The University as a Local Idea Space

Benefits from research links between colleagues

Lukas Kuld*

February 24, 2021

Recent research observes declining or no returns from the quality of local colleagues on a researcher's productivity but positive spill-overs between co-authors. This paper tests benefits from research links between department colleagues below the level of co-authorship. Based on data from the CVs of around 1,000 highly-cited economists, the success of a research article is linked to the personal and thematic connections of its authors. For these authors, non-routine articles receive around 30 percent more citations, if recent university colleagues are thematically connected. The number of (star) colleagues without a research link has no direct effect on individual papers, nor have future colleagues after the publication date. The timing of the impact, the type of affected research, and zero estimates for other personal links are used to show the causality of benefits from local, thematically connected colleagues.

JEL-codes: A14 • D83 • I23 • J24 • O31 • O32 • O33
Keywords: Academic research, scientific productivity, cluster, location

*Department of Economics, TU Dortmund, Vogelpothsweg 87 D-44227 Dortmund, Germany; kuldl@udo.edu
I thank John O'Hagan for extensive comments and for funding, with the Department of Economics, Trinity College Dublin, of interns Kate Hayes, Greg Mangan, Andrew Morrow, Sarah Mortell, Conor Parle and Michael Stone. I also thank Tara Bedi, Christian Fons-Rosen, Christiane Hellmanzik, Martina Kirchberger, Rogelio Mercado, Sara Mitchell, Carol Newman, David Lagakos, Clemens Struck, Alan Walsh, and the participants of the Trinity College Dublin PhD working group as well as conference participants in Bilbao, Vienna, Valladolid, Dortmund, and Le Mans for help and comments. The research is also generously supported by a Government of Ireland Postgraduate Scholarship from the Irish Research Council.

1 Introduction

Is there an advantage of local clusters of thematically linked colleagues and infrastructure for academic research? Such local idea spaces matter if a researcher's interaction with thematically linked colleagues, or shared infrastructure and institutions, benefits the quality, relevance, or dissemination process of published research. For instance, a local thematic cluster may help involved researchers to identify research gaps and understand a research question's link with concurrent research and its relevance to future research. It can provide access to novel research knowledge to answer the question at hand, including methods, data, and qualitative arguments. Subsequently, seminars and personal networks can increase the visibility of research output, before and after publication.

Most recent literature, however, finds no localized effects, such as peer effects, within university departments (see *inter alia* Kim et al. [2009], Waldinger [2012], Borjas and Doran [2015]). Direct collaborators, on the other hand, are shown to increase a researcher's productivity (see *inter alia* Azoulay et al. [2010], Waldinger [2010]). Our paper connects these results by focusing on thematic links in the research of university colleagues. Publications are shown to be more successful if their authors reference colleagues, who are at their university during the publication's research phase, that is the years before publication. University colleagues, who are referenced in an article, work on topics that are related to said article. At the same time, they are not directly involved as co-authors. These findings suggest that researchers at the same university affect each other positively, if their research is sufficiently close.

In order to evaluate the impact of the local environment empirically, we trace how articles fare after publication. In particular, the advantages of a local cluster may lead to, first, a higher quality at publication (for instance the innovation step or clarity of thought), second, a closer thematic connection to major future developments within the research field, and, third, more visibility after publication. All these factors are likely to increase an article's influence on future research. To measure influence, we use the number of citations that an article receives by later publications, some written by economists from the sample but most by other researchers.[1]

We identify a research article's thematic links to the local environment using yearly biographical data on around 1,000 economists. The data are collected from online CVs and begin at the undergraduate level. This re-

[1]The measure is discussed in the empirical strategy section, in particular possible biases and strategic citation behavior.

veals personal connections between researchers and also the timing of these connections relative to the publication of their articles. For instance, two economists might be former university colleagues or future co-authors. In addition, bibliometric data on publications by these economists allow to locate their articles thematically by the references it gives to previous research. For instance, an empirical study on music consumption might cite theoretical trade papers, as well as research on the estimation of gravity models and empirical evidence on cultural industries. Combined, a reference to a university colleague shows the overlap of the idea (or thematic) space and the local, personal space at the university.

The identification of causal effects, using these observational data, builds on three main components. First, the comparison of individual articles, instead of more often considered yearly aggregates, allows us consider the variation in citation counts within publications by the same author team, journal type and publication year. In this limited empirical space, the estimate of the counterfactual - the outcome of an initially similar research project but without a thematic link to university colleagues - is credible enough to show a causal impact, based on the variation in correlations observed and described in the following paragraphs. While articles that reference university colleagues have a high unconditional correlation with citation counts, the attribution of this difference to the local environment is hasty, if we ignore author quality and other characteristics.

Second, we vary the timing of personal connections to test our estimates within the empirical setting. For instance, a future colleague, who joins the university after the publication of the article is not part of the local space during the research phase. We will show that such colleagues do not significantly influence the success of an article even if they are referenced in the article. By contrast, colleagues, who join before publication, and long-term colleagues increase the citation count. In addition, we vary the timing of the references that an article gives to previous research. For instance, we can classify thematic relations while only using research published before two researchers become university colleagues. Thereby, we make sure that earlier, thematically connected publications are not a reflection of the same local environment.

Third, we compare the estimates for articles in relatively high-profile journals with the estimates for more routine work. We do not expect a significant impact on the quality of routine work (see the conceptual framework in Section 2).[2] For instance, relatively simple 'routine' research should require less

[2]We do not consider publication quantity. However, visibility and networks might be factors for the distribution of routine work. See Section 4 for a more detailed discussion.

cutting-edge research knowledge and receive less exposure to and feedback from local colleagues. Our empirical definition of routine work is based on the journal and therefore approximative. In turn, we vary the cut-off for routine journals, which is based on the eminence of the researchers. As a consequence, we are mostly interested in the effect observed for high-profile publications that are assumed to be challenging and high-effort for its authors with and without thematic links to local colleagues.

We find that articles in high-profile journals receive a significantly higher (around 30%) number of citations by future research if they reference university colleagues. This increase in citation counts is conditional, however, on the journal type, the researchers and their affiliations. Therefore, advantages of a university department that are not thematically linked, such as general reputation and working environment, are not considered. Also, the benefit of potentially easier access to high-profile journals is not considered, since we compare publications within a journal or journal type.

More routine research is not affected in our empirical setting. This indicates that our results are unlikely to be driven by visibility or other direct dissemination effects. The impact of colleagues, who have roughly the same seniority, is higher than the impact of more experienced or more published researchers. The overall effect increases with the number of colleagues referenced, when estimating one, two, and three(-plus) thematically connected colleagues separately. The number of not connected local star colleagues has no effect in our setting. Former co-authors are linked with the outcome of routine research but not key publications. Both non-results, for co-authors and overall number of authors, might be due to the use of fixed effects for the author teams, as both groups strongly reflect individual characteristics. In addition, some network effects such as preferential access to publication would lead to negative estimates if, for instance, weaker papers are published.

The novel results in this paper connect to the literature in two ways. First, they can be understood within the theoretical framework given in Bobtcheff et al. [2016]. Here, in a priority competition for research innovation, a beneficial research environment leads to a longer maturation process and higher quality at publication. This theoretical result is supported by the findings in this paper.

Second, the paper links earlier empirical results. It shows that localized effects depend on the thematic proximity between colleagues. Similar to previous findings, the paper sees no effect between local colleagues who are not thematically linked. This helps explain why, on the on hand, most studies find no overall effect of researchers on the productivity of local colleagues or vice versa. On the other hand, the findings in this paper agree with recent research that shows positive spill-overs between researchers who are both personally *and* thematically connected, for instance co-authors or PhD students and their supervisors, on individual productivity (see Figure A.1 and *inter alia* Agrawal et al. [2017], Borjas and Doran [2015], Azoulay et al. [2010], Waldinger [2010], Waldinger [2012] and Waldinger [2016]).[3]

Specifically with a focus on economics, the findings on thematic localized effects provide context to two recent articles that observe no general localized effects within university departments in the last twenty years. Kim et al. [2009] study the effect of being affiliated with a top American economics or finance department. In their empirical setting, a researcher's annual productivity is estimated including university fixed effects. These affiliation fixed effects are positive for the 1970s but are insignificant in the 1990s. As a consequence, they conclude that localized peer effects disappear with the decrease of communication costs. Similarly, Bolli and Schläpfer [2015] conclude that German economists' overall productivity did not profit on average from moves to new institutions between 2004 and 2008. These results are confirmed for thematically unconnected colleagues.

The remainder of the paper is organized as follows. Section 2 describes the theoretical framework for local effects in knowledge production and dissemination. It sets a conceptual framework in which the impact of local colleagues can be empirically evaluated. Sections 3 and 4 present the data used and describe the empirical identification strategy. Section 5 presents empirical findings and Section 6 concludes.

[3]Agrawal et al. [2017] decompose the effects of a department hiring a star researcher in evolutionary biology into effects on overall research output by related and unrelated incumbents as well as new hires. While no positive effect is shown on incumbents overall, incumbents who work on similar research questions as the hired star increase their annual article output on average. The paper's main focus then is on the effect on new hires. Borjas and Doran [2015] study the potential negative effects resulting from the exodus of mathematicians after the end of the Soviet Union on the overall article output of collaborators left behind and previously geographically or thematically close researchers. Out of these groups, only former collaborators of highly productive emigrants appear to have been negatively affected. Azoulay et al. [2010] show the negative effects of the unexpected death of a scientific star on the productivity of collaborators in life science. Waldinger [2010] shows the importance of the quality of the supervisor on the lifelong productivity of PhD students in a sample of pre-World War II mathematicians. Waldinger [2012], in a later paper, finds no aggregated localized peer effects within German physics, chemistry or mathematics departments before the Second World War. Conversely, Waldinger [2016] finds negative short and long term impacts on the overall output of departments of the dismissed scientists.

2 Conceptual framework

The following conceptual framework helps us to distinguish potential mechanisms based on two assumptions. First, researchers rely less on their colleague's help for routine work. Second, outside influences on the take-up of an article, for instance through extra visibility by a well known or connected author, has a larger relative impact on low-profile, routine publications, if not related to the quality of the article.

The model of scientific production tested in this paper assumes that a researcher engages in both of two types of research. First, "cutting-edge" research that leads to a high-profile publication and is challenging to the authors and, second, "routine-type" research that is less prominently published and does not require the same effort. The research process for either type starts with an idea. The researcher, then, decides whether to pursue the idea. Next, she decides on the level of effort for the project depending on the quality potential. Quality depends on the innovation step and relevance to the research area. Eventually, effort, length of the production process, and quality will correspond to the publication type: a high- or a low-profile journal publication.[4]

The local environment, including colleagues and infrastructure, can impact on this stylized research process. For example, common seminars can bring access to new ideas and feedback and, thereby, increase the publication quantity and quality. Focusing on an individual article, cutting-edge research that challenges the authors has a larger scope for quality improvements by local factors than routine-type research. First, the authors are more likely to seek feedback and over a longer time frame. Second, more parts of the research are based on new, potentially tacit knowledge and the authors are less likely to posses all necessary research knowledge beforehand. Third, more open questions in interpretation, theory or method give more scope for an impact by others, for instance through discussion or complementary knowledge. This quality improvement can also have a direct social component as, for example, linking with concurrent research and its relevance to future research. On the other hand, routine-type research requires less time, effort, and innovation and has, therefore, less direct scope for quality improvement.[5]

[4]The eventual impact and publication types are understood to be relative the author's standard.

[5]See Bobtcheff et al. [2016] for a formal model of a research process in which a beneficial research environment leads to higher publication quality. The authors model a researcher's decision between early publication and quality maturation in a priority contest. Within their framework, a factor that makes the research process more efficient leads to more quality maturation and, subsequently, to a higher quality of the published research.

Taken together, this motivates the assumption that, first, local research linkages increase the efficiency of the research process, and, second, a significant impact on quality is expected for high-profile research only. This assumption is posited for the empirical evaluation as follows:

Assumption 1. *The quality of an article is a function of author ability (A) and additional factors (X) as $h(A, X)$, with h strictly increasing in both up to an inherent quality maximum $\widehat{Q}$:*

$$Q = \min\left(\widehat{Q}, h(A, X)\right).$$

In addition, for a routine-type article: $\forall x : h(A, x) \geq \widehat{Q}$ and, therefore, $Q = \widehat{Q}$ and for a cutting-edge type article: $\forall x : h(A, x) < \widehat{Q}$ and, therefore, $Q = h(A, x)$.

Assumption 1 states that the knowledge and effort of the authors are sufficient to produce the quality at the time of publication for a routine-type article. Conversely, the quality of a cutting-edge article benefits from additional input which the authors are also more likely to seek.

The success or influence I of an article is expected to depend positively on the quality of the article: $I'(Q) > 0$. Then, Assumption 1 leads to the empirical hypothesis that if a factor x helps the authors to improve the quality of an article behind their isolated efforts, then the semi-elasticity of the influence of the article at a given value of x is higher for a high-profile than a low-profile article.

Corollary 1. *Following from Assumption 1, if a variable x affects the measured influence (take-up) of a research article foremost through making the research process more efficient and enabling the authors to improve the article's quality behind their isolated efforts, then,the semi-elasticity of the influence I of an article at x is greater for a high-profile (cutting-edge) than a low-profile (routine) article, that is with quality $Q(x)$ and journal type J:*

$$\frac{\partial log(I(Q(x, j^{high}), j^{high}))}{\partial x} > \frac{\partial log(I(Q(x, j^{low}), j^{low}))}{\partial x} = 0$$

Conversely, personal connections may be correlated with the success of an article without contributing to the article's quality. A researcher's connectedness could, for instance, correlate with her or her department's (sub-field specific) reputation.[6] Then the opposite pattern is expected. Routine-type

[6]Further examples of possible dissemination benefits include strategic citation behaviour or a higher awareness of the research of personally connected researchers.

research in a low-profile journal should have a *relatively* larger reputation or visibility gain from the author or author's department. Cutting-edge research if published in a high-profile journal should benefit relatively less from the additional reputation gain.

On the other hand, if a higher reputation gain affects overwhelmingly high-profile publications, we cannot distinguish between the two posited effects. To address this concern the estimates for local colleagues will be compared to analogous estimates using the network of former co-authors (See Section 4). Former co-authors serve here as an expression of former research and personal connectedness in the specific area and are used to test Assumption 2.

Assumption 2. *The influence (take-up by future research) of an article is set as the function $I = I(R(Z), J)$, where Z is a variable that affects I foremost through the reputation R of the authors (or affiliation) and $J = \{j^{low}, j^{high}\}$ is the profile of the article's journal. Then, the following relations are posited.*

- *A high-profile article has more influence than a low-profile-article all else equal.*
$$I(R(z, j^{low}), j^{low}) < I(R(z, j^{high}), j^{high})$$
- *A possible difference in the derivatives is dominated by this overall difference of a high- and low-profile article.*

$$\frac{I(R(z, j^{high}), j^{high})}{I(R(z, j^{low}), j^{low})} \gtrapprox \frac{\frac{\partial I(R(z, j^{high}), j^{high})}{\partial z}}{\frac{\partial I(R(z, j^{low}), j^{low})}{\partial z}}$$

Assumption 2 leads to the empirical hypothesis that if a factor Z helps the visibility of an article without benefiting high-quality research more strongly, then the semi-elasticity of the influence of the article at a given value of z is lower for a high-profile than a low-profile article (or similar).

Corollary 2. *Following from Assumption 2, if a variable Z affects the measured influence of a research article foremost through correlation to the author's or department's reputation without a direct impact on quality, the semi-elasticity of the influence I of an article at z is smaller for a high-profile (cutting-edge) than a low-profile (routine) article, that is with reputation $R(z)$ and journal type J:*

$$\frac{\partial log(I(R(z, j^{high}), j^{high}))}{\partial z} < \ or \ \lessapprox \frac{\partial log(I(R(z, j^{low}), j^{low}))}{\partial z}$$

Taken together, the assumptions 1 and 2 state that a positive impact of a local cluster on the quality of an article leads to a positive correlation between citations received and references to local colleagues in a high-profile article. However, no positive correlation is expected for low-profile publications.[7] Conversely, if the cluster mostly reflects the field-specific reputation of the authors, we would observe higher (or similar) estimates for low-profile publications.[8]

3 Data

The empirical analysis is carried out using 7,291 publications by 967 highly cited economists between 1996 and 2014. In total, 28,901 research articles are listed on Scopus for the economists in this study. However, the baseline estimation uses a restricted data set that excludes articles that are co-authored with researchers outside of the sample. In addition, articles in the Journal of Economic Literature and articles that cite less than five or more than 80 other articles are excluded. The former often represent later summaries of earlier work, while the latter may also indicate work that mainly summarizes the preceding literature.

For the economists selected, complete information on the research career from the undergraduate studies onward was compiled using on-line CVs and the encyclopedia Who's Who in Economics (Blaug and Vane [2003]). The inclusion of authors is based on citations to publications in all economic journals listed by Kalaitzidakis et al. [2011]. These over two hundred journals are supplemented with a number of other, highly ranked journals in Ideas RePEc, for instance the later established AEA American Economic Journals, which brings the total number to 255 journals.

We ranked all authors of articles in these journals by the number of citations received as reported by Scopus.[9] From this exercise, a total of 1,300 economists are chosen based on work published in the period 1996 to 2014 and most highly cited in this period. Out of the 1,300 top ranked economists, we have sufficient information on 967 economists to include in our final dataset. For the others, we either failed to locate a complete CV or to attribute the combination of initials and surname to a single economist.

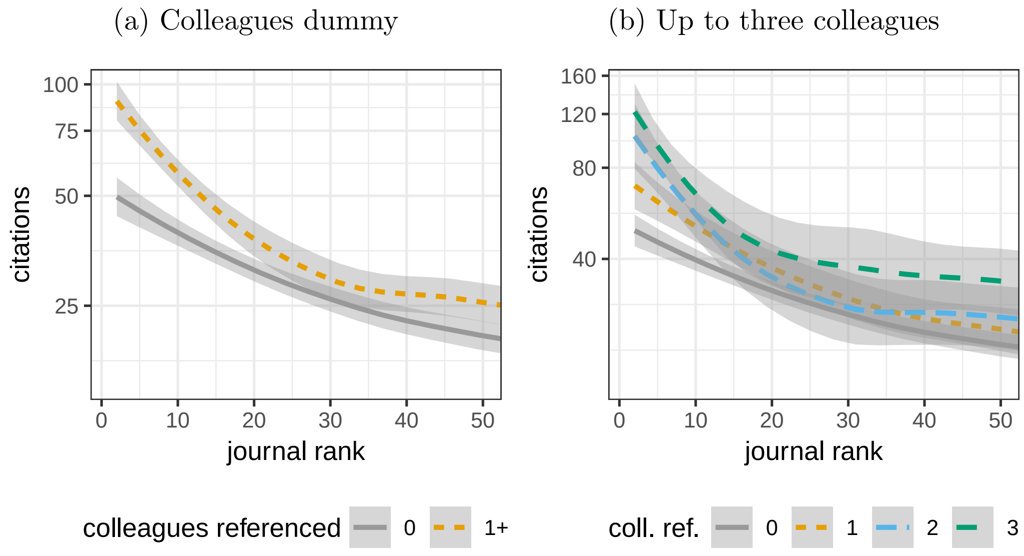
As an illustration of the spatial distribution of the universities studied,

[7]The impact might be more on quantity than quality for this category.

[8]The empirical results are robust to alternative 'high-profile' specifications and, therefore, not reliant on the theoretical framework.

[9]See `scopus.com`. This count includes citations by articles in any journal indexed on Scopus.

Figure 1: Citations by journal rank



Notes: Figure 1 shows smoothed estimates for citations per author received by articles in journals that are ranked by citations. Figure 1a shows mean estimates for article with no colleague referenced or one and more colleagues referenced. Figure 1b shows estimates for one to to three-plus colleagues. Smoothed means and confidence intervals are based on univariate loess regressions.

the affiliations of the selected economists are shown in Figure A.2. While many countries are observed, there is still a strong concentration on North-American universities. Within the sample, economists from Harvard and Berkeley account together for over ten per cent of the article output and the ten universities with the highest share of articles are in the USA. Education is even more concentrated; twenty-one per cent of the sample economists hold a PhD from Harvard or MIT and the top six universities account for forty-four per cent of the doctorates.

Stylized facts: university colleagues and citations counts in raw data

For the main regression data, Figure 1 shows the basic statistical relationship between citations received by an article, the journal in which the article is published, and the number of university colleagues referenced in the article. Publications in top journals receive on average twice as many citations per author if they reference university colleagues. The relationship is much weaker for less highly ranked journals. Looking at numbers of colleagues, we see that each additional colleague referenced is associated with an increase in citation counts. These estimates are obtained non-parametrically using local regressions.

4 Empirical strategy

The empirical estimation aims to identify a causal impact of the local environment. This section describes, first, the baseline estimation and construction of variables. Second, possible spurious effects due to omitted variables and selection, how these concerns are addressed and which exclusion assumptions are needed for identification.

Estimation

Equation 1 shows the baseline estimation of the total sum of citations I_i received by article i as a function of the authors personal links and article's i characteristics.[10]

$$\log\left(\mathbb{E}\left(I_i\right)\right) = \beta_1 C_i + \beta_2 C_i J_i + \beta_3 J_i + X_i \gamma + \alpha_i + \tau_i \tag{1}$$

In Equation 1, we estimate the number of citations I_i, that an article receives, based on a dummy variable C_i indicating whether a university colleague is referenced in this article. C_i is then interacted with a dummy for low-profile journals (J_i) to assess the correlation by type of research. In addition, X_i contains article specific variables, including the total number of references given, the number of references to earlier work by the authors, references given to economists at other universities, and on the ordering of the authors. Finally, we include fixed effects for the author team α_i and the publication year τ_i. In robustness checks, we additionally use the authors' affiliations and the individual journal as fixed effects.

We use a dummy variable as the main functional form for the effect of thematically linked colleagues since we, first, do not assume a constant effect for each additionally linked author. We will investigate this relationship later

[10]The empirical identification of local effects is based on the impact on the total sum of citations received per article (I_i). See, for instance, Tahamtan et al. [2016] for a literature review on factors that affect citation counts and Bornmann and Daniel [2008] or Osterloh and Frey [2014] for a discussion on impact and citation counts. Bornmann and Daniel [2008] give an overview of studies on the relation of citation counts and impact. While the citation behavior varies between researchers, the authors conclude that citation counts are generally a valid measure of impact. It is important to note that articles by the same authors are compared with a series of control variables. Conclusions based on citations would be problematic without the specific context.

though by using the number of colleagues referenced. Second, the number of references to university colleagues observed is small, mostly one or two. The limited number stems from the restriction to colleagues included in our sample of eminent economists. Only for these authors, we have full biographical data.

The construction of all colleague dummies is based on the following sets. E is the set of all authors of articles in our sample. E_i^{old} are all researchers in E that share an affiliation with an author of article i six to ten years prior to i's publication (shared affiliation in $\tau_i - 6, \tau_i - 7, \ldots$ or $\tau_i - 10$). For $E_i^{current}$, this time frame is one to five years prior to publication, and for E_i^{future}, one to five years after publication. R_i is the set of all economists referenced in article i. The dummy C_i in Equation 1 can then be written as

$$C_i = \begin{cases} 1 & \text{if } E_i^{current} \cap R_i \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Dummies for researchers who become colleagues during the periods before or after publication are defined as follows.

$$C_i^{new} = \begin{cases} 1 & \text{if } E_i^{current} \cap R_i \neq \emptyset \\ & \text{and } E_i^{old} \cap R_i = \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

and

$$C_i^{future} = \begin{cases} 1 & \text{if } E_i^{future} \cap R_i \neq \emptyset \\ & \text{and } (E_i^{old} \cup E_i^{current}) \cap R_i = \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

This implies that only researchers who become colleagues during the respective time period are counted in this alternative specification. For instance, researchers, who share an affiliation from four years before to two years after publication, would be counted as being new colleagues before publication but not after publication in this alternative classification.

Similarly, we construct variables that indicate whether former co-authors or other sample (and therefore eminent) economists are referenced. More detail on the construction of the variables and control variables is given in appendix C. In the main specification, J denotes whether the article is published outside a top 15, economics/finance, journal. Estimates are alternatively shown for other top definitions, the 'Top 5', and ranging from the five to the 75 highest ranked journals.

Identification issues and strategy

A causal interpretation of empirical effects of personal links to other researchers is not straightforward. The decision to link with the research of a colleague is not random. The following describes how the main concerns for an unbiased estimation, omitted variables and selection issues, are addressed. The identification of the effect is based on heterogeneity in three dimensions: and the timing of this personal link, the type of research published, and the publication date of the research referenced.

Selection into a department - quality

First, more productive researchers are likely to be selected into universities with other productive researchers. This motivates the focus on comparing a researcher's article to their other articles using author, university, and year dummies for differences in quality levels. Yearly quality fluctuations on the department level are addressed by including the overall number of eminent (sample) economists, and references to new colleagues directly after the publication date.

Selection into a department - thematic fit

Second, a researcher's forte might correspond to that of the department because of selection: The researcher selects or is selected into the department because of her prior research record or future research prospects. This can lead to her best research falling into the specialization area of her university colleagues. Importantly, this implies that the shared strength in a research area is not based on mutual support or advantages by the university environment but due to to selection. Without further controls, this would lead to a spurious estimate of the impact of the local environment.

This concern is addressed using two instruments. First, the estimates for university colleagues who share an affiliation during the research process, set as the five years prior to publication, are compared to estimates for colleagues who share an affiliation directly after the publication date or long before publication. Estimates for colleagues after the publication would correlate more strongly with a paper's success if thematic specialization accounts for most of the (spurious) effect. In turn, if we do not observe high estimates for new colleagues, then this channel is less likely.

In addition, references to former co-authors, and self-references, give further evidence for the role of research specialization for the estimates. An author's network of former co-authors relates closely to her past research record and personal links in a specific area.[11] References to former co-authors are, therefore, used to estimate the correlation between an article's success and its authors' specialization, personal connections and past research. If the estimated effect of colleagues is foremost an expression of research specialization, then these would lead to similar estimates for references to former and future co-authors.

Reflection

Positive effects between colleagues work both ways. Therefore, a colleague who works in a related area, might be as positively influenced by the authors of article i as they are by this colleague. In particular, high quality research by the authors of article i could lead to more related research by colleagues. While citations refer to earlier research[12], it cannot be excluded that the authors had influenced the work of colleagues before their own publication. Therefore, the timing of publications alone is not sufficient to rule out reverse causality. In addition, both the earlier related publications by colleagues and the higher quality of the authors' publication might both be a reflection of a common underlying factor within the university that is limited to field or topic.

We address this reflection by showing alternative estimates, that use only research published before the cited researchers become colleagues, to classify them as having worked on related topics. For comparison, we also count references to articles by colleagues that were published after they share, even temporarily, an affiliation with at least one author. The estimates obtained by these two ways of classifying colleagues are similar, as shown in the next section.

Field size and hot topics

A third potential issue arises from differing sizes of research areas. Suppose there are two main research fields in Economics. Field A and Field B. There are 10,000 researchers working in Field A, and there are 100 researchers working in Field B. Therefore, if you allocate a random sample of researchers in each department, you are 100 times more likely to have colleagues working in Field A and the odds of being cited also increase by 100 times, just because

[11] While these former co-authors are closely linked to other work by the researcher, they are not authors of the new article (the unit of observation). In addition, since the co-author network is an expression of the author's past research, the correlation with the success of an article does not necessarily indicate an influence of the co-authors.

[12] Recently, more and more citations might refer to contemporary publications.

so many more people are working on this field. This problem is aggravated, if researchers at prestigious universities - the typical colleague of our sample researchers - are more likely to work on widely studied research questions.

This concern is addressed by estimates for future colleagues, co-authors, not personally linked colleagues, and overall references. All of these would be affected by the omitted field size. They also work at similar universities and we see no significant difference in our estimates, when we include university fixed effects.

Unknown biases due to the construction of the colleague variable

Further mechanical effects due to the construction of the colleague variable C are addressed by several similarly constructed variables. The former co-author variable, for instance, reflects like C the decision to reference a personally connected researcher. If the estimation process leads to mechanical correlations with other factors, then an equivalent estimation process of this peer group and local colleagues leads to similar estimates. Diverging estimates support that the estimates are not dominated by common correlated aspects of these peer groups, time frame, or the estimation process.

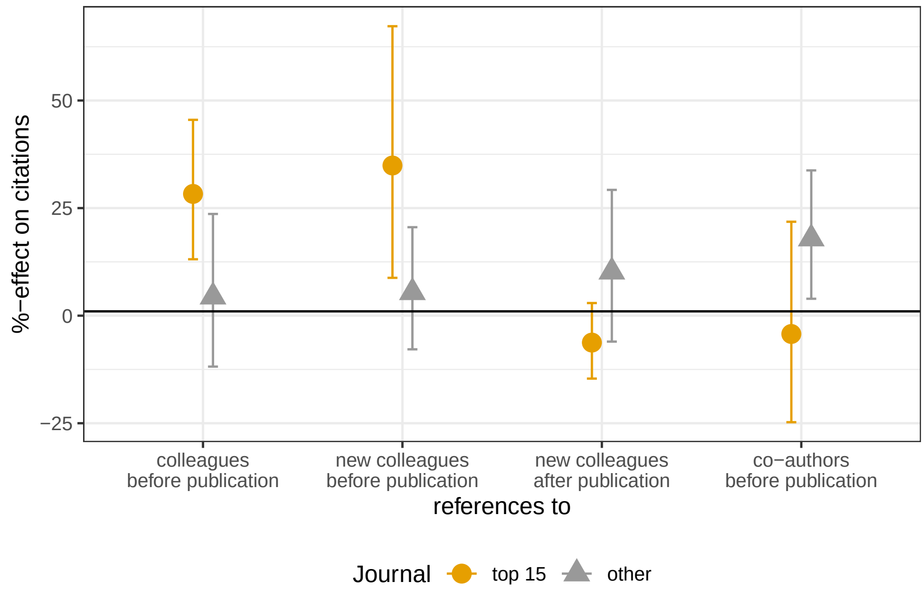
Third, the distinction made in the theoretical framework between a high-profile and a low-profile article is used to indicate possible general effects associated with the authors or universities that are expressed in the peer variables. For instance, the local colleague or co-author network may reflect the reputation of the authors which, in turn, increases the visibility of associated research. Relatively, this visibility gain would be more important for research that is published less prominently.

Conversely, exogenous factors that improve the quality of the paper and the immediate circulation and visibility are more likely to centered on high-profile publications, as described in Section 2. The two main arguments are the higher potential for improvement and the need for outside inputs on the one side, and more exposure of the research project to colleagues, for instance in seminars, on the other side.

Career age and research topics

Young researchers might be more likely to reference colleagues. First, they were hired more recently and their research interests should align, therefore, more closely with those of their colleagues'. Second, their research network is likely to be smaller than those of more established researchers. At the same time, researchers tend to be more productive at the beginning of their career (as measured in articles and citations received, see *inter alia* Oster

Figure 2: Key estimates from Table A.1



Notes: The dots and triangles in Figure 2 are point estimates (with 95% confidence intervals) for the statistical effect of references in a paper to different groups of personally linked researchers on the number of citations received by this paper. For all orange dots, this paper is published in a top 15 economics and finance journal. The grey triangles indicate a different publication outlet. Being a colleague or co-author is limited to either five to one year before the publication date or one to five years after publication. The first and last group includes references to any uni colleague or co-authors who shared an affiliation or co-authored a paper during said period. The two middle groups only include new colleagues, that is colleagues who did not share an affiliation in year six (before) or year zero (after publication). The regressions also include fixed effects per author team and year and further covariates as in shown in Table A.1.

and Hamermesh [1998]).

We, therefore, include career age in our estimation. We observe that the estimates hold for researchers within age groups and, in particular, for researchers who are more than five years past their PhD. Second, older and more experienced colleagues, relative to the authors, are estimated to have a lower impact on their colleagues than younger colleagues. Therefore, the results are not driven by young researchers emulating the topics of older or more experienced colleagues.

5 Empirical findings

The central finding of this paper is that thematically connected university colleagues impact on the success of challenging research. When comparing the publications from the same group of co-authors, we find that publications in top 15 journals receive around 30 percent more citations, if they reference researchers from our sample of eminent economists who are also recent university colleagues. We observe no effect on routine research. By contrast, researchers from our sample, who will become colleagues just after the publication, are not associated with higher citations counts for top publications in our setting, nor does connectedness in the specific field, proxied by the network of thematically connected, former co-authors.

These central findings are derived using quasi-Poisson (PPML) regressions shown in Table A.1. Figure 2 shows the main estimates from Table A.1. All estimates are derived while accounting for the general impact of the authors and specifics of the publications, including the number of references given to other researchers in our sample, the number of not-thematically connected colleagues at the university, and bibliometric measures.[13]

First, Figure 2 shows that a top 15 journal article receives around 30 % more citations if it is connected with the research of an eminent colleague who is at same university during the research period of the article.[14] This increase is relative to articles with similar characteristics, such as the number of references given and publication year, and the average of citations received by the group of authors, that is the authors publishing in this constellation. We understand this increase, therefore, as the benefit of having thematically connected university colleagues, a local idea space, during the research period.

We use the other estimates shown in Figure 2 to assess the causality and mechanism of this increase. First, we note that we see no positive correlation with citation counts in lower-profile journals. Our theoretical framework sets such publications as routine research for the sample of eminent economists, making their authors less likely to seek or benefit from assistance by local colleagues. A similar strong impact on routine work would have made it unlikely that our results are driven by quality improvements, for instance through access to tacit knowledge. However, we do not see an indiscriminate association between the local environment and citation counts of published

[13] All main estimates are derived using quasi-Poisson regressions using the natural logarithm as the link function. Alternative estimation techniques used (negative binomial, lognormal, normal with log transformed citation count) do not change any key result significantly. The regressions are carried out using the glm, feglm{alpaca}, felm{lfe}, and gamm{mgcv} functions in R. The packages multiwayvcov and alpaca (multi-way cluster-robust variance estimation after Cameron et al. (2011)) are used for standard error correction.

[14] The research period is set as five years prior to publication. Around 25 percent of the sample articles are published in a top 15 journal and of these 51 percent reference a university colleague. Twenty-seven percent of publications in other journals reference a university colleague.

research, which could have been caused by visibility, reputation, strategic citations and similar factors. This is again confirmed by the next estimate, that is references to future colleagues have no positive association with the number of citations received.

The second and third category in Figure 2 look at the effect of new and thematically connected colleagues. These researchers were not at the same university during the preceding five-year period. We see that researchers who become university colleagues during the five years preceding the article's publication impact citation counts, while those after publication have no impact. We take this as an indication that our main estimates are not caused by spurious effects due to general factors associated with thematically connected local colleagues. These contrasting findings support the argument for a causal impact of colleagues on the research and publication process, if we assume that colleagues directly before and after publication correlate similarly with these factors, for instance field-specific fit and quality in the department.

The last category shows estimates for former co-authors. This variable is as much an expression of the authors standing and past research in the field as of the potential influence of the former collaborators. There is, therefore, no clear interpretation for this variable but it shows the influence of connectedness in the research field on the estimates. In particular, a positive influence on top journal publications would be a concern for the validity of the estimates for local colleagues and the thematic framework. However, we observe a correlation only for publications in less eminent journals. Therefore, we can argue that in our empirical setting, in which we control for the general impact of the researchers involved, personal connectedness within the research area does not drive citation counts for top journal publications. Again, this supports the argument for a positive impact of local colleagues on the research process.

Table A.1 presents more detail on the regression estimates shown in Figure 2. In addition , the table provides further variations by including the key variables separately. We see that positive estimates for thematically connected colleagues before publication are robust to either including or excluding co-authors and future colleagues. The control variables used are the number of university colleagues five to zero years before publication in the sample. We use this variable in the form $log(x + 1)$ and find no effect in this and other transformations. We then use the number of sample economists referenced. The association with citations received is positive and significant though much smaller than for university colleagues. Articles that reference highly-cited economists are likely to be more in more researched areas, so this association is expected. This interpretation is supported by the observation that the estimate is similar for low- and high-profile journals. Next, we see a positive estimate for the overall number of references given. We see this variable as an indication for the size of the article. In fact, once we include references given, the number of pages per article is no longer significant in any specification and we do not retain both variables. We do not see a positive effect of self references given to earlier publications of the authors. Similarly, we do not see a significant effect for alphabetical ordering of author names.[15]

Thematically connected colleagues

Our baseline regression use dummies to estimate the impact of the local environment depending on whether a publication of university colleagues is referenced. In this section, we vary this key variable. First, we classify colleagues as thematically connected using only publications before they become colleagues to address reflection as a potential source for spurious estimates. Second, we estimate the impact of several thematically connected colleagues by using their number instead of a dummy.

Some positive effects within a department might also increase the number of earlier, thematically related publications by other members of the department. In turn, this would make references to colleagues more likely. Then, an article's references to local colleagues and its later success would be both expressions of this outside effect. Or, the authors of a later successful article could have helped earlier publications by their colleagues, which would make the estimate of thematic links on citation counts an expression of reverse causality, if still a local effect between colleagues. To sum up, the earlier quantity of related articles and the quality of the later article would both be caused by an unknown factor within the department.

We address these concerns in column (2) in Table A.2. Here, we introduce two alternative classifications. The first only uses publications before the authors share an affiliation (colleague classification before) to classify thematically connected colleagues. The second classification uses publications published after an author shares an affiliation (colleague classification after). The resulting estimates for both classifications are similar. This contradicts a significant impact by this kind of reflection on the estimates.

In our baseline estimates, we use a dummy to indicate one or more thematically connected colleagues for three reasons. First, few researchers have more than one thematically connected researcher. In the data used in the baseline regressions, three percent of articles have authors with more than

[15] See Kuld and O'Hagan [2018] for an interpretation of this variable. For instance, alphabetical ordering is convention in economic research and non-use might indicate an outsider.

three thematically connected colleagues, three percent have exactly three thematically connected colleagues, seven percent have two such colleagues, and 19 percent have one thematically connected author. Second, given these numbers the dummy variable gives us more robustness when estimating an effect for subgroups. Third, given our sample, one connected star colleague might already indicate a local cluster in their area, implying other thematically but less eminent researchers and related infrastructure.

Figure A.4 shows the positive link between an increasing number of uni colleagues referenced in a paper and the number of citations received by this paper.[16] Unlike Figure 2 and Table A.1, the number of colleagues is estimated for up to three colleagues. While already the first colleague is estimated to have a significant effect, the estimates for two and three-plus colleagues are considerably higher. All three point estimates for routine research are near and not significantly different from zero.

Top journals and top economists

Our estimation process uses two partly arbitrary definitions for 'top': journals and economists. First, our empirical framework makes a distinction between articles in high-profile or top journals and more routine research. In this section, we show variations of this distinction to gauge the robustness of the baseline results to different cut-off points. In addition, we estimate the effect size as a function over the rank of journals, defined by their median of citations received, to see how the impact of the local environment varies along the range of economic journals.

Column (1) in Table A.2 shows the effect size of the localized environment estimated as a linear function of journal rank. We see a positive effect estimated for highly ranked journals. The estimate decreases by 0.003 point for each rank. This implies a positive effect size up to rank 70.

In Figure A.4, seven estimates for different top journal definitions are shown on the right side. As expected, the estimates decrease for very wide definitions of what constitutes a top journal. However, the top journal definition is robust over a certain range and does not rely on the narrow definition made in the main regression tables shown above.

Second, we use a sample of researchers who are among the 1,300 most cited authors of articles in economics journals. In Figure A.5, we estimate the impact for different sub-samples to understand how our findings vary with the eminence of a researcher. We see no systematic difference in the estimates for

[16] The estimation is otherwise the same as in specification (1) in Table A.1, including covariates.

subsets of economists, classified by their overall number of citations received. However, the smaller sample size increases the uncertainty in the estimation. We conclude that the estimated effect is not driven by a subset of authors and, therefore, robust to sample selection, which supports the external validity of our findings.

Which colleagues matter: age, experience

In Table A.3, we compare, first, the effect of colleagues who have roughly the same career age (time since PhD) as the authors with the effect of colleagues who are six and more years older or younger in career age. Second we compare the effect of colleagues who have been at the university for (± 5 years) the same length as the authors with those who have been at the university either shorter or longer. Last, we compare the effect of colleagues who have more Top 5 journal articles published with those who have published as many or fewer than the authors of the article at this point in time.

The effects are the stronger for colleagues who are more similar in age and experience. Table A.3 show the estimates for the variations discussed. First, specification (1) repeats the baseline estimates as we have slightly different estimates using career age. Specification (2) estimates the effect of colleagues who have roughly the same academic age, that is received their PhD not more than five years before or after the respective author. Column (3) shows the effect of the opposite group that is either considerably older or younger. Columns (4-6) consider the time spent at a specific university. Column (4) shows the significant and similarly sized effect of colleagues hired after the authors. The last two columns distinguish between colleagues who have published more Top 5 journals up to the previous year or the same number and fewer.

For all the categorizations considered, the effect is stronger for colleagues who are more similar, that is in academic or university age, or in their previous publication record. These findings suggest that we do not observe a gatekeeping or mentoring effect of more senior researchers helping less experienced researchers. Instead, the estimates indicate that local groups of researchers at similar stages in their career might be more beneficial for an individual researcher to produce influential research.

Robustness of estimation to assumed distribution and additional fixed effects

None of the alternative estimations shown in Table A.2 changes the baseline results significantly. Column (3) fits a negative binomial distribution

($\theta = 1.36$) instead of the Poisson distribution used before (both with clustered and, therefore, re-estimated standard errors). Column (4) shows the estimates from an OLS regression with ($log(y + 1)$) transformed citation counts. Column (5) fits a negative binomial distribution using the same θ parameter as column (6) in a generalized additive model. This allows to model nonparametrically the relation between citation counts and the control variables using splines to gauge the effect of a potential misspecification in their functional form.

Table A.4 shows estimates that include fixed effects for universities and individual authors instead of author groups. The alternative estimation confirms the main estimates. In addition, university colleagues ten to six years before publication date are shown to have a lower influence. This underlines the importance of a shared affiliation in the years directly before publication.

Caveats about the interpretation

The empirical design implies caveats in the interpretation of the results. First, the empirical analysis focuses on the general impact of local resources and research links between local colleagues. Endogenous peer effects and the impact of exogenous characteristics of colleagues or departments are not explicitly distinguished (see Manski [1993]). Therefore, the main focus is on assessing whether correlated effects such as selection effects are expressed in the estimates. Second, this paper focuses on individual articles instead of the overall article output of a researcher. The interpretation of the results is, therefore, limited on effects on the quality and dissemination of individual articles. Third, the estimates for the co-author network are used to support the causal interpretation of local colleagues. The estimates do not justify a causal interpretation between former co-authors and research quality. In particular, learning through co-authorship may lead to an overall higher research performance which does not show in the comparison of individual research projects.[17] Fourth, the estimated impact of local colleagues is con-

[17]Overall, the co-author network might impact more positively on individual productivity, in particular as direct co-authorship is not studied here (see for instance Azoulay et al. [2010], Borjas and Doran [2015], Ductor et al. [2014] and Ductor [2015]). In addition, the co-author network is arguably a stronger reflection of a researcher's past productivity than local colleagues, if the affiliation is controlled for. In turn, this lower dependence on past performance is used here to estimate impact of local colleagues outside the researcher's co-author network. The negligible impact of university dummies on the estimated effect of colleagues reaffirms with the posited low dependence. On the other hand, co-authors may be important for the generation of new ideas but that the quality of high-profile publications depends more on the efficiency of the maturation (Bobtcheff et al. [2016]) process which may be more strongly influenced by local colleagues.

centrated around articles that are already highly cited.[18] However, the type of article affected is likely a product of the chosen sample of highly cited economists. Routine work means different things for different researchers. Last, some positive effects of colleagues, such as ease of publication, would lead to negative estimates here. If we control for the journal type, an easier way into a top journal may lead to quality which is relatively low compared to the journal's standards. Also, a resulting higher quantity might be to the detriment of individual quality. These benefits might be stronger for links to co-authors and more senior colleagues and explain the lower estimates of the latter group.

6 Conclusion

The economists studied in this paper produce more influential research when working on topics related to the research of their department colleagues. Therefore, this paper argues for localized benefits if researchers can link their research. This does not necessarily contradict the observed decline in localized peer effects. The absolute number of eminent colleagues is not associated with an increase in citations received if there is no connection in their research. Moreover, the decrease of localized effects in economics departments observed elsewhere (Kim et al. [2009]) might be partly explained by a decrease of interaction between local colleagues or an increase in specialization (Jones [2009]).

The productivity benefit observed for local colleagues who work on related research questions could be important in the organization of research entities. However, the empirical setting studies the benefits of research links, which might connect different research fields. The results show, in particular, the benefit of thematic links to local colleagues outside the authors' co-author network.

[18] For instance, more than 50 citations received on Scopus (Scopus typically reports fewer citations than for instance Google Scholar).

References

Agrawal, Ajay, John McHale, and Alexander Oettl, "How stars matter: Recruiting and peer effects in evolutionary biology," *Research Policy*, 2017, *46* (4), 853 – 867.

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

Blaug, M. and H.R. Vane, *Who's Who in Economics*, Edward Elgar Pub., 2003.

Bobtcheff, Catherine, Jérôme Bolte, and Thomas Mariotti, "Researcher's Dilemma*," *The Review of Economic Studies*, 2016.

Bolli, Thomas and Jörg Schläpfer, "Job Mobility, Peer Effects, and Research Productivity in Economics," *Scientometrics*, September 2015, *104* (3), 629–650.

Borjas, George J and Kirk B Doran, "Which peers matter? The relative impacts of collaborators, colleagues, and competitors," *Review of Economics and Statistics*, 2015, *97* (5), 1104–1117.

Bornmann, Lutz and Hans-Dieter Daniel, "What do citation counts measure? A review of studies on citing behavior," *Journal of Documentation*, 2008, *64* (1), 45–80.

Ductor, Lorenzo, "Does Co-authorship Lead to Higher Academic Productivity?," *Oxford Bulletin of Economics and Statistics*, 2015, *77* (3), 385–407.

_ , Marcel Fafchamps, Sanjeev Goyal, and Marco J. van der Leij, "Social Networks and Research Output," *The Review of Economics and Statistics*, December 2014, *96* (5), 936–948.

Jones, Benjamin F., "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?," *Review of Economic Studies*, 2009, *76* (1), 283–317.

Kalaitzidakis, Pantelis, Theofanis P. Mamuneas, and Thanasis Stengos, "An updated ranking of academic journals in economics," *Canadian Journal of Economics*, November 2011, *44* (4), 1525–1538.

Kim, E. Han, Adair Morse, and Luigi Zingales, "Are elite universities losing their competitive edge?," *Journal of Financial Economics*, 2009, *93* (3), 353–381.

Kuld, Lukas and John O'Hagan, "Rise of multi-authored papers in economics: Demise of the 'lone star' and why?," *Scientometrics*, March 2018, *114* (3), 1207–1225.

Manski, Charles, "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Oster, Sharon M. and Daniel S. Hamermesh, "Aging and Productivity Among Economists," *The Review of Economics and Statistics*, 1998, *80* (1), 154–156.

Osterloh, Margit and Bruno S. Frey, "Academic rankings between the 'republic of science' and 'new public management'," in Alessandro Lanteri and Jack Vromen, eds., *The Economics of Economists*, CUP, 2014, pp. 77 – 103.

Tahamtan, Iman, Askar Safipour Afshar, and Khadijeh Ahamdzadeh, "Factors affecting number of citations: a comprehensive review of the literature," *Scientometrics*, 2016, *107* (3), 1195–1225.

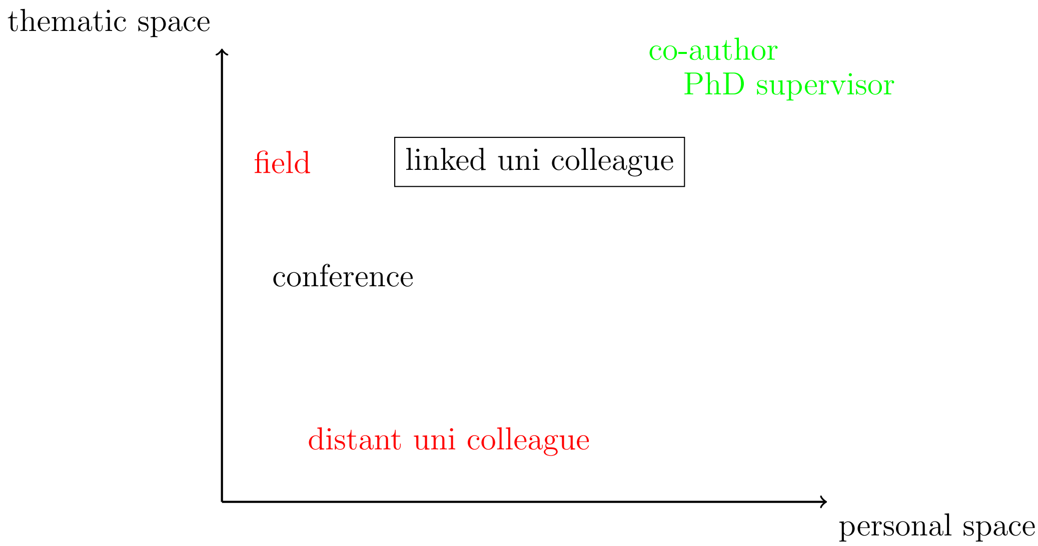
Waldinger, Fabian, "Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany," *Journal of Political Economy*, 08 2010, *118* (4), 787–831.

_ , "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany," *Review of Economic Studies*, 2012, *79* (2), 838–861.

_ , "Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge," *Review of Economics and Statistics*, December 2016, *98* (5), 811–831.

A Figures

Figure A.1: Productivity effects in science: thematic and personal space



Literature: Effect: co-authors (Azoulay et al. [2010], Borjas and Doran [2015]), PhD supervisors (Waldinger [2010]). No effect: field (Borjas and Doran [2015]), local/university (Kim et al. [2009], Waldinger [2012], Waldinger [2016], Borjas and Doran [2015], Bolli and Schläpfer [2015], Agrawal et al. [2017])

Figure A.2: The affiliations of the authors in the sample.

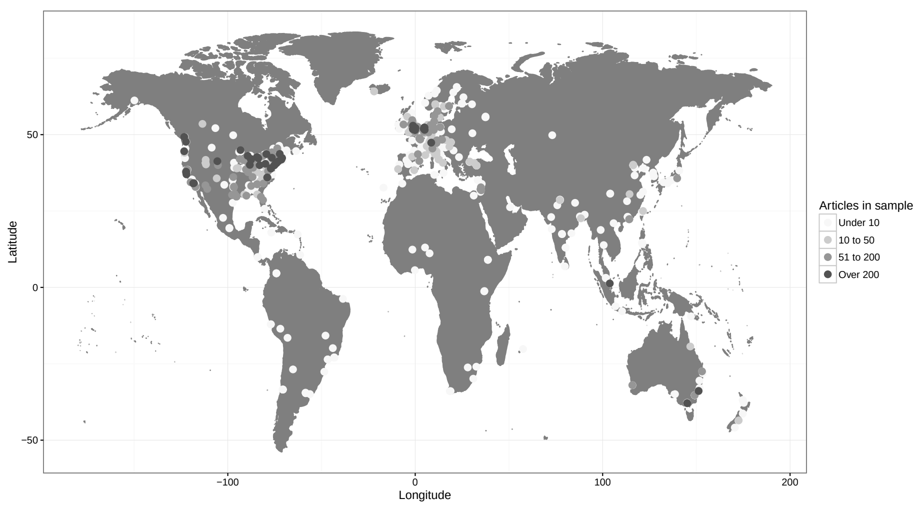


Figure A.3: Thematic and personal relations in an article citation graph.

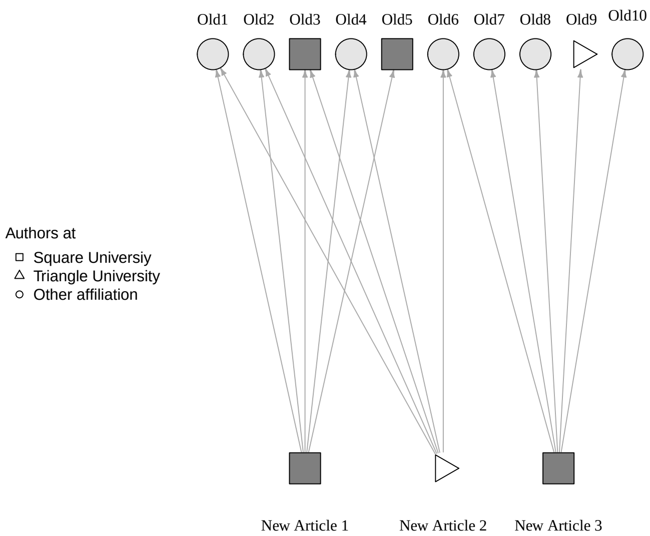
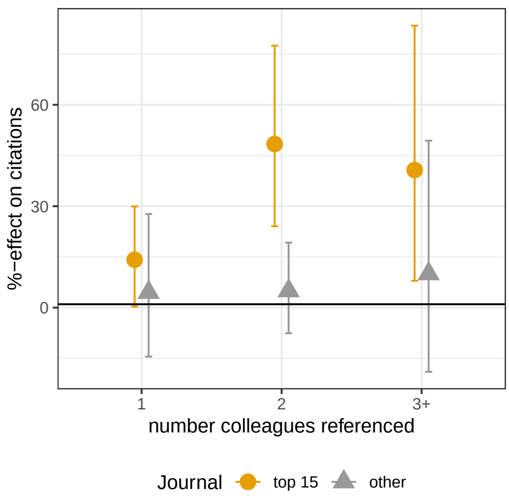
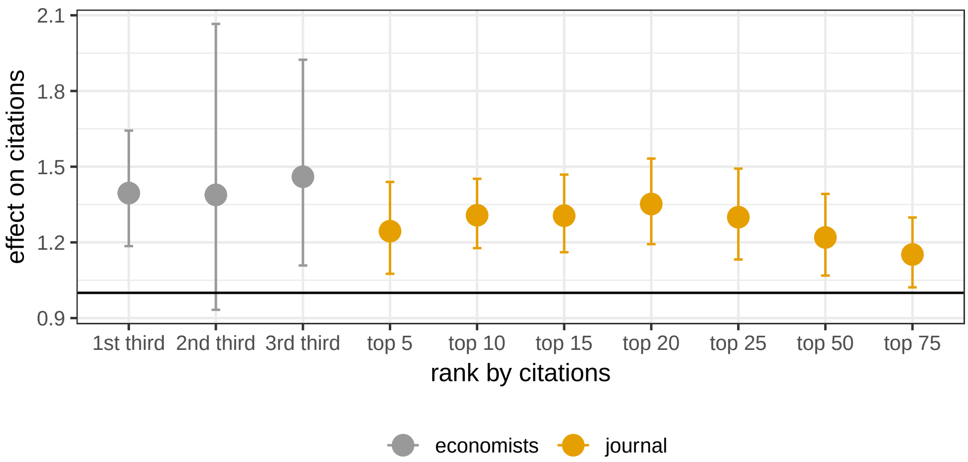


Figure A.4: Effect by number of colleagues



Notes: The dots and triangles in Figure 2 are point estimates (with 95% confidence intervals) for the statistical effect of different numbers of colleagues referenced in a paper on the number of citations received by this paper. For all orange dots, this paper is published in a top 15 economics and finance journal. The gray triangles indicate a different publication outlet. The regressions also include fixed effects per author team and year and further covariates as in shown in Table A.1, specification (1).

Figure A.5: Variation of top journal definition and sample split by economists' total citations



Notes: The dots in Figure 2 are point estimates (with 95% confidence intervals) for the effect of colleagues referenced in a top journal paper on the number of citations received by this paper. For all orange dots, the estimation is repeated in three samples defined by the total number of citations received. For instance, the first dot does not include estimates for publications with an author outside of the top 323. The gray dots indicate estimates for articles in top journals. The definition is based on average citations and varied from top 5 to 75. The regressions also include fixed effects per author team and year and further covariates as in shown in Table A.1, specification (1).

B Tables

Table A.1: Baseline estimates

	Dependent variable:				
	citations received *(link: log)*				
	(1)	(2)	(3)	(4)	(5)
colleague referenced (C)	0.27***	0.27***			
	(0.06)	(0.06)			
new colleague referenced (C^{new})			0.32**		0.33**
			(0.11)		(0.11)
future colleague referenced (C^{future})				−0.06	−0.10
				(0.07)	(0.07)
former co-author referenced (CA)		−0.06	−0.05	−0.03	−0.05
		(0.13)	(0.13)	(0.13)	(0.13)
lower profile journal (J)	−0.63***	−0.65***	−0.69***	−0.77***	−0.71***
	(0.10)	(0.10)	(0.11)	(0.11)	(0.10)
$C \times J$	−0.23	−0.26*			
	(0.12)	(0.12)			
$C^{new} \times J$			−0.28*		−0.29*
			(0.12)		(0.12)
$C^{future} \times J$				0.15	0.19
				(0.11)	(0.12)
$CA \times J$		0.22	0.22	0.19	0.22
		(0.12)	(0.12)	(0.12)	(0.12)
log(number colleagues + 1)	0.00	0.00	0.00	0.02	0.00
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)
sample economists referenced	0.04***	0.04***	0.04***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
sample economists referenced$\times J$	0.00	−0.00	−0.00	−0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
number references	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
number self-references	−0.00	−0.00	−0.01	−0.00	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
alphabetically ordered authors	0.17	0.15	0.16	0.16	0.16
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
N (articles)	7030	7030	7030	7030	7030
FE author group	yes	yes	yes	yes	yes
FE publication year	yes	yes	yes	yes	yes

Notes: This table reports estimated coefficients from quasi-Poisson regressions with standard errors in parentheses. The standard errors are clustered at the author team, year, and journal level. For more detail on the variables see Section C.

*p<0.05; **p<0.15; ***p<0.001

Table A.2: Robustness

	Dependent variable:				
	citations received *(link: log)*				
	(1)	(2)	(3)	(4)	(5)
	q-P	q-P	NB	OLS	GAM
colleague referenced	0.210***		0.246***	0.267***	0.210***
	(0.063)		(0.030)	(0.042)	(0.069)
colleague referenced*journal rank	−0.003**				
	(0.001)				
journal rank	−0.006***				
	(0.001)				
colleague referenced*low-profile journal			−0.208**	−0.273**	−0.124
			(0.076)	(0.083)	(0.077)
lower profile journal		−0.623***	−0.847***	−0.810***	−0.800***
		(0.098)	(0.110)	(0.130)	(0.055)
colleague classification before		0.173***			
		(0.048)			
colleague classification after		0.243**			
		(0.076)			
colleague classification before*low-profile journal		−0.230***			
		(0.061)			
colleague classification after*low-profile journal		−0.241			
		(0.142)			
log(number colleagues + 1)	0.032	0.004	0.017	0.047	* * *
	(0.058)	(0.059)	(0.031)	(0.037)	
sample economists referenced	0.043***	0.036***	0.034***	0.035**	* * *
	(0.010)	(0.010)	(0.010)	(0.011)	
sample economists referenced*journal rank	0.000				
	(0.000)				
sample economists referenced*low-profile journal		0.002	0.013	0.006	
		(0.013)	(0.011)	(0.013)	
number references	0.011***	0.013***	0.016***	0.018***	* * *
	(0.003)	(0.003)	(0.002)	(0.002)	
number self-references	0.001	−0.004	0.004	0.001	* * *
	(0.013)	(0.012)	(0.007)	(0.007)	
alphabetically ordered authors	0.236	0.157	0.082	−0.029	
	(0.155)	(0.147)	(0.087)	(0.085)	
N (articles)	5406	7038	7038	7058	7058
FE author group	yes	yes	yes	yes	RE
FE publication year	yes	yes	yes	yes	* * *

Notes: This table reports estimated coefficients from quasi-Poisson (1-2),negative binomial (3), OLS (4), and GAM (negative binomial) (5) regressions with standard errors in parentheses. The standard errors are clustered at the author team, year, and journal level. Column (4) uses the following transformation of citations received as the dependent variable: $log(citationsreceived + 1)$. For more detail on the variables see Section C.

*p<0.05; **p<0.15; ***p<0.001

Table A.3: Colleague variations

	Dependent variable:							
	citations received *(link: log)*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
coll definition	all	± 5 (age)	$\lessgtr 5$ (age)	sh. uni	lo. uni	sa. uni	$>$ top	$\leq$ top
coll.ref:top	0.24^{***}	0.29^{**}	0.25^{***}	0.21^{***}	0.15	0.28^{*}	0.14^{*}	0.28^{**}
	(0.06)	(0.10)	(0.05)	(0.05)	(0.08)	(0.11)	(0.07)	(0.10)
coll.ref:other	0.01	0.16	−0.06	−0.00	−0.12	0.04	0.04	0.02
	(0.09)	(0.12)	(0.08)	(0.08)	(0.07)	(0.12)	(0.09)	(0.08)
log(sample.coll)	0.00	0.01	0.01	0.01	0.02	0.01	0.01	0.01
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
min.age	−0.07	−0.08	−0.07	−0.07	−0.07	−0.06	−0.07	−0.07
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)
sample.econ.ref	0.04^{**}	0.03^{**}	0.04^{**}	0.04^{**}	0.04^{**}	0.03^{**}	0.03^{**}	0.03^{**}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
top	0.63^{***}	0.68^{***}	0.62^{***}	0.68^{***}	0.65^{***}	0.67^{***}	0.69^{***}	0.65^{***}
	(0.09)	(0.10)	(0.10)	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)
references	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
self.references	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
alphabetical	0.18	0.19	0.19	0.20	0.20	0.19	0.20	0.18
	(0.16)	(0.16)	(0.16)	(0.15)	(0.16)	(0.15)	(0.16)	(0.16)
sample.econ.ref:top	0.00	0.00	−0.00	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	6589	6589	6589	6589	6589	6589	6589	6589
author team	FE	FE	FE	FE	FE	FE	FE	FE
year	FE	FE	FE	FE	FE	FE	FE	FE

Notes: This table reports estimated coefficients from quasi-Poisson regressions with standard errors in parentheses. The standard errors are clustered at the author team, year, and journal level. The group of university colleagues is varied for each column. all: all colleagues; ± 5 (age): up to five years older or younger (academic age, years since PhD); $\lessgtr 5$ (age): outside this academic age range; sh. uni: shorter at uni; lo. uni: longer at uni; sa. uni: ± 5 years same time at uni; $>$ top: more top 5 publications up to previous year; $\leq$ top: less or same number of top 5 publications up to previous year.
coll.ref: colleague (5-1 years prior to publication) referenced, top: article is published in a top 15 journal, other: other journal, sample.coll: number of uni colleagues (5-1 years prior to publication) in sample, sample.econ.ref: number of sample economists referenced, references: total number of references given in article, self.references: number of references to earlier research by the authors, alphabetical: authors appear in alphabetical order, coauth.ref: co-author of another publication with an author referenced, new.coll.ref: colleague from 5-1 years before publication referenced who was not a colleague before this period, after.new.coll.ref: colleague from 5-1 years after publication referenced who was not a colleague before this period. For more detail on the variables see Section C.

*p<0.05; **p<0.15; ***p<0.001

Table A.4: Main estimates, alternative fixed effects

	Dependent variable:					
	SumCitationsReceived *(link: log)*					
	(1)	(2)	(3)	(4)	(5)	(6)
coll.ref:top	0.220**	0.237**	0.408***	0.231**		0.220*
	(0.102)	(0.102)	(0.084)	(0.113)		(0.117)
coll.ref:other	0.031	-0.018	0.003	-0.042		-0.027
	(0.070)	(0.069)	(0.089)	(0.079)		(0.075)
coauth:top		0.010	0.072	0.065		0.066
		(0.108)	(0.097)	(0.111)		(0.118)
coauth:other		0.193***	0.151*	0.116		0.073
		(0.074)	(0.085)	(0.078)		(0.083)
top 1	0.714***	0.728***	0.757***	0.732***	0.778***	0.732***
	(0.090)	(0.088)	(0.088)	(0.081)	(0.100)	(0.091)
log(sample.econ.ref)	0.249***	0.242***	0.242***	0.250***	0.252***	0.242***
	(0.055)	(0.056)	(0.055)	(0.054)	(0.052)	(0.054)
references	0.433***	0.431***	0.436***	0.432***	0.437***	0.432***
	(0.073)	(0.074)	(0.073)	(0.073)	(0.073)	(0.073)
alphabetical	0.483***	0.482***	0.470***	0.488***	0.494***	0.486***
	(0.090)	(0.089)	(0.090)	(0.090)	(0.091)	(0.089)
self.refernces	-0.009**	-0.010**	-0.011**	-0.009**	-0.011**	-0.010**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
after.coll:top			-0.140		-0.154	-0.176*
			(0.113)		(0.100)	(0.106)
after.coll:other			-0.116		-0.106	-0.076
			(0.082)		(0.065)	(0.064)
after.coauth:top			0.105		0.235**	0.206*
			(0.096)		(0.117)	(0.122)
after.coauth:other			0.136		0.188**	0.180**
			(0.095)		(0.082)	(0.089)
long.before.coll:top			-0.184**			
			(0.079)			
long.before.coll:other			0.015			
			(0.091)			
long.before.coauth:top			-0.207			
			(0.183)			
long.before.coauth:other			0.001			
			(0.141)			
Ind, Uni, & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Coll/Co-Auth Restr.	No	No	No	New	New	New
N (df)	7291 (6071)	7291 (6069)	7291 (6061)	7291 (6069)	7291 (6069)	7291 (6065)

Notes: This table reports estimated coefficients from quasi-Poisson regressions with standard errors in parentheses. The standard errors are clustered at the authors and journal level. Estimation is similar to Table A.1 but fixed effects are per author instead of author group and the estimates include university dummies.

*p<0.1; **p<0.05; ***p<0.01

C Data appendix: description of variables

Table A.5: Description of the regression variables

Name	Description	Range	Mean
citations received	The sum of citations received up to January 2015.	$[\![0,4200]\!]$	72.88
coll.ref	Article refers to at least one researcher who was a colleague of any author one to five years prior to publication.	$[\![0,1]\!]$	0.32
coauth.ref	Article refers to at least one researcher who was a co-author of any author one to five years prior to publication.	$[\![0,1]\!]$	0.23
top	A publication in one of the 15 most frequently cited economics journals (citations per article/divided by the yearly median).	$[\![0,1]\!]$	0.26
min.age	The time since completion of the PhD of the in this sense youngest authos.	[0,60]	16
references	The number of articles referenced, truncated at 5 and 80.	$[\![5,80]\!]$	28
self.references	The number of articles referenced that are written by an author of the article, truncated at 20.	$[\![0,20]\!]$	3.6
alphabetical	Authors are listed alphabetically.	$[\![0,1]\!]$	0.94

Notes: The Table shows a description, the range and the mean of all variables included in the regressions presented. The statistics refer to the main data set of 7,030 articles of which all authors are in the sample as described in the data section.

Variables

For each article in the data, variables on references to connected researchers are computed. The main variable of interest indicates the number of university colleagues who share an affiliation during the research phase and are referenced in the article. This research phase is set at five to one years before the publication date of the article. For ease of comparison, the variable is later aggregated to the dummy $R_i^{Colleague}$ indicating one or more colleagues referenced before being varied in several dimensions.

The main variations of the dummy for colleagues referenced $R_i^{Colleague}$ change the time of shared affiliation (in particular after or before publication date), the type of personal connection (colleague or co-author), and the relative seniority of the colleagues referenced (older, more publications, or longer at the university).

Following the conceptual framework, the colleague and co-author variables are evaluated separately for more and less high-profile journals. This classification is based on the mean number of citations received by all articles of a journal.[19] The estimation starts by setting the 15 top ranked economics and finance journals as high-profile The cut-off for top journal is varied in alternative specifications.

Control variables

The vector X contains a series of further characteristics of the article and its authors. First, academic age is counted as years since the completion of the PhD. This is included to account for career effects on productivity. Second, the number of affiliations, the number of authors and whether the authors are listed alphabetically is used to complement the individual and university dummies. A non-alphabetical ordering is unusual in economics[20] and can indicate a different background of the authors or authors added without a full contribution which would overestimate the number of authors. Third, the number of references and its squared value are included. More references can indicate a bigger project, more interest in the research area or increase the visibility of the publication independently.[21] The number of pages is not significant if the number of references is used and, subsequently, not used in the estimation.

Importantly, the total number of eminent sample economists referenced is included. Since the sample of economists was selected based on citations in economic journals, a high number of cited eminent economists indicates a research field that attracts a high interest by economic researchers. The interest in the topic could in turn cause a positive effect of peers referenced. However,, the main estimates for the influence of peers are not

[19] More specifically, to rank the 255 economics journals, the set of all research articles between 1996 and 2014 as described in the data section is used. First, all citations are divided by the yearly median. Then, the mean of these adjusted citations per article is used to rank the journals.

[20] In the time span studied, 92 per cent of articles published in top 5 journals that have more than one author list the authors alphabetically.

[21] For example, on-line databases make it possible to search citing papers. Therefore, the more references an article lists the more such searches include the article.

changed significantly after the introduction of this control variable. Finally, the number of self-references is counted to indicate prior experience in the area: self.references $= |M_i|$, where M_i is the set of references in i to articles by an author of i. Prior publications could indicate a higher visibility and linked prior experience should be helpful in the research process. On the other hand, self-references may be partially arbitrary or indicate follow-up work to main publications. There is, therefore no clear interpretation of this variable.