contributions, Dedications, Founder of the study

Table 16 Summary statistics: Control and Treated group (k-to-k matched) for an unbalanced panel for the types of channels

Variable	Control	Treated	Diff in mean	P-value
Conceptual: Unbalanced panel of 468 matched authors, with 234 in each group				
Year	2006.282	2006.278	0.004	0.994
Subject	9.145	9.145	0.000	1.000
Country	2.068	2.068	0.000	1.000
Total Career Age	20.389	20.389	0.043	0.963
Total Career Age Bins	5.697	5.697	0.000	1.000
Cumulative Publication experience	8.175	8.124	0.051	0.932
Cumulative Publication experience Bins	2.171	2.171	0.000	1.000
Cumulative citations received per cumulative publications	56.448	57.712	− 1.264	0.960
Cumulative citations received per cumulative publications Bins	2.154	2.154	0.000	1.000
Technical: Unbalanced panel of 300 matched authors, with 150 in each group				
Year	2006.907	2006.880	0.027	0.967
Subject	10.880	10.880	0.000	1.000
Country	1.987	1.987	0.000	1.000
Total Career Age	20.193	20.547	− 0.353	0.763
Total Career Age Bins	5.687	5.687	0.000	1.000
Cumulative Publication experience	9.193	9.113	0.080	0.912
Cumulative Publication experience Bins	2.420	2.420	0.000	1.000
Cumulative citations received per cumulative publications	32.358	32.660	− 0.302	0.902
Cumulative citations received per cumulative publications Bins	1.813	1.813	0.000	1.000
Materials: Unbalanced panel of 176 matched authors, with 88 in each group				
Year	2006.818	2006.795	0.023	0.980
Subject	11.148	11.148	0.000	1.000
Country	1.955	1.955	0.000	1.000
Total Career Age	21.989	22.068	− 0.080	0.961
Total Career Age Bins	6.023	6.023	0.000	1.000
Cumulative Publication experience	9.966	9.727	0.239	0.808
Cumulative Publication experience Bins	2.523	2.523	0.000	1.000

Table 16 (continued)

Variable	Control	Treated	Diff in mean	P-value
Cumulative citations received per				
cumulative publications	43.074	44.368	− 1.294	0.809
Cumulative citations received per				
cumulative publications Bins	2.250	2.250	0.000	1.000
Funds & Support: Unbalanced panel of 60 matched authors, with 30 in each group				
Year	2010.300	2010.233	0.067	0.964
Subject	11.967	11.967	0.000	1.000
Country	2.000	2.000	0.000	1.000
Total Career Age	18.633	19.267	− 0.633	0.798
Total Career Age Bins	5.467	5.467	0.000	1.000
Cumulative Publication experience	10.367	10.367	− 0.167	0.928
Cumulative Publication experience Bins	2.633	2.633	0.000	1.000
Cumulative citations received per cumulative publications	40.707	42.375	− 1.668	0.864
Cumulative citations received per cumulative publications Bins	2.100	2.100	0.000	1.000
Other Types: Unbalanced panel of 248 matched authors, with 124 in each group				
Year	2008.234	2008.226	0.008	0.991
Subject	13.694	13.694	0.000	1.000
Country	2.008	2.008	0.000	1.000
Total Career Age	20.548	20.871	− 0.323	0.811
Total Career Age Bins	7.738	7.738	0.000	1.000
Cumulative Publication experience	9.790	9.895	− 0.105	0.903
Cumulative Publication experience Bins	2.500	2.500	0.000	1.000
Cumulative citations received per cumulative publications	37.064	36.567	0.497	0.895
Cumulative citations received per cumulative publications Bins	1.935	1.935	0.000	1.000

Reports the t-test for the mean difference between control and treated groups one year before forming the star- relation

Table 17 Dynamic Star Help Effects for Raw publications output for five types of interactions. Estimation Window: 1996–2017

	Conceptual (1)	Technical (2)	Materials (3)	Funds & Support (4)	Other Types (5)
$Staracknwtype_{i,t\text{-}3}$	− 0.00830	− 0.00461	0.0634	− 0.128	− 0.0412
	(0.0297)	(0.0388)	(0.0521)	(0.0946)	(0.0465)
$Staracknwtype_{i,t\text{-}2}$	− 0.0408	− 0.0000	0.0305	− 0.0843	− 0.0543
	(0.0295)	(0.0386)	(0.0504)	(0.110)	(0.0344)
$Staracknwtype_{i,t}$	0.219***	0.188***	0.220***	0.193**	0.173***
	(0.0264)	(0.0301)	(0.0437)	(0.0829)	(0.0352)
$Staracknwtype_{i,t+1}$	0.0366	0.0290	0.0818	− 0.00607	− 0.0110
	(0.0295)	(0.0360)	(0.0539)	(0.113)	(0.0484)
$Staracknwtype_{i,t+2}$	0.0103	− 0.0153	0.120**	− 0.00258	− 0.00547
	(0.0325)	(0.0384)	(0.0570)	(0.0965)	(0.0531)
$Staracknwtype_{i,t+3}$	0.0127	0.0165	0.133**	− 0.0621	− 0.0534
	(0.0328)	(0.0432)	(0.0602)	(0.103)	(0.0496)
Constant	0.404***	0.473***	0.378***	0.428***	0.420***
	(0.0311)	(0.0333)	(0.0472)	(0.0738)	(0.0459)
R-squared	0.051	0.051	0.071	0.095	0.057
Observations	7,665	4,942	3,001	947	4,152
Number of authors	468	300	176	60	248
Author FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

The dependent variable is raw publications

The table reports the estimates based on the model specification in Eq. 4. Also, the 1996 cohort (always treated) is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

Table 18 Dynamic Star Help Effects for Field Normalised Total Citations excluding the authors with co-authored publications with a star afterinitial contact

	Overall star help effect (1)	One-time star help effect (2)	Multiple star help effect (3)
$Staracknw_{i,t-3}$	0.0267	0.0303	− 0.0550
	(0.0303)	(0.0313)	(0.105)
$Staracknw_{i,t-2}$	− 0.0144	− 0.0128	− 0.0383
	(0.0288)	(0.0297)	(0.110)
$Staracknw_{i,t}$	0.245***	0.233***	0.384***
	(0.0285)	(0.0295)	(0.101)
$Staracknw_{i,t+1}$	0.00949	− 0.00568	0.170
	(0.0306)	(0.0321)	(0.103)
$Staracknw_{i,t+2}$	0.0257	− 0.00409	0.279**
	(0.0328)	(0.0335)	(0.115)
$Staracknw_{i,t+3}$	− 0.00123	− 0.0327	0.263**
	(0.0337)	(0.0339)	(0.124)
Constant	0.424***	0.400***	0.643***
	(0.0292)	(0.0297)	(0.105)
R-squared	0.022	0.024	0.061
Observations	14,408	12,946	1,462
Number of authors	900	816	84
Author FE	YES	YES	YES
Year FE	YES	YES	YES

Estimation Window: 1996–2017. The dependent variable is Field-Normalised Total Citations

The table reports the estimates based on the model specification in Eq. 1. Also, the 1996 cohort (always treated) is not dropped in this case. Robust standard errors are clustered at an individual level and reported in parentheses. *, **, and *** represent significance levels at the 10%, 5% and 1% respectively

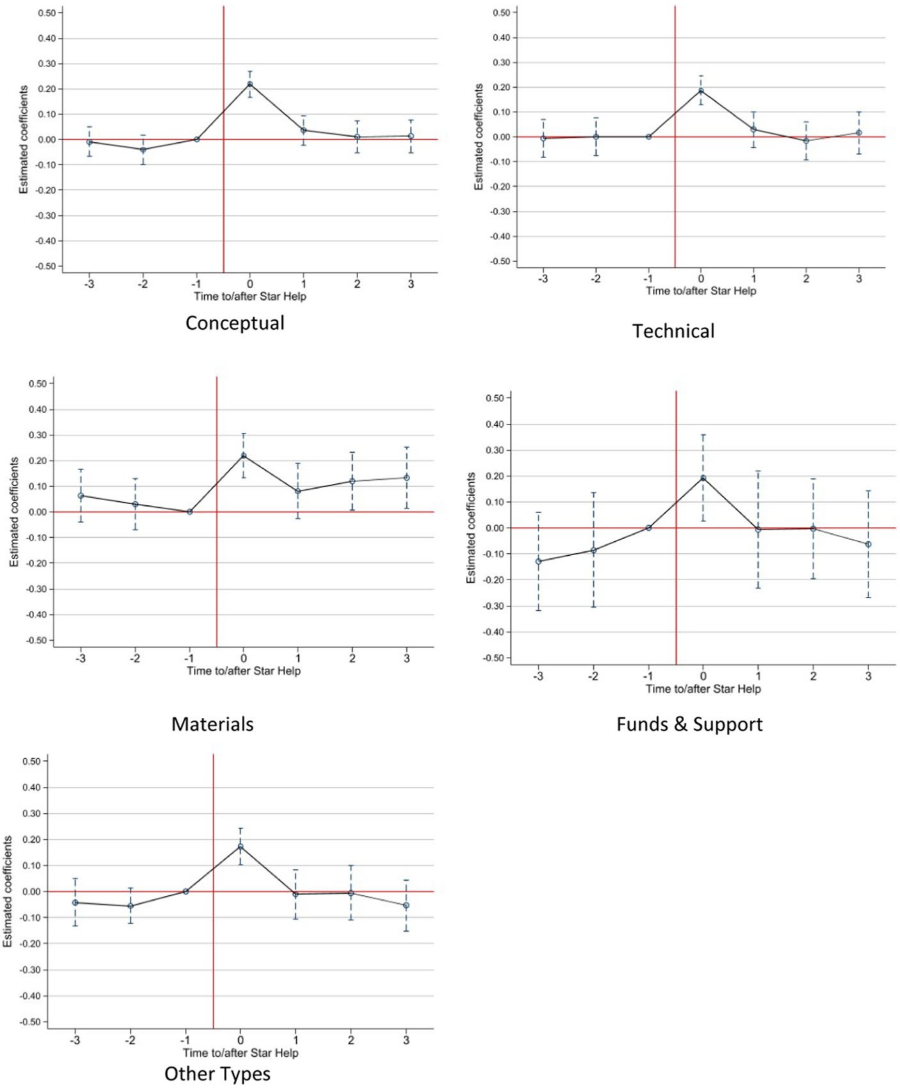

Fig. 12 Event study model with types of channels identified from the acknowledgement texts. Conceptual (row 1, left), Technical (row 1, right), Material (row 2, left), Funds & Support (row 2, right), and Other types (row 3). The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is raw publications

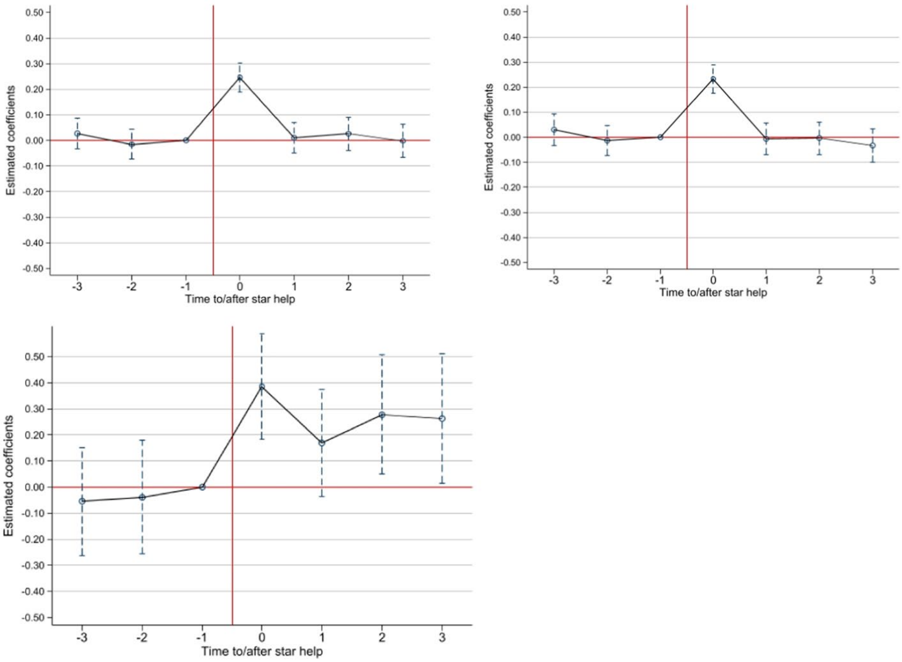

Fig. 13 Event study model with homogenous star help effects at an individual level:Excluding the authors with future star co-authored publications that happen after the star help (t = 0): Overall Star help Effect (Row1, Left), One-Time Star Help Effect (Row1, Right), Multiple Star Help Effect (Row 2, Left). The figure plots the dynamic effect of star interactions with an author at 95% confidence intervals. The event and observation window is from 1996–2017. The dependent variable is field normalised total citations

Acknowledgements We are indebted to Jen-Chung Mei, Anil Yadav and Galetti Jefferson for the valuable discussions and helpful comments received throughout the research. We also thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

Funding Open Access funding provided by the IReL Consortium. We gratefully acknowledge funding from Science Foundation Ireland under the SFI Science Policy Research Programme, Grant Number 17/SPR/5329.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Agrawal, A., McHale, J., & Oettl, A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy, 46*(4), 853–867.

Azoulay, P., & Zivin, J. G. (2005). *Peer effects in the workplace: evidence from professional transitions for the superstars of medicine* (No. y: 2006: x: 1).

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar extinction. *The Quarterly Journal of Economics, 125*(2), 549–589.

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography, 28*(1), 31–56.

Berg, N., & Faria, J. (2008). Negatively correlated author seniority and the number of acknowledged people: Name-recognition as a signal of scientific merit? *The Journal of Socio-Economics, 37*(3), 1234–1247.

Bergman, E. M., & Schubert, U. (2005). Spillovers and innovation, environment and space: policy uncertainties and research opportunities. In *Spillovers and Innovations* (pp. 157–177). Springer, Vienna.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). cem: Coarsened exact matching in Stata. *The Stata Journal, 9*(4), 524–546.

Blumenthal, D., Causino, N., Campbell, E., & Louis, K. S. (1996). Relationships between academic institutions and industry in the life sciences—an industry survey. *New England Journal of Medicine, 334*(6), 368–374.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies, 39*(1), 61–74.

Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy, 33*(4), 599–616.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2020). Peer effects in networks: A survey. *Annual Review of Economics, 12*(1), 603–629.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200–230.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128–152.

Cronin, B. (1991). Let the credits roll: a preliminary examination of the role played by mentors and trusted assessors in disciplinary formation. *Journal of Documentation*.

Cronin, B., McKenzie, G., & Rubio, L. (1993). The norms of acknowledgement in four humanities and social sciences disciplines. *Journal of Documentation*.

Desrochers, N., Paul-Hus, A., Haustein, S., Costas, R., Mongeon, P., Quan-Haase, A., & Larivière, V. (2018). Authorship, citations, acknowledgments and visibility in social media: Symbolic capital in the multifaceted reward system of science. *Social Science Information, 57*(2), 223–248.

Fini, R., Jourdan, J., & Perkmann, M. (2018). Social valuation across multiple audiences: The interplay of ability and identity judgments. *Academy of Management Journal, 61*(6), 2230–2264.

Fini, R., Perkmann, M., Kenney, M., & Maki, K. M. (2023). Are public subsidies effective for university spinoffs? Evidence from SBIR awards in the University of California system. *Research Policy, 52*(1), 104662.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there). *Journal of Economic Geography, 3*(1), 75–99.

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

Grigoriou, K., & Rothaermel, F. T. (2014). Structural microfoundations of innovation: The role of relational stars. *Journal of Management, 40*(2), 586–615.

Groysberg, B., Lee, L. E., & Nanda, A. (2008). Can they take it with them? The portability of star knowledge workers' performance. *Management Science, 54*(7), 1213–1230.

Haeussler, C., Jiang, L., Thursby, J., & Thursby, M. (2014). Specific and general information sharing among competing academic researchers. *Research Policy, 43*(3), 465–475.

Hellqvist, B. (2010). Referencing in the humanities and its implications for citation analysis. *Journal of the American Society for Information Science and Technology, 61*(2), 310–318.

Horta, H., Veloso, F. M., & Grediaga, R. (2010). Navel gazing: Academic inbreeding and scientific productivity. *Management Science, 56*(3), 414–429.

Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association, 106*(493), 345–361.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*(1), 1–24.

Kehoe, R. R., & Tzabbar, D. (2015). Lighting the way or stealing the shine? An examination of the duality in star scientists' effects on firm innovative performance. *Strategic Management Journal, 36*(5), 709–727.

King, G., Nielsen, R., Coberley, C., Pope, J. E., & Wells, A. (2011). Comparative effectiveness of matching methods for causal inference. *Unpublished manuscript, Institute for Quantitative Social Science, Harvard University, Cambridge, MA*.

King, G., & Nielsen, R. (2019). Why propensity scores should not be used for matching. *Political Analysis, 27*(4), 435–454.

Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation, 11*(1), 3–15.

Li, W., Aste, T., Caccioli, F., & Livan, G. (2019). Early co-authorship with top scientists predicts success in academic careers. *Nature Communications, 10*(1), 1–9.

Mackintosh, K. H. (1972). *Acknowledgement patterns in sociology*. University of Oregon.

Maier, G., Kurka, B., & Trippl, M. (2007). Knowledge spillover agents and regional development: spatial distribution and mobility of star scientists. *Dynamic Regions in a Knowledge-Driven Global Economy, 17*, 35.

McCain, K. W. (1991). Communication, competition, and secrecy: The production and dissemination of research-related information in genetics. *Science, Technology, & Human Values, 16*(4), 491–516.

McHale, J., Harold, J., Mei, J. C., Sasidharan, A., & Yadav, A. (2022). Stars as catalysts: an event-study analysis of the impact of star-scientist recruitment on local research performance in a small open economy. *Journal of Economic Geography, 23*(2), 343–369.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.

Mohnen, M. (2022). Stars and brokers: Knowledge spillovers among medical scientists. *Management Science, 68*(4), 2513–2532.

Murray, F. (2010). The oncomouse that roared: Hybrid exchange strategies as a source of distinction at the boundary of overlapping institutions. *American Journal of Sociology, 116*(2), 341–388.

Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. The Belknap Press of Harvard University Press.

Newman, M. (2018). *Networks* (2nd ed.). Oxford University Press.

Nooteboom, B. (2000). Learning by interaction: Absorptive capacity, cognitive distance and governance. *Journal of Management and Governance, 4*(1), 69–92.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science, 58*(6), 1122–1140.

Paruchuri, S. (2010). Intraorganizational networks, interorganizational networks, and the impact of central inventors: A longitudinal study of pharmaceutical firms. *Organization Science, 21*(1), 63–80.

Paul-Hus, A., & Desrochers, N. (2019). Acknowledgements are not just thank you notes: A qualitative analysis of acknowledgements content in scientific articles and reviews published in 2015. *PLoS ONE, 14*(12), e0226727.

Paul-Hus, A., Desrochers, N., & Costas, R. (2016). Characterization, description, and considerations for the use of funding acknowledgement data in Web of Science. *Scientometrics, 108*(1), 167–182.

Paul-Hus, A., Díaz-Faes, A. A., Sainte-Marie, M., Desrochers, N., Costas, R., & Larivière, V. (2017). Beyond funding: Acknowledgement patterns in biomedical, natural and social sciences. *PLoS ONE, 12*(10), e0185578.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy, 94*(5), 1002–1037.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy, 98*(5 Part 2), S71–S102.

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

Schiller, D., & Diez, J. R. (2010). Local embeddedness of knowledge spillover agents: Empirical evidence from German star scientists. *Papers in Regional Science, 89*(2), 275–294.

Schmidheiny, K., & Siegloch, S. (2019). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Equivalence, and Generalization, 38*, 695–713.

Storper, M., & Venables, A. J. (2004). Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography, 4*(4), 351–370.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175–199.

Torre, A., & Rallet, A. (2005). Proximity and localization. *Regional Studies, 39*(1), 47–59.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *The Review of Economic Studies, 79*(2), 838–861.

Walsh, J., & Cohen, W. (2008). Real Impediments to Biomedical Research. *Innovation Policy and the Economy, 8*, 1–30.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics, 113*(2), 331–360.
Yadav, A., McHale, J., & O'Neill, S. (2023). How does co-authoring with a star affect scientists' productivity? Evidence from small open economies. *Research Policy, 52*(1), 104660.
Zucker, L. G., & Darby, M. R. (2006). Movement of star scientists and engineers and high-tech firm entry.