Research Policy 52 (2023) 104844

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Local peer communities and future academic success of Ph.D. candidates

Mignon Wuestman, Iris Wanzenböck *, Koen Frenken

Copernicus Institute of Sustainable Development, Utrecht University, Princetonlaan 8a, 3584, CS, Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Early-career scientists
Academic careers
Peer communities
Support
Competition
Academic mentorship
Genealogy

ABSTRACT

Compared to senior scientists, early-career scientists have largely been neglected in the literature on academic success. This study aims to identify the effects of local peer communities of Ph.D. candidates on their future careers. We argue that local communities of Ph.D. candidates may offer both supportive and competitive environments depending on the nature of the relationships between its members. While Ph.D. candidates generally learn from and support each other in their local peer communities, they may also compete for their mentor's attention and future academic positions. We analyse such complex peer effects for 90,264 Ph.D. candidates in the field of mathematics in a genealogical way, by measuring a candidate's academic career success by the number of next-generation Ph.D. candidates supervised later on. To capture both the supportive and competitive peer effects, we distinguish between local peers who share mentors (co-mentees) and other local peers. Our result suggests that competition exists primarily among peers who share mentors, and only at the start of one's career. We also find supportive effects among peers who do not share mentors, particularly those from the same cohort. Our results highlight the importance of universities supporting informal interactions among Ph.D. candidates.

1. Introduction

Developing support structures for the organization of science is important to promote the academic success of scientists and foster scientific advancements (Sverdlik et al., 2018). However, some aspects of the organization of science have received more attention than others in previous literature. While we know that team composition (Head et al., 2018; Lee et al., 2010; Singh and Fleming, 2010), formal mentorship (Berry, 1981; Chariker et al., 2017; Malmgren et al., 2010; Sugimoto et al., 2011), and individual circumstances (Ramesh and Singh, 1998) significantly affect academic success, studies thus far mainly focused on scientists in later stages of their academic career. The determinants of academic success in the early career phase remained under-researched (Broström, 2019; Cameron and Blackburn, 1981; Dundar and Lewis, 1998).

We argue that a clearer distinction between early and later career stages is important in studying academic success, as the success factors identified for senior scientists may not carry over to the context of early career scientists (Clemente, 1973). Early career scientists do not yet have significant professional networks to form new collaborations, or do not yet have an established presence and reputation. Instead, early career scientists, and Ph.D. candidates in particular, rely more on their local research environment composed of their direct mentors and peers within their university (Paglis et al., 2006; Tartari et al., 2014). Typically, these mentors and peers perform tutoring tasks for early career scientists, introduce them to external scientific networks, give advice on research practices, help secure funding, or provide emotional support.

There is already considerable evidence that high-quality Ph.D. training is a strong determinant of a future academic career rich in scientific and societal contributions (Horta and Santos, 2016; Long et al., 1979; Paglis et al., 2006). We understand Ph.D. training here quite literally as the transmission of scientific knowledge and practices from mentors to mentees (Delamont et al., 1997; Hackett, 1990). This process is sometimes referred to as 'imprinting' (Andrew, 2013), analogous to the bonding reaction of young birds to their first exposure. Arguably, institutionalized factors such as formal supervision and research funding play a crucial role in future career development (Baruffaldi et al., 2016; Cameron and Blackburn, 1981). Less clear, however, is how the local informal environment composed of fellow Ph.D. students affects the development of their careers later on (Broström, 2019).

This paper focuses on the influence of an individual's local academic peer community during the Ph.D. stage on one's future academic career. Specifically, we analyse how these local peer communities in the Ph.D. stage affect the likelihood that an individual will become a mentor of a next generation of Ph.D. students. We understand peer communities as the collective of co-located Ph.D. candidates who are active within a

* Corresponding author.
E-mail address: i.wanzenbock@uu.nl (I. Wanzenböck).

https://doi.org/10.1016/j.respol.2023.104844
Received 30 April 2020; Received in revised form 13 December 2022; Accepted 22 June 2023
Available online 8 July 2023

university at a specific point in time. Our study aims to disentangle some of the complex peer community effects, depending on the nature of the relationship between members of a peer community.

To this aim, we distinguish between *general peers* active in the same discipline and co-located within the same university, and *co-mentees*, that is, members of the local community who share the same formal mentor. Earlier studies found peer community size to be negatively related to the career outcomes of Ph.D. candidates (Broström, 2019; Conti and Visentin, 2015; Heinisch and Buenstorf, 2018). However, we argue that it may not be community size per se, but the presence of peers who share mentors within that community that may be detrimental to academic careers. The reason is that those who share mentors are more likely to compete for the attention of their mentors during the Ph.D, which can induce the negative externalities from peer community size. Co-mentees are also more likely to compete for academic jobs after graduation due to the cognitive similarities developed through their joint Ph.D tractories. However, such competitive effects might be less pronounced or missing in the wider peer community of Ph.D. candidates. A lively community of peers who do not share a mentor might instead be supportive for future careers due to the better opportunities for socialization in academia, learning and support during the Ph.D., and the development of social ties on which academics can draw later again (Katz, 1982).

To analyse peer effects, we use data of 90,264 Ph.D. candidates in the field of mathematics over the period 1945 to 2003. Our dataset allows us to identify a Ph.D. candidate's peer community size based on the Ph.D. graduates of a distinct university at a particular point in time. We further distinguish between community members with whom a Ph.D. does and does not share mentors. Using these data, we test how the local peer community size influences the chances of, first, acquiring an academic position that allows for mentoring next-generation Ph.D. students, and second, maintaining a successful academic position by accumulating mentees over time.

By focusing on the informal research environment of early career scientists, we contribute to the literature on academic careers and academic support structures in three ways. First, we analyse academic success specifically for Ph.D. candidates; a group that is often neglected in empirical studies based on academic networks and publication records (Fortunato et al., 2018). Second, while most studies on career success focus on the formal research environment, characterized by mentorship (Baruffaldi et al., 2016; Liu et al., 2018a), or funds and equipment (Cameron and Blackburn, 1981; Ramesh and Singh, 1998), we consider informal local community effects as determinants of early career success. Third, we contribute to the literature on mentors and peers (Broström, 2019; Conti and Visentin, 2015) by introducing a distinction between peer community members who share mentors and those who do not. By distinguishing competition effects (more likely via co-mentees) and support effects (arising from the wider peer community), we create new insights for the organization of science at universities that help develop inspiring and supportive research environments for early career scientists to facilitate scientific progress.

The paper is structured as follows. In Section 2, we discuss factors of the local peer community of Ph.D. students that influence future academic success and develop our reasoning regarding support and competition among different types of peer community members. Next, in Section 3, we introduce our dataset on mentor-mentee relationships in the field of mathematics, and explain the variables and our empirical modelling strategy. In Section 4, we describe our regression results, and in Section 5, we discuss our findings in light of the literature on academic careers and research environments, and provide pointers for future research.

2. Theoretical background

Research has identified several aspects of the research environment that influence the academic success of Ph.D. candidates. These factors are related to the quality of the university or university department, such as its prestige (Agrawal et al., 2017; Long and McGinnis, 1981) or the access to funds and equipment (Cameron and Blackburn, 1981; Dundar and Lewis, 1998; Kyvik, 1993; Ramesh and Singh, 1998) on the one hand, and the organization of Ph.D. supervision or the quality of Ph.D. mentors (Baruffaldi et al., 2016; Liu et al., 2018b; Waldinger, 2010) on the other hand.

While these formal conditions of research environments are relatively well understood, few have studied the effects of informal conditions, such as the presence of peer communities within university departments, on the future academic success of Ph.D. candidates. This is somewhat surprising as Ph.D. candidates interact, besides their mentors, primarily with members of their co-located peer community (Broström, 2019). Compared to established scholars, Ph.D. candidates have limited professional networks outside of their own university department on which they can rely for collaborations, ideas, feedback and other research opportunities.

We understand peer communities within research environments as social structures which include all Ph.D. candidates that are active within a university in a particular discipline at a certain point in time. Within these social structures, Ph.D. candidates repeatedly interact and learn from each other. For example, members of peer communities may encounter one another in shared offices, during courses, seminars and social events hosted by the department or university, or deliberately spend time together on collaborative work or tutoring arrangements. Peer community members are also subject to the same organizational context and may share mentors, research facilities, equipment or funding. These regular interactions generate social dynamics, which, as we argue, will affect the Ph.D. candidates' knowledge and passion for pursuing an academic career.

Moreover, local peer communities can be considered a social network involving interactions ranging from deep connections among some members to infrequent or casual acquaintances among others. Besides direct effects from interactions, these communities may also generate indirect peer effects, so-called network externalities (Katz and Shapiro, 1985), in that the benefits from social embedding in a local peer community likely depend on the size of that community. Generally speaking, the more members actively interact in the local peer community, the more opportunities for learning and knowledge spillovers (i.e. positive externalities) may arise for an individual member. However, some peers in the local research environment may compete more than others, for instance, for their mentor's attention or position in the community. The resulting ‘congestion effects’ (Ductor, 2015) likely generate negative externalities among these network members. In what follows, we discuss our arguments for the existence of both supportive and competitive influences in the local research environments of young scholars in more detail.

2.1. Support dynamics in the local peer community

Peer communities of Ph.D. candidates can be thought of as supportive environments. The reason is that participating in co-located peer communities resembles a process of socialization involving learning, coaching and teaching among its members and the transmission of tacit knowledge through social interactions (Collins, 1985; Gertler, 2003; Rutten, 2017). Such an academic ‘sharing’ culture, in which frequent interactions in the local peer community is the social norm, is typically more or less developed, depending on the department and discipline in question (Shibayama et al., 2012). Then, if a lively community culture has been established, being part of it allows members of the peer community to repeatedly interact in classes, seminars, shared offices or coffee rooms, and to benefit from knowledge spillovers in their local research environment (Howells, 2002; Olson and Olson, 2000; Polanyi, 1967; Rutten, 2017). Socialization within a local peer community helps Ph.D. candidates to learn faster about the ‘rules of the game’, such as the academic norms and practices, and to develop academic skills and networks, which will enable them to perform better also in later stages of their career (Horta and Santos, 2016).

Related to the social dynamics in local peer communities, we identify two mechanisms of how the interactions in peer communities during the Ph.D. phase can influence the individuals' ability to acquire and maintain a senior position in academia:

First, *coaching and learning* among members of the local peer community may occur when Ph.D. candidates engage in 'peer tutoring' of other members, thus taking on an advisory or motivational role (Ensher et al., 2001; Kram and Isabella, 1985). In this way, local Ph.D. community members can directly support each other during the Ph.D. track. The role of peer tutors may be similar to that of Ph.D. mentors, although more informal and with a stronger focus on emotional and social support than this is the case for formal mentors (Delamont and Atkinson, 2001).

Besides these direct benefits of peer coaching or tutoring, more accidental forms of learning similar to incidences of 'local buzz' (Bathelt et al., 2004) may occur in a peer community through recurring interactions on the work floor. Although potentially more casual, such interactions enable Ph.D. candidates to pick up on the 'hidden curriculum': the knowledge, beliefs, values and practices implicit in both the organizational and academic practice or culture (Gilbert, 2009, p. 56). Resulting benefits can be viewed in terms of knowledge spillovers or positive externalities stemming from the peer community, as peers typically engage in generalized exchange by freely sharing their knowledge and experiences with others. Arguably, both forms of learning during a Ph.D. track - the more direct or intentional and the more accidental ones - can activate the Ph.D. candidates' passion for science and motivate future participation in the academic system (Horta and Santos, 2016; Roach and Sauermann, 2010).

Second, local peer communities allow members to form strong and long-lasting *social networks*, which can be called upon also in later career stages. Members of the local peer community develop their position within the community and form ties with other community members. Peers may form strong ties and share complex knowledge with peers working in the same research group, and weaker ties to share less complex knowledge with peers from other groups (Hansen, 1999). Then, later in the career, similarity in professional positions can further improve interpersonal understanding and likely lead to an intensification of the social ties established in the early career stages (Dahlander and McFarland, 2013; Ruef et al., 2003). For instance, it is known that long-standing ties among individuals provide more trust, certainty and reciprocity than new ties, thus supporting knowledge-sharing and effective communication (Breschi and Lissoni, 2009; Katz, 1982; Uzzi, 1997). While not all connections will persist among alumni (Dahlander and McFarland, 2013), those that do may affect performance in the later stages of an individual's career (Agrawal et al., 2006).

As such, social networks that form early in the career can be a crucial resource for a Ph.D. candidate's future academic success and ability to accumulate own mentees over time. Arguably, such network benefits may be stronger if two individuals are cognitively close to each other (Boschma, 2005), for instance, if they work on similar projects or research topics. Cognitively close peers may also be of greater help with finding new job opportunities in and outside academia, or with establishing future collaborations, research projects or acquiring additional funds later on. However, there might also be decreasing marginal returns from such strong social ties in a peer community, particularly if Ph.D. cohorts are large or the peers in a research group are cognitively too similar. Both may lead to redundancies in the network (Granovetter, 1973), reducing the opportunities to generate future benefits from the local peer community.

Moreover, we argue that, in principle, peers may benefit from any other peer in the local peer community. However, the effects of peer coaching, learning and social networks will depend on the Ph.D. candidate's experience, role and position in the group (Baker and Lattuca, 2010). Novice group members typically have a more peripheral position in a community, having not yet mastered the practices of the community, while experienced members occupy more central positions (Lave and Wenger, 1991). The latter can provide valuable knowledge and guidance to their novice peers. Ph.D. candidates in the early stage would then particularly benefit from the local presence of more experienced peers. In turn, members of the same Ph.D. cohort are assumed to be in similar positions and to interact more frequently within the peer community (Baker and Lattuca, 2010; Lave and Wenger, 1991). Suppose Ph.D. students start their projects in the same period. In that case, they will interact over a longer period, follow introductory and training courses together, and exchange experiences over their full Ph.D. trajectory. Having a similar position in the peer community makes it more likely that they benefit from each other's emotional support. Strong social bonding during the Ph.D. phase will make individuals who conducted their Ph.D. at the same time also more likely to benefit from long-lasting social networks.

2.2. Competition dynamics in the local peer community

Besides supportive dynamics, the local research environment can also be a source of competition for its members (Broström, 2019). Competition effects may more likely occur among peers with similar knowledge bases, competencies and skills, or among peers who interact closely in a mentor's research group. We expect competition among local peer community members to stem from at least two different sources related to two different stages in early careers. First, during the Ph.D. phase, Ph.D. candidates compete for the recognition and attention of mentors, and second, at a later stage of their career, the Ph.D. graduates compete for more senior positions in academia.

Competition for attention during a Ph.D. may arise as one-to-one interaction between mentor and mentee over an extended period of time is necessary for effective knowledge transfer (Delamont et al., 1997). Academic mentors play a role as content advisors, advisors of research practices, emotional supporters, financial supporters and role models (Broström, 2019; Guberman et al., 2006; Paglis et al., 2006). Good mentors thus have a lasting effect on the academic success of their mentees. Apart from transferring knowledge and research skills, mentors may provide their mentees with relevant scientific networks (Ensher et al., 2001), help in acquiring more senior academic positions (Baruffaldi et al., 2016), or pass on their mentoring skills (Malmgren et al., 2010) and their ability to recognize novelty (Trapido, 2015).

It follows that a lack of attention by the mentor is likely to reduce the transfer of knowledge, skills and networks from mentor to mentee (Conti and Visentin, 2015; Dundar and Lewis, 1998). In support of this, an analysis by Long and McGinnis (1985) suggests that the influence of the mentor on the mentee is stronger when mentor and mentee actively co-author papers than when there is no such collaboration. However, the more Ph.D. candidates a mentor is actively supervising, the less time will be available for interaction and knowledge transfer to the individual Ph.D. candidate. In this context, Broström (2019) argued that: "a star scientist already advising a large number of students may be a less useful supervisor than a somewhat less prominent but less overburdened professor" (p. 1655). This suggests that the more 'co-mentees', i.e. mentees attached to the same mentor, the more competition between mentees exists to attract a mentor's attention.

We understand such 'competition for attention' effects as negative externalities or congestion effects that increase with the number of co-mentees in the local peer community. They may emerge mainly among co-mentees and negatively affect them because of the resource and time constraints of a shared mentor (Conti et al., 2014; Shibayama and Baba, 2015). It is reasonable to assume that such competition dynamics during the Ph.D. phase might affect not only the Ph.D. candidates' academic skills or networks but also their aspirations to pursue an academic career, to build up their own research group, and to accumulate many mentees themselves.

While competition for attention is likely to be present between all co-mentees, the extent of the effect may depend on the division of time and

Table 1
Descriptive statistics for mentees within 15 years and peer community size variables.

Variable	N	Mean	St. Dev.	Min	Pctl (25)	Pctl (75)	Max
Mentees within 15 y (DV)	90,264	0.548	1.632	0	0	0	48
Peers	90,264	125.900	142.400	0	29	166	708
Co-mentees	90,264	4.319	5.101	0	1	6	52
Novice peers	90,264	55.320	66.050	0	11	73	327
Novice co-mentees	90,264	1.865	2.610	0	0	3	32
Cohort peers	90,264	15.130	17.450	0	3	20	89
Cohort co-mentees	90,264	0.587	1.001	0	0	1	12
Experienced peers	90,264	55.480	65.400	0	11	73	327
Experienced co-mentees	90,264	1.868	2.571	0	0	3	34

tasks within a mentor's research group. More experienced Ph.D. candidates could be disadvantaged if a mentor's time investment in the younger goes at the expense of support and knowledge transfer for the more experienced. However, more experienced Ph.D. candidates can also take over more responsibilities in larger research groups over time (Feldon et al., 2019; Shibayama et al., 2015). In this informal mentoring role, they will learn to supervise and transfer knowledge effectively early in their careers. Thus, informal mentoring across Ph.D. generations can generate positive career effects for both the more experienced co-mentees (due to early academic supervisory experience) and the younger ones (due to additional knowledge transfer). Consequently, one may expect that the adverse network externalities or congestion effects discussed before are more prevalent among co-mentees of the same cohort, and less so between more novice or experienced peers.

Moreover, *competition for academic jobs* may arise as some Ph.D. peer community members compete for the same academic positions after graduation (Waaijer et al., 2018). Shifts in the balance of labour supply and demand in academia, a bottleneck in the number of tenured positions, and a policy shift away from long-term contracts have all increased uncertainty and added to a shortage of academic career opportunities (Petersen et al., 2012). Hence, while local peer communities can offer support, its members are also competitors for a limited number of academic positions once they enter the job market.

One can further expect that the competition for academic positions is particularly strong among Ph.D. candidates who share mentors compared to those who do not. Being exposed to the same mentor during the Ph.D. track implies higher cognitive proximity due to similar knowledge bases, skills and networks (Nooteboom, 2000), at least to the extent that mentors can transfer these effectively and equally to students. Having a similar specialization thus makes co-mentees also more likely to apply for the same jobs. Following this reasoning, being part of a large research group connected to one mentor would reduce someone's chances of acquiring an academic position later on to the extent that this is in the interest of the individuals. Hence, crowding out due to 'competition for academic jobs' might appear more likely among co-mentees than general peers, and may be strongest among co-mentees in a similar career stage. The recent study of Xing et al. (2022) seems to support this reasoning, showing that Ph.D. candidates trained in big groups face higher risks of academic dropout. This implies that Ph.D. candidates trained in large co-mentee cohorts could also more likely strive for a career outside academia.

3. Methods

3.1. Data

To analyse the determinants of academic success of early career scientists, we use data on mentor-mentee relationships of scientists as retrieved from the Mathematics Genealogy Project run by the NDSU Department of Mathematics.[1] The project employs an online database as a record of individuals who received a Ph.D. in mathematics or related disciplines in the natural sciences, and creates a genealogical network of mathematicians based on mentor-mentee relationships, which is continuously expanded. Individuals can be identified through a unique ID, and are listed together with the names of one or more supervisor(s), university, country, year and, in some cases, the mathematical sub-field in which they obtained their Ph.D. degree. Note that supervisor names concern not only university professors, but also associate professors and people outside academia, as the crowdsourced nature of data collection allows for the inclusion of supervisors other than professors.

For this project, individuals can record themselves and others. Submissions are checked by editors before being released online. Although the data is largely self-registered and therefore prone to biases, the project is recognized as one of the globally most complete and accurate historical records of all Ph.D. degrees within a scientific discipline (Gargiulo et al., 2016; Rossi et al., 2017). The dataset was also used in several quantitative science studies (Malmgren et al., 2010; Rossi et al., 2017; Sugimoto et al., 2011; Wuestman et al., 2020) and has been an inspiration for the collection of similar records in other scientific disciplines (David and Hayden, 2012; Tenn, 2016).

We consider the field of mathematics to be an appropriate context to study informal peer community effects among Ph.D. candidates, since intellectual work in mathematical research is traditionally done individually with very low levels of division of labour in formal teams (Hagstrom, 1964) and little use of assistants and technical support (Hargens, 1975). Moreover, Ph.D. candidates in mathematics are more autonomous and anomic compared to other disciplines, apart from the choice of subject, which is often suggested by the supervisor (Hargens, 1975). This sets mathematics as an academic field apart from other natural sciences where research in formal teams became the norm in the post-war period (Barabási et al., 2002; Hagstrom, 1964; Hargens, 1975). The continuous individualist organization of research in mathematics is also apparent from a study on representative disciplinary journals for the period 1960–2010, which found that the average number of authors in mathematics during this period increased from 1.1 to 1.8, while the average in chemistry increased from 2.4 to 5.1 and in physics from 2.2 to 19 (Huang, 2015). The rarity of team science in mathematics would imply that the informal effects of peers are relatively higher than in other natural sciences, where Ph.D. candidates work more often in pre-structured teams and research infrastructures.

One important development in mathematics, though, has been the rise of more applied fields of mathematics, in particular computer sciences, which rely much more on teams with complementary skills and advanced research infrastructures. These research fields are arguably more 'Mode 2' type of research fields, working more often with industry and on more practical applications alongside theoretical issues (Bonaccorsi, 2008). As our data is partly classified into mathematical sub-field, we are able to compare peer effects for 'pure' mathematics and more applied mathematics.[2]

For this study, we use data on 90,264 Ph.D. candidates who received their degree between 1945 and 2003. We restrict our dataset to those who graduate after 1945, as data from earlier years may be more subject to distortions (Gargiulo et al., 2016). Using this data set, we assume that individuals included in the Mathematics Genealogy Project did a Ph.D. in mathematics or a related subject. We further assume that two

[1] We have obtained the data through recursive queries to the Mathematics Genealogy Project using R and the R-package 'rvest' (Wickham, 2016). Queries ran between the 27th of June 2018 and the 10th of July 2018.

[2] Subjects categorized as 'pure' relate to the study of foundations and logic; discrete mathematics/algebra; analysis; and geometry/topology, according to the Mathematical Sciences Classification System (MSC, 2020).

individuals were part of the same local peer community if both did a Ph.D. in mathematics or a related subject at the same university around the same time. Information on the mathematical subject field is available for 30,656 Ph.D. candidates.

3.2. Measuring academic success

There are various ways to measure the academic success of individual scientists. Many studies focus on academic productivity as measured by one's publishing rate in peer-reviewed journals (Baruffaldi et al., 2016; Dubois and Schlenker, 2014), while academic impact is typically measured by the citations scientists receive to their publications (Mukherjee et al., 2017; Wang et al., 2015). Even though high publication rates and citation counts have been associated with fast career progression and academic status (Long et al., 1993; Park and Gordon, 1996), they are not the only signs of a successful academic career (Milojević et al., 2018). Another way for academics to have a lasting effect on the scientific system is by raising future generations of academics. By mentoring novice academics in field-specific knowledge, scientists can pass on their knowledge and encourage the next generation to build upon and further develop this knowledge (Malmgren et al., 2010).

Hence, we use the number of mentees an individual supervises over time as a key feature of academic productivity and a successful academic career. This indicator also captures part of one's academic impact as mentees typically build on their mentors' knowledge to create new knowledge (Hackett, 1990; Delamont et al., 1997). Moreover, an individual's number of mentees has been found to correlate with other indicators of academic productivity and recognition, including publication rate or being a member of a National Academy of Science (Cruz-Castro and Sanz-Menéndez, 2010; Malmgren et al., 2010).

Note that the number of mentees is a variable that captures two aspects for each individual in our dataset: whether or not they have been able to acquire a position within academia as a mentor viz. a Ph.D. supervisor (regardless of whether this position is at a university or beyond) and, conditional upon becoming a supervisor, the extent to which they have been engaged in the reproduction of their research line over their professional academic career, as reflected by their accumulation of mentees over time. We consider individuals in the zero-mentee group not as unsuccessful individuals per se. Indeed, graduates may leave academia voluntarily to pursue a career elsewhere, possibly, with great success. However, others in the zero-mentee group may have wished to become a supervisor but failed in this ambition. Such individuals left academia involuntarily or remained in academia without a supervising role. Hence, we can understand peer effects on failing to become a supervisor as competition effects only to the extent that graduates with zero mentees aspired to become a supervisor.

We restrict the count of mentees to the first 15 years after receiving their Ph.D. degree for each individual in our dataset. We limit this count to 15 years, because we expect supportive and competitive effects of local peer communities on future academic careers to materialize within a relatively short period of time after the Ph.D. phase. Furthermore, not limiting the time frame of this variable would result in a considerably smaller dataset and create a bias against still-active scientists. In constructing the variable, we use full counting meaning that every mentee counts as one, regardless of whether one or more mentors supervise this mentee. Note here that of the 90,264 individuals in our sample, 88.7 % have only one Ph.D. mentor, 11.1 % have two mentors, 0.2 % three, and only 0.01 % have a maximum of four mentors.

Table 1 presents descriptive statistics for our dependent variable *Mentees within 15 years* as well as for our peer community size variables. The total number of mentees within 15 years varies considerably between 0 and 48, with a mean of 0.55 and a standard deviation of 1.63. Interestingly, a large majority (80 %) of our observations have not mentored any mentees within 15 years. The excess number of scientists with zero mentees mainly reflects that many leave academia shortly after completing their Ph.D.[3] Histograms showing the distribution of our dependent variable and our core independent variable for the local peer community size are provided in Appendix A.

3.3. Measuring local peer community size

We are interested in the effect of a local peer community on the future academic success of Ph.D. candidates. In line with earlier literature (Broström, 2019; Conti and Visentin, 2015), we understand peer communities as a set of individuals who have received their Ph.D. degree from the same university, in the same field, and within a short time from one another.

In contrast to the previous studies, we distinguish between two different types of local peer community members: co-mentee peers and general peers (hereinafter referred to as 'co-mentees' and 'peers'). The peer community size variables are constructed through the following steps. First, for each individual observation, we count all individuals in our dataset who received their Ph.D. degree from the same university, and no more than four years before or after the observed individual.[4] Second, we split and count the resulting group of peer community members into two variables (1) *co-mentees*, consisting of the count of those peer community members who, according to genealogical data from the Mathematics Genealogy Project, share at least one mentor with the observed individual, and (2) *peers*, or the count of the peer community members that do not share a mentor with the observed individual.

Moreover, to explore differences between more experienced and more novice local peer community members, we split our co-mentee and peer variables into three groups: peers who receive their Ph.D. degree in the same year (*cohort members*), peers who received their Ph.D. degree up to four years later (*novice members*), and peers who received their Ph.D. degree up to four years earlier (*experienced members*). This results in six count variables representing peer community size (see Table 1): *novice peers; novice co-mentees; cohort peers; cohort co-mentees; experienced peers, and experienced co-mentees*. Note that we do not intend to understand our peer community size variables to represent separate communities. We only imply that these community subsets may affect an individual's academic success differently through different externalities of learning, support and competition.

Looking at the descriptive statistics in Table 1, we observe that individuals have, on average, 130 peers, that is 126 general peers and 4 co-mentees. Splitting these into the three groups, we see that individuals have, on average, 15 cohort peer members and 55 more novice and 55 more experienced peers. We also see that peer community sizes vary widely: for each count, the standard deviation exceeds the mean.

3.4. Mentor-specific variables

Ph.D. candidates interact primarily with their immediate research environment, consisting of both their mentors and peers. Regarding the mentors, empirical evidence is inconclusive on the effects of mentor seniority. While some studies showed that mentors' performance

[3] If we release the 15-year time constraint, 77 % of our observations have zero mentees, which is close to the 80 % found when setting the 15-year constraint. Further note that 9 % of our observations did not have any mentees graduate within 15 years after their own graduation, but did mentor others in later years.

[4] Given the lack of information of Ph.D. completion times in our dataset, the four-year window is to be considered as a rough estimation of the time it takes to complete a Ph.D. program in mathematics. It is in line with Espenshade and Rodriguez (1997) who report a mean time of 5.7 years to complete a full graduate program in mathematics in the United States in the period 1962–1986. Given that U.S. graduate programs include both a master degree and a doctoral degree, we consider the four-year window for doing the Ph.D. research appropriate.

positively correlates with the performance of their mentees (Malmgren et al., 2010; Trapido, 2015), others found a negative association between a mentor's career stage and the number of graduates supervised before (Heinisch and Buenstorf, 2018). A positive effect may reflect that mentees with experienced mentors learn more valuable knowledge and skills, contributing to their success later in their careers. Furthermore, the more mentees a mentor has supervised before, the more effective a mentor may become in supervising subsequent mentees. Nonetheless, senior mentors are more likely to take over managerial or representative tasks in the organization and academic community, which decreases the attention and time they can spend on actual research or supervision. Also, they might more likely be 'locked-in' in established research lines, thus less open to the latest theory and method developments, potentially negatively affecting their students' career prospects (Heinisch and Buenstorf, 2018).

It is important to note that the matching between mentees and mentors is not a random process. While mentees may not deliberately select their peer community, they are, at least to some extent, able to select their mentors and university environment. For example, proficient or high-potential candidates are more likely able to choose prestigious universities, or to select reputed scientists as mentors, and vice versa. Consequently, mentees of more proficient mentors may also have a more proficient co-mentee community than mentees of less proficient mentors. This means that mentor-mentee observations might not be entirely independent (see e.g. Azoulay et al., 2017).

To better disentangle the effect of peer communities from mentorship effects, we include two mentor-specific control variables, one for *mentor experience* and one for *mentor prestige*. For mentor experience, we cannot simply count the number of prior mentees of an individual's mentor, as this count would overlap with our peer community size variable. We, therefore, consider only prior mentees, who received their Ph.D. degree at least four years before the observed individual and are thus not included in our variable 'more experienced co-mentees'. Note that prior mentees included in mentor experience may have obtained their Ph.D. from different universities than the observed individual. Observations of the mentor experience variable vary widely (mean = 4.697, standard deviation (sd) = 7.799), with one mentor having mentored as many as 118 individuals before our observed individual.[5]

We further include a variable for high mentor prestige by counting the number of prizes awarded to a mentor. To this end, we collected data on the prize winners of all major awards in mathematics (including computer science) using the list of most prestigious international academic awards compiled by Meho (2020). This list includes the Abel Prize, Crafoord Prize, Fields Medal, Maryam Mirzakhani Prize (formerly NAS Award), Rolf Schock Prize, Shaw Prize, Turing Award, and Wolf Prize. In total, 211 mentors received at least one of the prizes. The variable we constructed counts the number of awards per mentor, ranging from 0 to 5 as its maximum.

3.5. Control variables

We include variables for individual-level attributes, university-level characteristics and control for country and time heterogeneity when estimating local peer community effects on academic success. At the level of the individuals, we control for the gender of the Ph.D. candidate. To construct the gender control variable, we matched the name and country of each Ph.D. candidate in our dataset with data from the worldwide gender-name dictionary developed by Raffo (2016) and Martínez et al. (2016). In this way, we can identify the gender (male/female) for 82 % of all individuals in our sample.

Moreover, we control for potential differences in academic careers between pure mathematics and applied fields, as we assume that graduates in applied fields of mathematics are more likely to pursue a career outside academia. We construct a dummy variable indicating whether a Ph.D. candidate graduated in a mathematical sub-field related to pure mathematics using two-digit subject codes as in the Mathematical Sciences Classification System (MSC, 2020). Our database includes field-specific data for about 44 % of Ph.D. graduates.[6]

In addition, we control for variations at the university level related to prestige, size and university type. For university prestige in mathematics and related fields, we include two dummy variables, one indicating whether a Ph.D. candidate graduated from a university with a winner of the *Fields Medal* in the past, and one indicating whether a Ph.D. candidate graduated from a university with a *Nobel Prize* winner in physics or chemistry in the past. For both variables, we used the affiliation of the award winners and the year the university received the award for the first time. Moreover, we construct a dummy variable for the *Top 1 % universities* in terms of the accumulated number of graduates in our database to control for prominence and reputation effects of the Ph.D. programme. This dummy indicates whether an individual graduated from one of the 15 most prestigious universities in our data. Twelve of the top 1 % universities are located in the US, one university in the UK (Cambridge), one university in Switzerland (ETH) and one in Russia (Moscow State University). We assume that highly prestigious universities with established Ph.D. programmes most likely attract proficient early career scientists (Burris, 2004; Clauset et al., 2015). Furthermore, to distinguish between universities that are more of an applied nature and universities more oriented towards scientific research, we include a dummy variable if a university is a medical or technical university following the classification by Frenken et al. (2017). We assume that more Ph.D. candidates would choose to pursue a career outside academia in such more applied universities.

To further consider country-specific differences in Ph.D. training systems, we include fixed effects for the *country* in which the Ph.D. graduated. To construct these dummy variables, we select only the countries in the upper quartile according to the number of graduates in the sample (24 countries in total). Other countries are included in a residual category 'other'. This restriction was necessary to achieve sufficient variation for our model computation over countries and time, given the skewed distribution of candidates over countries. Finally, we include fixed effects for the *decade* of an individual's Ph.D. graduation to control for time-specific differences. Descriptive statistics of all control variables are provided in Appendix A.

3.6. Empirical model specification

Our dependent variable *mentees within 15 years* is a count variable with a large number of zero observations, a low mean and a high standard deviation (see Table 1). The large number of zeros is not surprising, as our dataset includes individuals who obtain a Ph.D. degree but do not continue their academic career (structural zeros). We therefore assume that our outcome variable reflects two different processes: First, the Ph.D. graduates' susceptibility to acquiring a mentoring position at a

[5] We run a robustness check with an alternative measure of mentor experience to avoid collinearity problems between the positively correlated mentor experience and the more experienced co-mentees variable (cor = 0.464). For this alternative measure, we replace our variable mentor experience with a count of the mentor's prior mentees, but excluding all prior mentees who obtained their degree at the same university as the observed individual. This alternative measure correlates with our original mentor experience variable (cor = 0.511), but much less with more experienced co-mentees (0.097). Note that this alternative mentor experience count is biased against individuals with mentors who have not been affiliated to other universities. Nevertheless, main results of our regression models do not change, and are available upon request.

[6] We run robustness checks for our model using fixed effects for the 63 different mathematics fields included in our database. Effect size and significance of our peer community variables do not change, and results are available upon request.

university (within 15 years) which allows them to mentor a new generation of Ph.D. candidates, and second, the ability of these individuals to accumulate more mentees over time.

To account for the excess zeros and overdispersion in our data, we estimate zero-inflated negative binomial (ZINB) regression models. A ZINB consists of two parts: a logistic component to predict the probability of having zero mentees, and a negative binomial component to model the count outcome, that is, the number of mentees above zero (Cameron and Trivedi, 1998). A negative sign in the logit part, for instance, would indicate that a respective variable decreases the likelihood of having 0 mentees within 15 years, or put differently, a higher chance of pursuing an academic career involving mentorship. Then, a positive sign in the count part would indicate a higher expected number of mentees within 15 years.

To analyse the relationship between early career scientists' peer community and their academic success, we estimate eleven regression models in total. Each model includes the mentor-specific variables, different control variables and the two peer community size variables, *co-mentees* and *peers*. First, we run a general peer community model on the full sample of 90,264 Ph.D. candidates, controlling for mentor performance, university prestige and university type. Second, we want to distinguish the peer community by experience and explore differences in Ph.D. candidates' positions within their peer communities. To this aim, we specify one model for the number of *more novice peer community members (co-mentees* and *general peers)*, one for *cohort members*, and one for *more experienced peer community members.*[7] Lastly, we use a smaller set of observations to test whether our general results for peer community effects hold when controlling for gender-specific or mathematics field-specific heterogeneities. We additionally include interaction terms between the mathematics dummy variable and the two peer community variables. In this way, we test for varying peer community effects in different branches of mathematics. All our models include fixed effects for country and decade.

Next to the main models discussed in the next section, we run a series of robustness checks for which we provide the results in Appendix B. First, we test for potential influences of the country and university a Ph.D. student was hired after graduation. The first work environment after graduation might influence the career of young scholars even more than their Ph.D. training university (Way et al., 2019). To do so, we considered the size of the university an individual was active in the longest within the first 15 years of their career, as early career scholars might be more productive in larger, more prosperous universities (Way et al., 2019). We construct a categorical variable including five size groups according to the number of Ph.D. graduates in the sample. Note that we only have data for individuals who become mentees themselves, and therefore run this check only for the count part of our models. As a second robustness check, we calculate clustered standard errors at the university level for our models to test whether a potential sampling bias at the university level – for instance, due to the crowd-sourced nature of the data – influences the significance of the observed peer community effects. Finally, regarding our dependent variable, we run the general model with an alternative specification measuring the number of mentees within 10 years, and the number of mentees within 20 years, to test for different definitions of academic success of early career scientists.[8]

[7] We estimate separate models because of the high correlation between variables (see Appendix A) which would cause multicollinearity problems if they were included in the same model. We considered VIF-values for each of our models; none of these are problematic according to standard rules (Hair et al., 2010).

[8] We also tested for a number of non-linearities in the relationship between an individual's academic career outcome and our continuous independent variables (peers, co-mentees and mentor experience). The obtained significances for the quadratic terms in the different models were not conclusive to derive strong interpretations. The results for our linear term estimates do not change for all three models, and are available upon request.

4. Results

4.1. General models

Table 2 shows the coefficients of our ZINB regression models (models 1–4) performed on the full sample. Model 1 includes only the mentor-specific variables, while all other model variants include the local peer community variable. In this way, we can better understand the effects of the mentor versus the effects of peers. The logit parts of the models in Table 2 suggest that *mentor experience* is negatively associated with academic success: it suggests that Ph.D. candidates with a more experienced mentor are less likely to supervise Ph.D.'s themselves within 15 years than Ph.D. candidates with a less experienced mentor. In contrast, the positive sign of the *mentor prestige* variable suggests that Ph.D. candidates supervised by prestigious mentors are more likely to become mentors themselves. These results might indicate that Ph.D. candidates with very successful mentors are also more likely to pursue an academic career, which is not necessarily the case for Ph.D. candidates supervised by experienced mentors. Ph.D. candidates with experienced mentors might also seek good career opportunities outside academia; they may leave academia and not supervise any Ph.D. candidates within 15 years. In contrast, we find positive coefficients for both variables, mentor experience and mentor prestige, in the count part of the model, implying that mentor performance is positively associated with a high number of mentees for those staying in academia. Further note that the coefficients of both mentor variables are stable in the logit part over all four models. However, once we include our local peer community variables (model 2–4), we do not find significant effects of mentor-specific variables on the number of mentees in the count part.

For the peer community variables, we find a negative sign for *peers* in the logit part, indicating that a one-unit increase in the number of general peers decreases the odds of having 0 mentees within 15 years between 0.1 and 0.2 % (OR $= e^{-0.001} = 0.999$, and $e^{-0.002} = 0.998$, respectively). In the count part of our models, we find a positive sign: a one-unit increase in the number of general peers increases the expected number of mentees within 15 years between 0.03 % and 0.1 % ($e^{0.0003} = 1.0003$, and $e^{0.001} = 1.001$, respectively). Even though both effects are small, our results indicate that with an average number of 126 (sd = 142), general peers can make a considerable difference in the future academic career of Ph.D. candidates. The negative effect in the logit part suggests support in the form of learning, emotional support and socialization in academia among local peers, positively contributing to staying in academia and acquiring a mentor role later on. The supportive effect observed in the count part of our model confirms the positive association between the size of the local peer community and academic success, conditional on acquiring a mentoring position.

In contrast to general peers, we find positive and significant effects for *co-mentees* in both parts of models 2–4. Also, these effects are larger than the effects of *peers*. A one-unit increase in co-mentees increases the odds of having 0 mentees within 15 years of about 3 % (OR $= e^{0.033} = 1.034$), which may point to a competition effect among co-mentees. To the extent that graduates wish to take on a mentoring role later on, co-mentees may more likely compete for such positions due to their cognitive similarity. Alternatively, it seems that Ph.D. candidates with many co-mentees are more likely to pursue a career outside academia. For those acquiring a mentor position, we find an increase in the number of mentees of about 1.3 % ($e^{0.013} = 1.013$) in the count part. This positive effect would point to support, rather than competition, among co-mentees in accumulating a higher number of mentees over time, indicating the relevance of long-lasting social ties in academia established during the Ph.D. phase.[9]

[9] We also tested a model including only peer community variables, but not the mentor-specific variables. Coefficient size and significance of peers and co-mentees do not change. Results are available upon request.

Table 2
ZINB regression model coefficients, general model (dependent variable: mentees within 15 years).

	Logit - Zero mentees				Count - Number of mentees			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Peers		−0.002***	-0.002***	−0.001***		0.001***	0.0005***	0.0003**
		(0.0002)	(0.0002)	(0.0002)		(0.0001)	(0.0001)	(0.0001)
Co-mentees		0.034***	0.033***	0.033***		0.013***	0.013***	0.012***
		(0.004)	(0.004)	(0.004)		(0.003)	(0.003)	(0.003)
Mentor experience	0.011***	0.007***	0.008***	0.011***	0.004**	−0.0005	−0.0003	−0.0003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Mentor prestige	−1.613***	−1.326***	−1.218***	−1.170***	0.067**	0.044	0.045	0.048
	(0.187)	(0.178)	(0.159)	(0.161)	(0.029)	(0.030)	(0.030)	(0.031)
Top 1 % university (dummy)			−0.355***				0.042	
			(0.059)				(0.039)	
Nobel prize (dummy)				−0.507***				0.053*
				(0.049)				(0.031)
Fields medal (dummy)				−0.535***				0.054
				(0.071)				(0.041)
Technical/Medical (dummy)				−0.024				0.069*
				(0.054)				(0.036)
Constant	−2.294**	−2.487***	−2.252***	−2.342***	−0.073	−0.064	−0.066	−0.109
	(0.295)	(0.274)	(0.264)	(0.274)	(0.105)	(0.105)	(0.106)	(0.106)
Observations	90,264	90,264	90,264	89,640	90,264	90,264	90,264	89,640
Log Likelihood	−74,967	−74,599	−74,808	−74,242	−74,967	−74,599	−74,808	−74,242

Note: Coefficients of a zero-inflated negative binomial model. All models include control variables for country and decade (dummy). Standard errors in brackets. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Regarding our university-level controls, we observe significant negative effects in the logit part for top 1 % Ph.D. graduation programmes (model 3) and for prize-winning universities (model 4), confirming that graduates from established or prestigious universities are more likely to pursue an academic career which includes mentoring the next generation. However, we do not find conclusive evidence of graduation from prestigious universities on the number of future mentees in model 3 and 4. While graduation in mathematics from a technical or medical university does not significantly influence the chances of having 0 mentees, we find a positive albeit less significant influence on the number of own mentees within 15 years.

4.2. *Peer community experience*

Table 3 shows the results of the regressions in which we split the peer community variables into three different groups according to the experience level of peers. Here we go one step further in analysing local peer community effects and explore whether different peer groups, or cohorts, exert different effects on an individual's mentoring probability. First, regarding *peers,* we find that the number of same *cohort peers* who do not share a mentor (model 6) negatively and significantly affects the likelihood of having 0 mentees within 15 years ($e^{-0.009} = 0.991$). This supportive effect is weaker for more *novice peers* (model 5) and more *experienced peers* (model 5). Accordingly, the peer community experience models confirm the support effect among local peers in pursuing an academic career as observed in our general models in Table 2. This general support effect seems to be driven particularly by peers who graduate in the same year (model 6). We also find in the count part of our experience models that the positive effect of *peers* may be stronger associated with *cohort peers* than with more *novice peers* or more *experienced peers*. Both model results are in line with the idea that same-cohort peers establish stronger support and social ties in their peer community than Ph.D. candidates from different cohorts.

Moreover, for *co-mentees*, we again find a positive effect of the number of co-mentees on the likelihood of having 0 mentees within 15 years, regardless of experience. While model 2 showed us the average effect over all co-mentees, we now find that this competition effect is strongest for *cohort co-mentees*, and weakest for more *novice co-mentees*. We can interpret this finding as an indication of considerable competition for acquiring mentoring positions among co-mentees, particularly among co-mentees from the same cohort. Regarded the expected number of mentees (count part), the positive effect of *co-mentees,* as shown in model 2, seems to be mostly due to more *novice co-mentees*, whereas more *experienced co-mentees* and *cohort co-mentees* do not have any significant effect on the expected number of mentees. We interpret this finding as an indication of support effects due to the experience gained by informally mentoring younger cohorts during the Ph.D. phase. Instead, we do not find evidence for the idea of persisting co-mentee networks that support future academic success.

Based on these results, we can conclude that it may be mainly *cohort peers* who provide support to their peers, for instance, due to social bonding occurring in introductory courses and over a Ph.D. trajectory. *Co-mentees,* particularly from the same or earlier cohorts, could instead be a source of competition, due to either direct competition for academic jobs (same cohort peers), or indirectly, as earlier cohort co-mentees might occupy academic positions that are also desirable for later peers. Remarkable is that all co-mentees seem to exert competitive (negative) effects to become a mentor, while peers are more likely to support and increase the likelihood of acquiring a mentorship role. Both competition and support effects are strongest for the same or earlier cohorts.

4.3. *Analysis of individual- and field-specific characteristics*

We perform further analyses on how gender (model 8) and field-specific characteristics (model 9–11) influence the academic success of early career scientists. Due to failed name-gender matching for individuals, or missing data on mathematic sub-field in our database, the number of observations is about 1/3 lower in model 8, and about 2/3 lower in models 9, 10 and 11. All results are shown in Table 4.

Despite the reduced sample, we do not see any considerable differences in effect size and significance of our local peer community variables compared to the baseline models in Table 2 or the experience models in Table 3. Looking at the categorical control variables in more detail, model 8 shows that female graduates, and those with unknown gender, show significantly higher chances of 0 mentees compared to male graduates in the logit part, and also a significantly lower expected number of mentees in the count part. These results confirm the often observed gender disparities in science (Huang et al., 2020; Larivière et al., 2013), which seem to be reproduced over different generations of Ph.D. candidates.

We also distinguish between Ph.D. degrees in core areas of mathematics and Ph.D. degrees in more applied sub-fields such as computer

Table 3
ZINB regression model coefficients: experience (dependent variable: mentees within 15 years).

	Logit - Zero mentees			Count - Number of mentees		
	(5)	(6)	(7)	(5)	(6)	(7)
Novice peers	−0.003***			0.001***		
	(0.0004)			(0.0003)		
Novice co-mentees	0.036***			0.028***		
	(0.007)			(0.005)		
Cohort peers		−0.009***			0.003***	
		(0.002)			(0.001)	
Cohort co-mentees		0.110***			0.013	
		(0.017)			(0.013)	
Experienced peers			−0.003***			0.001***
			(0.0004)			(0.0003)
Experienced co-mentees			0.065***			0.008
			(0.007)			(0.006)
Mentor experience	0.013***	0.013***	0.008***	0.001	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Mentor prestige	−1.249***	−1.232***	−1.229***	0.047	0.057*	0.051*
	(0.160)	(0.158)	(0.157)	(0.030)	(0.030)	(0.030)
Top 1 % university (dummy)	−0.425***	−0.454***	−0.425***	0.053	0.078**	0.066*
	(0.055)	(0.058)	(0.059)	(0.037)	(0.038)	(0.038)
Constant	−2.213***	−2.267***	−2.229***	−0.103	−0.084	−0.055
	(0.267)	(0.275)	(0.272)	(0.106)	(0.106)	(0.107)
Observations	90,264	90,264	90,264	90,264	90,264	90,264
Log Likelihood	−74,592	−74,613	−74,582	−74,592	−74,613	−74,582

Note: Coefficients of a zero-inflated negative binomial model. All models include control variables for country and decade (dummy). Standard errors in brackets. *p < 0.1; **p < 0.05; ***p < 0.01.

science, mathematical physics or economics. Our mathematics dummy in model 9 shows that graduates in mathematics show a significantly lower expected number of future mentees (count part), but no significant effect in the logit part. Our results might point to similar entry barriers to pursuing an academic career but lower chances of accumulating mentees in 'pure mathematics' compared to more applied areas of mathematics. This finding could result from generally smaller group sizes in mathematical research compared to more applied natural sciences (Huang, 2015).

To further test for differences in the influence of local peer communities in different branches, we interact our peer community variables with the mathematics dummy (model 10 and 11). In the logit part, we observe that the positive effect of a larger peer community is stronger for mathematics than more applied fields. This result seems to confirm our assumption that Ph.D. candidates in mathematics rely on informal peer interaction more than in other fields in which team science is more common. In the count part, we see a slightly higher support effect of co-mentees in mathematics than in more applied fields, while there is no

Table 4
ZINB regression model coefficients: Gender- and field-specific characteristics (dependent variable: mentees within 15 years).

	Logit				Count			
	(8)	(9)	(10)	(11)	(8)	(9)	(10)	(11)
Peers	−0.002***	−0.001***	−0.001***	−0.001***	0.0004**	0.001***	0.001***	0.001***
	(0.0002)	(0.0003)	(0.0004)	(0.0003)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Co-mentees	0.028***	0.051***	0.050***	0.063***	0.017***	0.014***	0.013***	0.009
	(0.004)	(0.006)	(0.006)	(0.007)	(0.003)	(0.005)	(0.004)	(0.005)
Mentor experience	0.006**	0.012***	0.012***	0.012***	−0.003	0.001	0.001	0.001
	(0.003)	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)
Mentor prestige	−1.132***	−0.795***	−0.810***	−0.787***	0.040	0.112*	0.119*	0.105*
	(0.159)	(0.231)	(0.240)	(0.231)	(0.033)	(0.063)	(0.062)	(0.063)
Top 1 % university (dummy)	−0.364***	−0.291***	−0.283***	−0.279***	0.042	−0.117*	−0.116*	−0.118*
	(0.064)	(0.106)	(0.109)	(0.105)	(0.043)	(0.066)	(0.066)	(0.066)
Gender (unknown)	0.393***				0.138			
	(0.129)				(0.102)			
Gender Female	0.416***				−0.372***			
	(0.062)				(0.048)			
Math (dummy)		0.028	0.217**	0.217***		−0.275***	−0.220***	−0.306***
		(0.070)	(0.093)	(0.082)		(0.040)	(0.050)	(0.049)
Peers x math			−0.002***				−0.0005	
			(0.001)				(0.0003)	
Co-mentees x math				−0.043***				0.014*
				(0.011)				(0.008)
Constant	−2.260***	−2.289***	−2.571***	−2.284***	−0.051	0.168	0.103	0.196
	(0.329)	(0.582)	(0.627)	(0.546)	(0.141)	(0.174)	(0.176)	(0.174)
Observations	65,712	30,650	30,650	30,650	65,712	30,650	30,650	30,650
Log Likelihood	−55,548	−28,758	−28,754	−28,742	−55,548	−28,758	−28,754	−28,742

Note: Coefficients of a zero-inflated negative binomial model. All models include control variables for country and decade (dummy). Standard errors in brackets. *p < 0.1; **p < 0.05; ***p < 0.01.

significant difference between fields for peers.

4.4. Robustness

We checked the robustness of the dependent variable (Table A2.1 in Appendix B). We find that counting the number of mentees acquired within 10 or 20 rather than 15 years has little effect on our model results. Our main findings have remained unchanged. The robustness check suggests that the effects of competition and support among peer communities of Ph.D. students may also materialize within 10 years, and can last until at least 20 years after receiving the Ph.D. degree.

Moreover, we controlled for potential influences of the university and country an individual is hired, besides controlling for the individual's university of training. In particular, department size is assumed to be positively correlated with our academic success measures, due to favourable opportunities to accumulate mentees. Note that this robustness check can only be performed for individuals who become a mentor of at least one Ph.D. candidate, referring to the count part of our model. The results show that the support effect arising from local peers and co-mentees are robust, also if we control for the size and country of the hiring university (Table A2.2 in Appendix B). We find significant and with the university size increasing effects of the hiring environment on the expected number of mentees of early career scholars in their first 15 years.

Finally, we report in Table A2.3 in Appendix B the results of our general models using clustered standard errors to control for sampling biases at the university level. We observe a slightly decreasing significance of the general peer and mentor experience variables. Not surprisingly, the same is true for our university control variables. However, overall the results of the different model variants are robust, confirming the supportive and competitive effects found in our main analyses.

5. Discussion

In this paper, we looked at how the local academic peer community can be associated with the future academic success of Ph.D. candidates. We argued that peer community effects are likely ambiguous as peers may provide support through learning and networking but also generate competition among their members during and after the Ph.D. phase. To separate the effects of local academic peer communities, we distinguished between local peers who share mentors (co-mentees), thus competing for their mentor's attention and similar jobs in academia, and other local peers who we considered to be primarily a source of support. We further consider the various support and competition effects as a function of the local peer community size. The results of our regression analysis, in which we distinguish between co-mentees and general local peers, support these ideas.

We found evidence that a high number of co-mentees in the local peer community reduces the chances of acquiring an academic mentoring position later on, particularly when these co-mentees are part of the same cohort. We argued that co-mentees might, indeed, be the strongest competitors for the time of their mentors during the Ph.D. phase (Broström, 2019; Long and McGinnis, 1985). To the extent that Ph.D. graduates wish to take on an academic career, co-mentees are also the most likely to compete for academic jobs later on (Waaijer et al., 2018). However, the finding of such crowding-out effects in academia does not rule out the possibility that Ph.D. candidates trained in larger cohorts have higher chances of finding attractive job opportunities outside academia, for instance, facilitated by the social connections made during their Ph.D. phase.

Moreover, we did not find evidence for competition effects in one's later academic career observed in terms of the number of mentees supervised within 15 years. If anything, the local presence of co-mentees during the Ph.D. stage may positively affect academic success in the later mentoring stage. Overall, our results suggest that Ph.D. candidates trained in larger groups may become more successful in, and potentially also outside academia, because of their cognitive similarities, their joint embedding in similar professional networks and their mutual connection to a proficient mentor.

Our analysis also pointed to supportive effects among peers who do not share mentors. The finding that peers, particularly those from the same cohort, are significantly associated with career success point to the importance of long-lasting acquaintances in science and across different career stages. Also, more novice and more experienced peers with whom one does not share mentors are positively associated with academic success. This is in line with the idea that a dynamic training environment of young scholars with opportunities for learning, networking or tutoring among the local peer community provides support during the Ph.D. track, and may influence one's decision to pursue an academic career and to further succeed in that career.

All things considered, our research suggests that besides features of the mentor and the university (Baruffaldi et al., 2016; Cameron and Blackburn, 1981; Shibayama, 2019), the local peer community of Ph.D. candidates within universities can significantly affect long-term academic success due to higher learning and networking opportunities. This is relevant, because whereas aspiring Ph.D. candidates can, at least to some extent, select their mentors and their affiliations, their influence on the selection of their peers is at most limited. The organization of supportive research environments, in particular linking same cohort Ph.D.'s but also more experienced with novice candidates, can be taken up as a responsibility of the university and should be considered in developing Ph.D. training programmes (Dericks et al., 2019).

Our results indicate that it is not community size per se but competition between co-mentees (peers sharing a mentor) that may account for some of the adverse effects of peer community size found in earlier research (Broström, 2019; Conti and Visentin, 2015). Hence, large cohorts may not be as harmful as long as there is considerable diversity among cohort members by distributing mentoring tasks among various mentors.[10]

For the literature on academic success, scientific teams and collaborations, our distinction between co-mentees and general peers illustrate the potential for more complex analyses of mentorship relationships in academia. Our study could be extended by including not only direct mentor-mentee relationships but a more complex measure of ‘distance’ between individual peers based on their entire ‘mentorship pedigree’, as common in genealogical approaches (Rossi et al., 2017; Sugimoto et al., 2011). Similarly, it would be interesting to analyse how direct and indirect mentorship relations affect the performance and composition of academic teams (Singh and Fleming, 2010), or determine job mobility inside and outside academia (Cruz-Castro and Sanz-Menéndez, 2010).

Our work has some limitations. Measuring academic success with the number of Ph.D. candidates supervised provides a partial understanding of career success. Despite some evidence that the number of mentees correlates with other indices of academic quality (Malmgren et al., 2010), it is not unlikely that some high-performing academics do not have a large number of mentees, and vice versa. Given that mentoring relations and local peer communities are social in nature, more so than scientific publications and citations, our data focuses primarily on social dynamics in science. Comparing our results with results based on other performance indicators would thus be insightful. Publication and citation data can be helpful to better qualify the relationship between successful mentors and successful mentees. In particular, the number of co-publication and citations between mentor and mentees can indicate the time invested by a mentor in a mentee's project relative to other mentees.

Moreover, our analysis focuses on the case of mathematics and related fields in natural science. This choice may limit the generalizability of our findings to the extent that the role and size of peer

[10] Some shared mentorship is likely not harmful, though. Recent research (Shibayama, 2019) has shown that overly involved supervision can harm the mentee's autonomy and future career success, too.

communities of Ph.D. candidates vary between disciplines. For example, as academic job scarcity may vary, so may job competition (Waaijer et al., 2018). Despite these limitations, we believe that the more general mechanisms suggested here, related to support and competition in local Ph.D. communities of general peers and co-mentees, may apply to all scientific disciplines. A recent study by Xing et al. (2022) using large-scale genealogical data on mentor-mentee relationships in Neuroscience, Chemistry and Physics seems to support this claim. Their findings suggest that mentees trained in large groups of co-mentees show higher academic dropout rates but higher academic performance if they survive. However, more in-depth research would be needed to shed light on specific mechanisms in large peer communities and research groups, for example, to disentangle academic job competition from mentor's time and resource competition; or knowledge transfer, tutoring or emotional support during the Ph.D. phase from the more casual social acquaintances among peers. In this regard, it would also be interesting to study country-specific peer support and competition effects in more detail, for instance, related to different Ph.D. training systems or other institutional differences.

CRediT authorship contribution statement

Mignon Wuestman: Conceptualization, Methodology, Software, Investigation, Formal analysis, Data curation, Writing – original draft. **Iris Wanzenböck:** Conceptualization, Methodology, Investigation, Formal analysis, Validation, Writing – original draft, Writing – review & editing, Supervision. **Koen Frenken:** Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing 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.

Data availability

Data will be made available on request.

Acknowledgements

Mignon Wuestman and Koen Frenken received funding from the Netherlands Organization for Scientific Research (NWO) within the Vici program (number 453-14-014). We thank Jarno Hoekman and the participants of the SEI doctoral consortium 2019 at KU Leuven for useful comments.

Appendix A. Descriptive statistics

Correlation matrix of all continuous variables.

	M	G	C	NG	NC	CG	CC	EG	EC	ME	MP
Mentees within 15 y (M)	1										
Peers (G)	0.045	1									
Co-mentees (C)	0.019	0.182	1								
Novice peers (NG)	0.056	0.951	0.195	1							
Novice co-mentees (NC)	0.041	0.150	0.857	0.197	1						
Cohort peers (CG)	0.035	0.959	0.163	0.861	0.115	1					
Cohort co-mentees (CC)	0.002	0.107	0.653	0.106	0.432	0.106	1				
Experienced peers (EG)	0.032	0.961	0.156	0.83	0.098	0.952	0.098	1			
Experienced co-mentees (EC)	−0.004	0.167	0.86	0.146	0.516	0.164	0.469	0.172	1		
Mentor experience (ME)	−0.008	0.161	0.42	0.135	0.262	0.160	0.266	0.171	0.464	1	
Mentor prestige (MP)	0.057	0.116	0.070	0.109	0.060	0.109	0.045	0.114	0.060	0.094	1

Descriptive statistics of continuous and binary control variables.

Variable	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Mentor experience	90,264	4.697	7.799	0	0	6	118
Mentor prestige	90,264	0.038	0.257	0	0	0	5
Top 1 % University (dummy)	90,264	0.232	0.422	0	0	0	1
Nobel prize (dummy)	90,264	0.245	0.430	0	0	0	1
Fields medal (dummy)	90,264	0.132	0.338	0	0	0	1
Technical/medical (dummy)	89,640	0.116	0.320	0	0	0	1
Math (dummy)	30,650	0.580	0.494	0	0	1	1

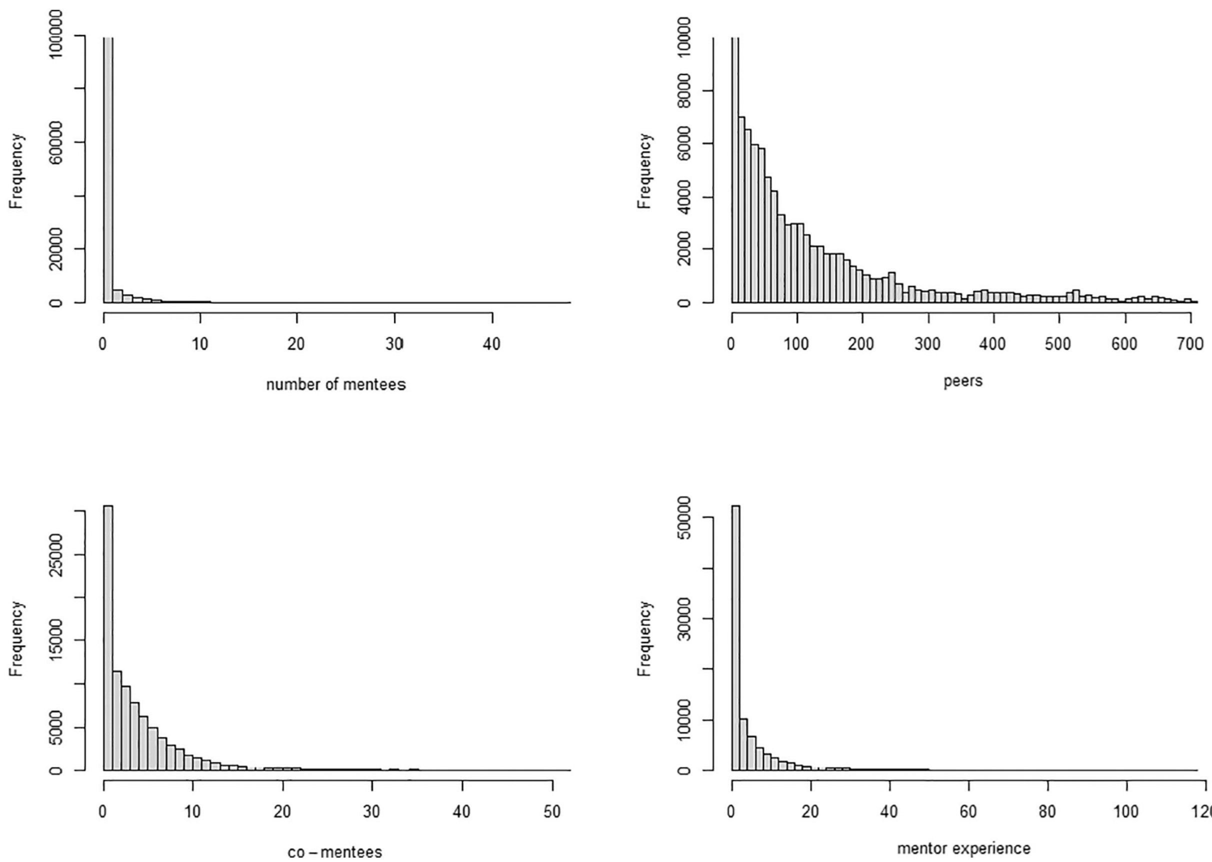
Frequency distribution of categorical variables: Ph.D. graduates over countries, decades and gender.

Category	Frequency	Category	Frequency	Category	Frequency
Country		*Country*		*Decade*	
US	57,284	SE	859	1940s	1,287
DE	13,864	MR	773	1950s	4,942
GB	5,902	BR	761	1960s	13,921
CA	4,458	AT	760	1970s	19,530
NL	2,887	UA	562	1980s	22,506
FR	2,409	IN	526	1990s	37,565
CH	1,825	IE	498	2000s	11,430
ES	1,727	FI	476	*Gender*	

(continued on next page)

(*continued*)

Category	Frequency	Category	Frequency	Category	Frequency
IT	1,262	RS	411	Male	62,307
AU	1,247	RO	399	Female	11,725
RU	949	other	3,252	Unknown	1,812
PL	919				



Distribution of core variables. Number of mentees after 15 years (DV, top left); number of peers (top right); number of co-mentees (bottom left); mentor experience (bottom right).

Appendix B. Robustness checks

Table A2.1
ZINB regression model coefficients for model 3. DV: Mentees within 10 years and 20 years.

	Logit - zero mentees		Count - number of mentees	
	(10y)	(20y)	(10y)	(20y)
Peers	−0.001***	−0.002***	0.0002**	0.0005***
	(0.0001)	(0.0002)	(0.0001)	(0.0001)
Co-mentees	0.008***	0.033***	0.008***	0.013***
	(0.002)	(0.004)	(0.002)	(0.003)
Mentor experience	0.009***	0.008***	0.001	−0.0003
	(0.002)	(0.002)	(0.001)	(0.002)
Mentor prize	−0.349***	−1.218***	−0.004	0.045
	(0.032)	(0.159)	(0.022)	(0.030)
Top 1 % university (dummy)	−0.259***	−0.355***	0.019	0.042
	(0.039)	(0.059)	(0.026)	(0.039)
Constant	1.274***	−2.252***	0.494***	−0.056
	(0.105)	(0.264)	(0.079)	(0.106)
Observations	90,264	90,264	90,264	90,264
Log Likelihood	−30,239	−30,239	−17,040	−74,570

Note: Coefficients of a zero-inflated negative binomial model. All models include control variables for country and decade (dummy). Standard errors in brackets. *p < 0.1; **p < 0.05; ***p < 0.01.

Table A2.2
NegBin regression model coefficients. Hired university controls.

	Count – Number of mentees				
	(2)	(3)	(4)	(8)	(9)
Peers	0.0002***	0.0003***	0.0001	0.0002***	0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Co-mentees	0.011***	0.011***	0.011***	0.013***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Mentor experience	−0.002	−0.002	−0.002	−0.001	−0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Mentor prestige	0.003	0.004	−0.003	0.001	0.017
	(0.019)	(0.019)	(0.019)	(0.021)	(0.043)
Hired university small	0.070	0.070	0.075	0.112	0.135
	(0.080)	(0.080)	(0.080)	(0.105)	(0.116)
Hired university medium	0.175**	0.175**	0.183***	0.154*	0.179*
	(0.069)	(0.069)	(0.069)	(0.089)	(0.102)
Hired university large	0.229***	0.228***	0.237***	0.217**	0.273***
	(0.066)	(0.066)	(0.066)	(0.085)	(0.098)
Hired university very large	0.502***	0.503***	0.507***	0.505***	0.475***
	(0.065)	(0.065)	(0.066)	(0.084)	(0.098)
Top 1 % university (dummy)		−0.028		−0.026	−0.142***
		(0.023)		(0.025)	(0.041)
Nobel prize (dummy)			0.015		
			(0.018)		
Fields medal (dummy)			0.045*		
			(0.024)		
Technical/medical (dummy)			0.044*		
			(0.023)		
Gender (unknown)				0.055	
				(0.062)	
Gender Female				−0.151***	
				(0.030)	
Math (dummy)					−0.102***
					(0.027)
Constant	2.833***	2.845***	2.822***	0.961	3.012***
	(0.596)	(0.596)	(0.596)	(0.938)	(0.620)
Theta	3.410***	3.410***	3.407***	3.391***	3.172***
	(0.085)	(0.085)	(0.085)	(0.095)	(0.124)
Observations	12,697	12,697	12,660	10,087	4,695
Log Likelihood	−26,129	−26,128	−26,069	−20,875	−9,782
Akaike Inf. Crit.	52,477	52,478	52,365	41,952	19,781

Note: Coefficients of a negative binomial model. All models include control variables for the hired country, for the country of training, and the decade (all dummies). Standard errors in brackets. *p < 0.1; **p < 0.05; ***p < 0.01.

Table A2.3
ZINB regression model coefficients for model 2–4. Clustered Standard Errors at university level. (dependent variable: mentees within 15 years)

	Logit - zero mentees			Count - number of mentees		
	(2)	(3)	(4)	(2)	(3)	(4)
Peers	−0.002	−0.002	−0.001	0.001***	0.0005**	0.0003**
	(0.0002)	(0.001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Co-mentees	0.034***	0.033***	0.033***	0.013***	0.013***	0.012***
	(0.004)	(0.006)	(0.004)	(0.003)	(0.004)	(0.003)
Mentor experience	0.007*	0.008	0.011**	−0.0005	−0.0003	−0.0003
	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
Mentor prestige	−1.326***	−1.218***	−1.170***	0.044	0.045	0.048
	(0.308)	(0.523)	(0.246)	(0.034)	(0.036)	(0.035)
Top 1 % university (dummy)		−0.355			0.042	
		(0.224)			(0.076)	
Nobel prize (dummy)			−0.507			0.053*
			(0.062)			(0.032)
Fields medal (dummy)			−0.535***			0.054
			(0.085)			(0.042)
Technical/medical (dummy)			−0.024			0.069*
			(0.057)			(0.037)
Constant	−2.487***	−2.252***	−2.342***	−0.064	−0.056	−0.092
	(0.481)	(0.869)	(0.472)	(0.119)	(0.189)	(0.120)
Observations	90,264	90,264	89,640	90,264	90,264	89,640
Log Likelihood	−74,599	−74,570	−74,041	−74,599	−74,570	−74,041

Note: Coefficients of a zero-inflated negative binomial model. All models include control variables for country and decade (dummy). Standard errors in brackets. *p < 0.1; **p < 0.05; ***p < 0.01.

References

Agrawal, A., Cockburn, I., McHale, J., 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. J. Econ. Geogr. 6, 571–591.

Agrawal, A., Mchale, J., Oettl, A., 2017. How stars matter: recruiting and peer effects in evolutionary biology ☆. Res. Policy 46, 853–867. https://doi.org/10.1016/j.respol.2017.02.007.

Andrew, N., 2013. Nurse education in practice clinical imprinting : the impact of early clinical learning on career long professional development in nursing. Nurse Educ. Pract. 13, 161–164. https://doi.org/10.1016/j.nepr.2012.08.008.

Azoulay, P., Liu, C.C., Stuart, T.E., 2017. Social influence given (partially) deliberate matching: career imprints in the creation of academic entrepreneurs. Am. J. Sociol. 122 (4), 1223–1271.

Baker, V.L., Lattuca, L.R., 2010. Developmental networks and learning : toward an interdisciplinary perspective on identity development during doctoral study. Stud. High. Educ. 35, 807–827. https://doi.org/10.1080/03075070903501887.

Barabási, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T., 2002. Evolution of the social network of scientific collaborations. Physica A 311, 590–614.

Baruffaldi, S., Visentin, F., Conti, A., 2016. The productivity of science & engineering PhD students hired from supervisors ' networks. Res. Policy 45, 785–796. https://doi.org/10.1016/j.respol.2015.12.006.

Bathelt, H., Malmberg, A., Maskell, P., 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. Prog. Hum. Geogr. 28, 31–56. https://doi.org/10.1191/0309132504ph469oa.

Berry, C., 1981. The Nobel scientists and the origins of scientific achievement. Br. J. Sociol. 32, 381–391.

Bonaccorsi, A., 2008. Search regimes and the industrial dynamics of science. Minerva 46, 285–315. https://doi.org/10.1007/s11024-008-9101-3.

Boschma, R., 2005. Proximity and innovation: a critical assessment. Reg. Stud. 39 (1), 61–74.

Breschi, S., Lissoni, F., 2009. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. J. Econ. Geogr. 9, 439–468. https://doi.org/10.1093/jeg/lbp008.

Broström, A., 2019. Academic breeding grounds : home department conditions and early career performance of academic researchers. Res. Policy 48, 1647–1665. https://doi.org/10.1016/j.respol.2019.03.009.

Burris, V., 2004. The academic caste system: prestige hierarchies in PhD exchange networks. Am. Sociol. Rev. 69, 239–264.

Cameron, C., Trivedi, P.K., 1998. Regression Analysis of Count Data. Cambridge University Press, New York.

Cameron, S.W., Blackburn, R.T., 1981. Sponsorship and academic career success. J. High. Educ. 52, 369377 https://doi.org/10.2307/1981284.

Chariker, J.H., Zhang, Y., Pani, J.R., Rouchka, E.C., 2017. Identification of successful mentoring communities using network-based analysis of mentor–mentee relationships across Nobel laureates. Scientometrics 111, 1733–1749. https://doi.org/10.1007/s11192-017-2364-4.

Clauset, A., Arbesman, S., Larremore, D.B., 2015. Systematic inequality and hierarchy in faculty hiring networks. Sci. Adv. 1 https://doi.org/10.1126/sciadv.1400005.

Clemente, F., 1973. Early career determinants of research productivity. Am. J. Sociol. 79.

Collins, H., 1985. Changing Order: Replication and Induction in Scientific Practice. Chicago University Press, Chicago.

Conti, A., Visentin, F., 2015. A revealed preference analysis of PhD students ' choices over employment outcomes. Res. Policy 44, 1931–1947. https://doi.org/10.1016/j.respol.2015.06.009.

Conti, A., Denas, O., Visentin, F., 2014. Knowledge specialization in PhD student groups. IEEE Transaction on Engineering Management 61, 52–67.

Cruz-Castro, L., Sanz-Menéndez, L., 2010. Mobility versus job stability: assessing tenure and productivity outcomes. Res. Policy 39, 27–38.

Dahlander, L., McFarland, D.A., 2013. Ties that last: tie formation and persistence in research collaborations over time. Adm. Sci. Q. 58, 69–110. https://doi.org/10.1177/0001839212474272.

David, S.V., Hayden, B.Y., 2012. Neurotree: a collaborative, graphical database of the academic genealogy of neuroscience. PLoS One 7. https://doi.org/10.1371/journal.pone.0046608.

Delamont, S., Atkinson, P., 2001. Doctoring uncertainty: mastering craft knowledge. Soc. Stud. Sci. 31, 87–107.

Delamont, S., Parry, O., Atkinson, P., 1997. Critical mass and pedagogic continuity: studies in academic habitus. Br. J. Sociol. Educ. 18, 533–549. https://doi.org/10.1080/0142569970180404.

Dericks, G., Thompson, E., Roberts, M., Phua, F., 2019. Determinants of PhD student satisfaction: the roles of supervisor, department, and peer qualities. Assess. Eval. High. Educ. 44, 1053–1068.

Dubois, P., Schlenker, J.R.J., 2014. Productivity and mobility in academic research : evidence from mathematicians. Scientometrics 98, 1669–1701. https://doi.org/10.1007/s11192-013-1112-7.

Ductor, L., 2015. Does co-authorship lead to higher academic productivity? Oxf. Bull. Econ. Stat. 77 (3), 385–407.

Dundar, H., Lewis, D.R., 1998. Determinants of research productivity in higher education. Res. High. Educ. 39, 607–631.

Ensher, E.A., Thomas, C., Murphy, S.E., 2001. Comparison of traditional, step-ahead, and peer mentoring on proteges' support, satisfaction, and perceptions of career success: a social exchange perspective. J. Bus. Psychol. 15, 419–438.

Espenshade, T.J., Rodriguez, G., 1997. Completing the Ph. D.: comparative performances of US and foreign students. Soc. Sci. Q. 593–605.

Feldon, D.F., Litson, K., Jeong, S., Blaney, J.M., Kang, J., Miller, C., Griffin, K., Roksa, J., 2019. Postdocs' lab engagement predicts trajectories of PhD students' skill development. Proc. Natl. Acad. Sci. 116, 20910–20916. https://doi.org/10.1073/pnas.1912488116.

Fortunato, S., Bergstrom, C.T., Börner, K., Evans, J.A., Helbing, D., Petersen, A.M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., Barabási, A., 2018. Science of science. Science 359. https://doi.org/10.1126/science.aao0185.

Frenken, K., Heimeriks, G.J., Hoekman, J., 2017. What drives university research performance? An analysis using the CWTS Leiden ranking data. Journal of Informetrics 11, 859–872. https://doi.org/10.1016/j.joi.2017.06.006.

Gargiulo, F., Caen, A., Lambiotte, R., Carletti, T., 2016. The classical origin of modern mathematics. EPJ Data Sci. https://doi.org/10.1140/epjds/s13688-016-0088-y.

Gertler, M.S., 2003. Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). J. Econ. Geogr. 3, 75–99. https://doi.org/10.1093/jeg/3.1.75.

Gilbert, R., 2009. The doctorate as curriculum: A perspective on goals and outcomes of doctoral education. In: Changing Practices of Doctoral Education. Routledge, New York.

Granovetter, M.S., 1973. The strength of weak ties. American Journal of Sociology1 78, 1360–1380.

Guberman, J., Saks, J., Shapiro, B., Torchia, M., 2006. Making the Right Moves: A Practical Guide to Scientific Management for Postdocs and new Faculty. Burroughs Wellcome Fund, Research Triangle Park.

Hackett, E.J., 1990. Science as a vocation in the 1990s: the changing organizational culture of academic science. J. High. Educ. 61, 241–279.

Hagstrom, W.O., 1964. Traditional and modern forms of scientific teamwork. Adm. Sci. Q. 9, 241–263. https://doi.org/10.2307/2391440.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2010. Multivariate Data Analysis, 7 ed. Prentice-Hall, Inc, Upper Saddle River, NJ, USA.

Hansen, M.T., 1999. The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits. Adm. Sci. Q. 44 (1), 82–111.

Hargens, L.L., 1975. Patterns of Scientific Research. American Sociological Association, Washington, DC.

Head, K., Li, Y.A., Minondo, A.. Geography, ties, and knowledge flows: evidence from citations in mathematics. CEPR Discussion Paper No. DP12942. Available at SSRN. https://ssrn.com/abstract=3182405.

Heinisch, D.P., Buenstorf, G., 2018. The next generation (plus one): an analysis of doctoral students' academic fecundity based on a novel approach to advisor identification. Scientometrics 117, 1–30. https://doi.org/10.1007/s11192-018-2840-5.

Horta, H., Santos, J.M., 2016. The impact of publishing during PhD studies on career research publication, visibility, and collaborations. Res. High. Educ. 57, 28–50. https://doi.org/10.1007/s11162-015-9380-0.

Howells, J., 2002. Tacit knowledge, innovation and economic geography. Urban Stud. 39, 871–884. https://doi.org/10.1080/0042098022012835.

Huang, D., 2015. Temporal evolution of multi-author papers in basic sciences from 1960 to 2010. Scientometrics 105, 2137–2147.

Huang, J., Gates, A.J., Sinatra, R., Barabási, A.-L., 2020. Historical comparison of gender inequality in scientific careers across countries and disciplines. Proc. Natl. Acad. Sci. 117, 4609–4616.

Katz, M.L., Shapiro, C., 1985. Network externalities, competition, and compatibility. Am. Econ. Rev. 75, 424–440.

Katz, R., 1982. The effects of group longevity on project communication and performance. Adm. Sci. Q. 27, 81–104.

Kram, K.E., Isabella, L.A., 1985. Mentoring alternatives: the role of peer relationships in career development. Acad. Manag. J. 28, 110–132.

Kyvik, S., 1993. Academic staff and scientific production. High. Educ. Manag. 5, 191–202.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., Sugimoto, C.R., 2013. Bibliometrics: global gender disparities in science. Nature 504, 211–213.

Lave, J., Wenger, E., 1991. Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, New York.

Lee, K., Brownstein, J.S., Mills, R.G., Kohane, I.S., 2010. Does collocation inform the impact of collaboration? PLoS One 5, 1–6. https://doi.org/10.1371/journal.pone.0014279.

Liu, J., Tang, T., Kong, X., Tolba, A., Al-Makhadmeh, Z., Xia, F., 2018a. Understanding the advisor–advisee relationship via scholarly data analysis. Scientometrics 116, 161–180.

Liu, L., Wang, Y., Sinatra, R., Giles, C.L., Song, C., Wang, D., 2018b. Hot streaks in artistic, cultural, and scientific careers. Nature 559, 396–399. https://doi.org/10.1038/s41586-018-0315-8.

Long, J.S., McGinnis, R., 1981. Organizational context and scientific productivity. Am. Sociol. Rev. 46, 422–442. https://doi.org/10.2307/2095262.

Long, J.S., McGinnis, R., 1985. The effects of the mentor on the academic career. Scientometrics 7, 255–280.

Long, J.S., Allison, P.D., McGinnis, R., 1979. Entrance into the academic career. Am. Sociol. Rev. 44, 816–830. https://doi.org/10.2307/2094529.

Long, J.S., Allison, P.D., McGinnis, R., 1993. Rank advancement in academic careers: sex differences and the effects of productivity. Am. Sociol. Rev. 58, 703–722.

Malmgren, R.D., Ottino, J.M., Nunes Amaral, L.A., 2010. The role of mentorship in protégé performance. Nature 465, 622–626. https://doi.org/10.1038/nature09040.

Martínez, G.L., Raffo, J., Saito, K., et al., 2016. Identifying the Gender of PCT Inventors. WIPO.

Meho, L.I., 2020. Highly prestigious international academic awards and their impact on university rankings. Quantitative Science Studies 2020 1 (2), 824–848.

Milojević, S., Radicchi, F., Walsh, J.P., 2018. Changing demographics of scientific careers: the rise of the temporary workforce. Proc. Natl. Acad. Sci. 115, 12616–12623.

MSC, 2020. Mathematical Reviews and zbMATH (Mathematics Subject Classification System).

Mukherjee, S., Uzzi, B., Jones, B.F., Stringer, M., 2017. How atypical combinations of scientific ideas are related to impact: the general case and the case of the field of geography. Knowledge and Networks. 243–267. https://doi.org/10.1007/978-3-319-45023-0.

Nooteboom, B., 2000. Learning by interaction: absorptive capacity, cognitive distance and governance. J. Manag. Gov. 4 (1), 69–92. https://doi.org/10.1023/A:1009941416749.

Olson, G.M., Olson, J.S., 2000. Distance matters. Human-Computer Interaction 0024, 139–178. https://doi.org/10.1207/S15327051HCI1523.

Paglis, L.L.L., Green, S.G., Bauer, T.N.T.N., 2006. Does adviser mentoring add value? A longitudinal study of mentoring and doctoral student outcomes. Res. High. Educ. 47, 451–476. https://doi.org/10.1007/s11162-005-9003-2.

Park, S.H.O., Gordon, M.E., 1996. Publication records and tenure decisions in the field of strategic management. Strateg. Manag. J. 17, 109–128.

Petersen, A.M., Riccaboni, M.H., Stanley, E., Pammolli, F., 2012. Persistence and uncertainty in the academic career. PNAS 109, 5213–5218. https://doi.org/10.1073/pnas.1121429109.

Polanyi, M., 1967. The tacit dimension. In: Prusak, L. (Ed.), Knowledge in Organizations. Routledge & Kegan Paul, London, UK, pp. 135–146. https://doi.org/10.1016/B978-0-7506-9718-7.50010-X.

Raffo, J., 2016. Worldwide Gender-Name Dictionary, WIPO Economics & Statistics Related Resources, 10. World Intellectual Property Organization - Economics and Statistics Division.

Ramesh, A., Singh, Y.P., 1998. Determinants of research productivity. Scientometrics 43, 309–329.

Roach, M., Sauermann, H., 2010. A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. Res. Policy 39 (3), 422–434.

Rossi, L., Freire, I.L., Mena-Chalco, J.P., 2017. Genealogical index: a metric to analyze advisor–advisee relationships. Journal of Informetrics 11, 564–582. https://doi.org/10.1016/j.joi.2017.04.001.

Ruef, M., Aldrich, H.E.H.E., Carter, N.M.N.M., 2003. The structure of founding teams: homophily, strong ties, and isolation among U.S. entrepreneurs. Am. Sociol. Rev. 68, 195–222. https://doi.org/10.2307/1519766.

Rutten, R., 2017. Beyond proximities: the socio-spatial dynamics of knowledge creation. Prog. Hum. Geogr. 41, 159–177.

Shibayama, S., 2019. Suitable development of science and scientists: academic training in life science labs. Res. Policy 48, 676–692. https://doi.org/10.1016/j.respol.2018.10.030.

Shibayama, S., Baba, Y., 2015. Impact-oriented science policies and scientific publication practices : the case of life sciences in Japan. Res. Policy 44, 936–950. https://doi.org/10.1016/j.respol.2015.01.012.

Shibayama, S., Walsh, J.P., Baba, Y., 2012. Academic entrepreneurship and exchange of scientific resources: material transfer in life and materials sciences in Japanese universities. Am. Sociol. Rev. 77, 804–830.

Shibayama, S., Baba, Y., Walsh, J.P., 2015. Organizational design of university laboratories: task allocation and lab performance in Japanese bioscience laboratories. Res. Policy 44, 610–622.

Singh, J., Fleming, L., 2010. Lone inventors as sources of breakthroughs: myth or reality? Manag. Sci. 56, 41–56. https://doi.org/10.1287/mnsc.1090.1072.

Sugimoto, C.R., Ni, C., Russell, T.G., Bychowski, B., 2011. Academic genealogy as an Indicator of Interdisciplinarity: an examination of dissertation networks in library and information science. J. Am. Soc. Inf. Sci. Technol. 62, 1808–1828. https://doi.org/10.1002/asi.21568.

Sverdlik, A., Hall, N.C., McAlpine, L., Hubbard, K., 2018. The PhD experience: a review of the factors influencing doctoral students' completion, achievement, and well-being. Int. J. Dr. Stud. 13, 316–388. https://doi.org/10.28945/4113.

Tartari, V., Perkmann, M., Salter, A., 2014. In good company: the influence of peers on industry engagement by academic scientists. Res. Policy 43, 1189–1203.

Tenn, J.S., 2016. Introducing astrogen: the astronomy genealogy project. Journal of astronomical history and heritage 19, 298–304.

Trapido, D., 2015. How novelty in knowledge earns recognition: the role of consistent identities. Res. Policy 44, 1488–1500. https://doi.org/10.1016/j.respol.2015.05.007.

Uzzi, B., 1997. Social structure and competition in interfirm networks : the paradox of embeddedness. Adm. Sci. Q. 42, 35–67.

Waaijer, C.J.F., Teelken, C., Wouters, P.F., van der Weijden, I.C.M., 2018. Competition in science: links between publication pressure, Grant pressure and the academic job market. Higher Education Policy 31, 225–243. https://doi.org/10.1057/s41307-017-0051-y.

Waldinger, F., 2010. Quality matters : the expulsion of professors and the consequences for PhD student outcomes in Nazi Germany. J. Polit. Econ. 118, 787–831.

Wang, J., Thijs, B., Glänzel, W., 2015. Interdisciplinarity and impact: distinct effects of variety, balance, and disparity. PLoS One 10, 1–18. https://doi.org/10.1371/journal.pone.0127298.

Way, S.F., Morgan, A.C., Larremore, D.B., Clauset, A., 2019. Productivity, prominence, and the effects of academic environment. Proc. Natl. Acad. Sci. 116, 10729–10733. https://doi.org/10.1073/pnas.1817431116.

Wickham, H., 2016. Rvest: easily harvest (scrape) web pages. R package version 0 (3), 2.

Wuestman, M., Frenken, K., Wanzenböck, I., 2020. A genealogical approach to academic success. PLoS One 15, e0243913.

Xing, Y., Fan, Y., Sinatra, R., Zeng, A., 2022. Academic Mentees Succeed in Big Groups, But Thrive in Small Groups. arXiv preprint. arXiv:2208.05304.