Zsófia Viktória Vida – István Péter Járay – Balázs Lengyel

PhD Students in Life Sciences Can Benefit from Team Cohesion

Introduction

As international competition in science accelerates, there has been a growing interest in the determinants of individual success in academia (Sinatra et al. 2016; Clauset et al. 2017; Fortunato et al. 2018). A general notion is that success breed success because recognized publications open new opportunities for funding and collaboration (Bol et al. 2018). This has directed attention to young scholars because achievements at early stage might generate future success (Wang et al. 2019). The relation between scientific collaboration and success has been investigated for long (de Solla Price – Beaver 1966; Luukkonen et al. 1993; Melin–Persson 1996; Katz–Martin 1997; Sonnenwald 2007; Hood–Wilson 2001; Sarigöl et al. 2014) and it is especially useful to understand early career success because young scholars need mentors to learn from and are more likely to stand out in case they work with successful supervisors (Sekara et al. 2018; Ma et al. 2020; Li et al. 2019). However, early careers might learn from more people at once, but it is less understood how future success of early career researchers depend on the team they work with.

Teams in scientific research are gaining dominance across all fields (Wuchty et al. 2007; Ziman, 1994). Research teams typically include postdocs, graduate and undergraduate students who collaborate with the principal investigator and other seniors of the group (Mali et al. 2012). Co-authorship across team members is frequently used to map collaboration networks (Beaver 2001; Glänzel–Schubert 2004), which are thought to influence success of projects in the sociology and management literatures (Uzzi–Spiro 2005). Two counteracting mechanisms are important in this respect. On the one hand, the project can create more novelty in case it combines diverse expertise by bringing together those who have not collaborated before (De Vaan et al. 2015; Vedres 2017; Zeng et al. 2021). On the other hand, team cohesion generated by shared co-authors, strong and persistent collaboration, trust and previous success can provide an environment, in which knowledge sharing are efficient (Uzzi–Spiro 2005, Aral – Van Alstyne 2011; Mukherjee et al. 2019). Thus, the question, whether early career researchers benefit more from diverse than from cohesive teams, is important because striving for novelty in scientific research and efficient learning for doctoral students are difficult achieve at once.

In this paper, we take a social network analysis approach to investigate co-authorship networks of early career researchers. To quantify diversity and coherence in the collaboration network of students and across the author teams they belong to, we apply the network constraint measure developed by Burt (1992, 2000, 2001). This measure takes high values in case co-authors of the PhD student work frequently together and the measure takes low values in case the PhD student works with co-authors who are otherwise not collaborating with each other. This measure has been very widely used to capture diverse knowledge access through connections in various contexts including creative industries (Juhász et al. 2020), innovation (Tóth–Lengyel 2021) and to capture the role of network cohesion in knowledge transfer (Reagens–McEvily 2003; Tortoreillo et al. 2012).

Our empirical case concerns researchers who have had a successful defense in any Hungarian doctoral school between 1993 and 2010. Our data contains information on the dissertation, including the scientific field and year of defense, and bibliometric information data comes from publication records of egos and their co-authors. We estimate the accumulated number of citations at the eighth year following defense that gives us a simple measure of success at the end of the early phase of academic career (Van Balen et al. 2012).

Cross-sectional linear regressions with year and scientific field dummies show that the number of papers published until the second year after the defense correlates negatively with accumulated citations but the impact of these papers correlates strongly with future impact. This finding indicates that thorough work focusing on a few but important papers is a much better strategy than producing many papers during doctoral studies. We find that in case of life science students, both the number of co-authors and most importantly the constraint measure correlates positively with future impact. These latter two co-efficients are not significant for other science fields and the significance in the case of life science also fade away at later stages in the career. These results provide new evidence that PhD students can benefit from working in a cohesive research team probably because this provides a better learning environment.

Materials and Methods

Data

We combine two data sources to collect information about early-career scholars. Data on doctoral defenses have been collected from www.doktori.hu, an openly available collection of all successful PhD theses defended in Hungarian doctoral schools starting from 1993, the year when the PhD system was introduced in the country. We downloaded data from the website in January 2017. This data contains 16,151 Hungarian PhD students who defended their theses until that date and information include the ID and name of every PhD student, the title of their thesis, the year of defense, scientific area, the name of supervisors. Our second data source is the Hungarian Scientific Bibliography database (MTMT) that contains the scientific publications' metadata of all active Hungarian researchers.

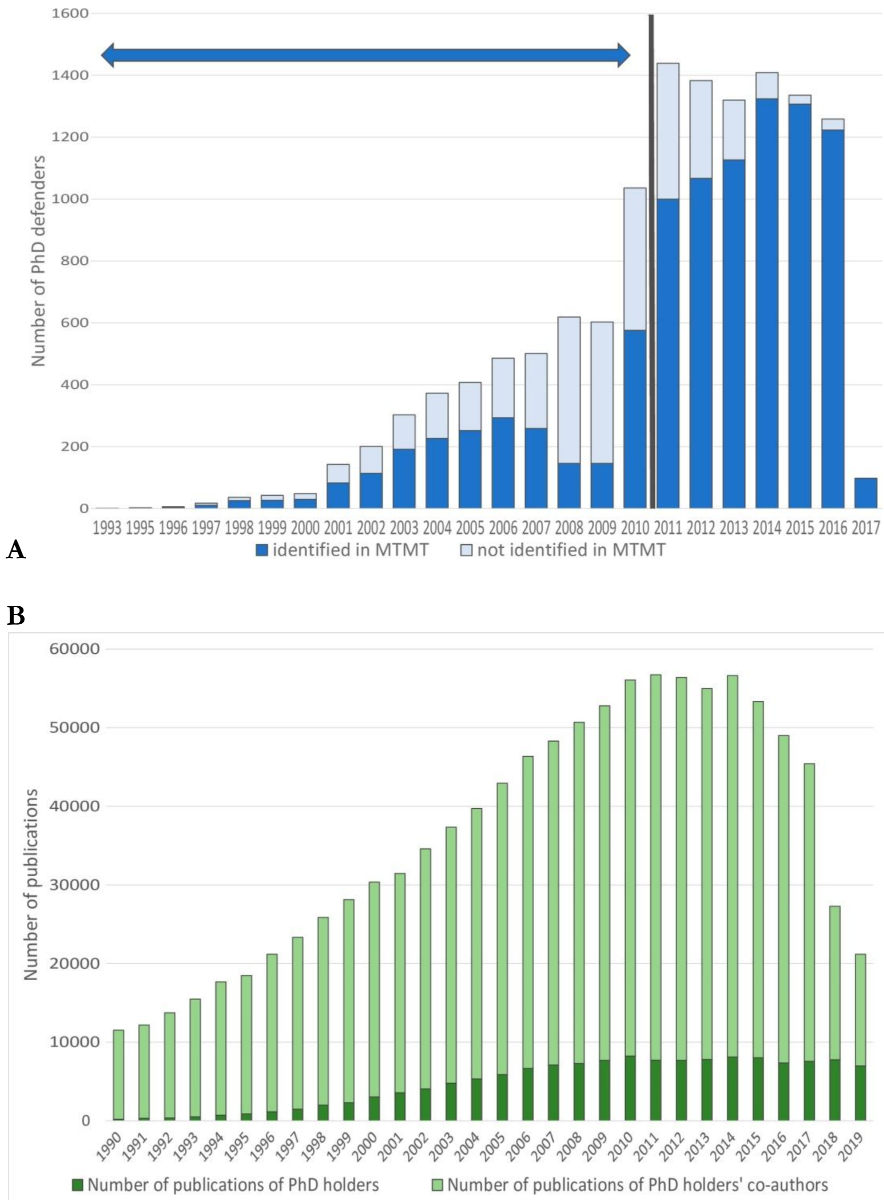


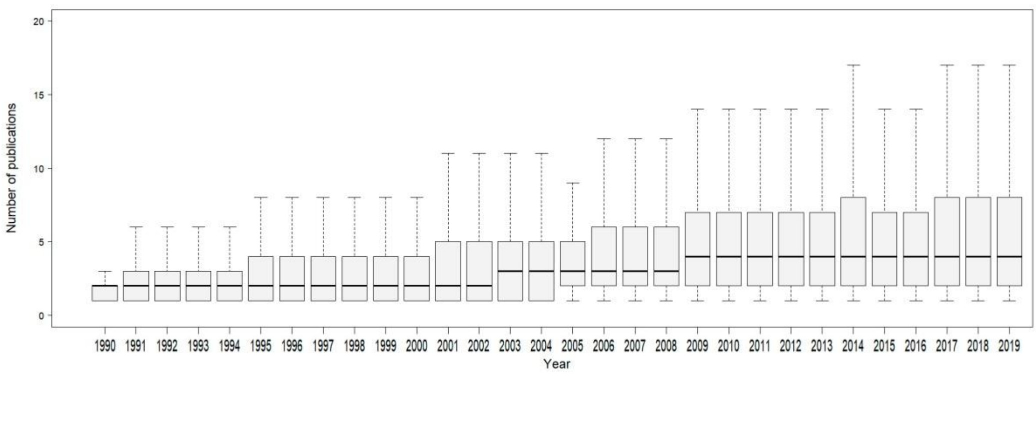
Figure 1: PhD students and publications in the data

A. The number of PhD students by year in www.doktori.hu (light blue) and the successfully identified PhD students in the MTMT database (navy blue). Data includes all students who defended between 1993 and 2017 but the analysis will focus on those who defended in the 1993–2010 period. **B.** The number of publications by the Hungarian PhD holders (year of defenses between 1993–2010) and their co-authors between 1990 and 2019.

The two databases can be matched on the individual student level. The doktori.hu database even contained student IDs in the MTMT data for 23% of students. The rest of students were matched by hand using name and scientific field. We could identify 60% of the PhD students in the MTMT data. The number of PhD students who can be matched with an MTMT profile is illustrated in *Figure 1A* by the year of defense. In our regression exercise, we focus on the future impact of PhD thus restrict the analysis to the 2,061 PhD students who defended thesis in the 1993-2010 period.

Bibliometric data have been downloaded after the identification of PhD students in MTMT. This happened in two steps. First, we have downloaded all 272,954 publication records of the identified 9,415 PhD students in 2017. Then, we identified 20,139 co-authors of PhD students in MTMT and downloaded their publication records in 2020. This final bibliometric dataset contains records of 1,205,184 papers published by 43,485 authors altogether between 1990–2019. Note that only those authors are included who are affiliated in Hungarian institutions and must have registered on MTMT. There are around 50 thousand MTMT accounts altogether, meaning that our data collection has covered around 86% of the total scientific community in the country. *Figure 1B* illustrates the number of all publications from the entire career of those PhD students who defended between 1993 and 2010 and their co-authors. As we are interested in the production of the PhD holders, we analyzed only papers published between 1990 and 2019.

A



B

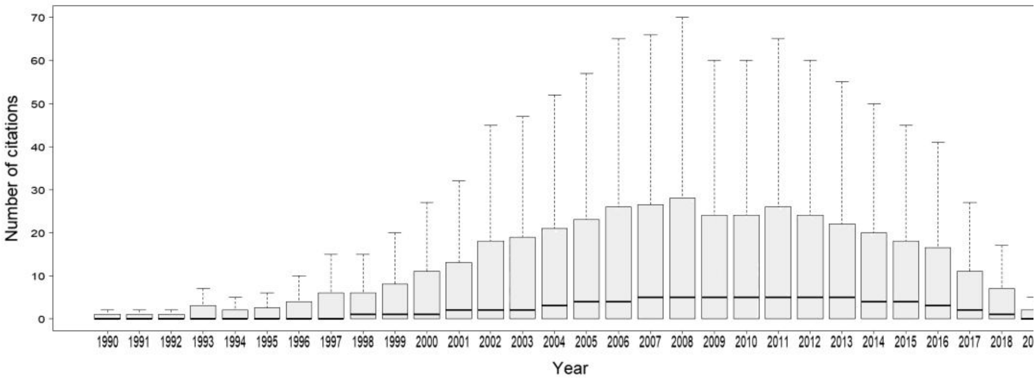


Figure 2: Citations

A. The distribution of number of publications by Hungarian PhD holders (year of defense 1993–2010) between 1990 and 2019. **B.** The distribution of number of citations between 1990 and 2019 by Hungarian PhD holders (year of defense 1993–2010) in 2020.

Methods

Publication variables

Measuring scientific success, especially individual scientific performance is a complex problem. Traditionally, it is based on production (publication) numbers, scientific impact (citation numbers) and structural measurements for example the network characteristics of authorship (Van Balen et al. 2012; Glänzel et al. 2019). However, the raw citation number depends on several factors, such as the year of publication, the research field, the document type (e.g. research article, review article or proceedings), the journal characteristics (e.g. frequency of occurrence, number of articles in the journal). It is easy to see that for example the earlier an article has appeared, the more citations it could receive. The citation habits are different in individual research fields, so to compare two citation measures we must do it in the same research area. The various document types use different number of references. Thus, the comparison is more accurate if it is made within the same document type. Moreover, the journal characteristic also can cause a bias on raw citation numbers. The solution for these problems is using normalized citation numbers.

A

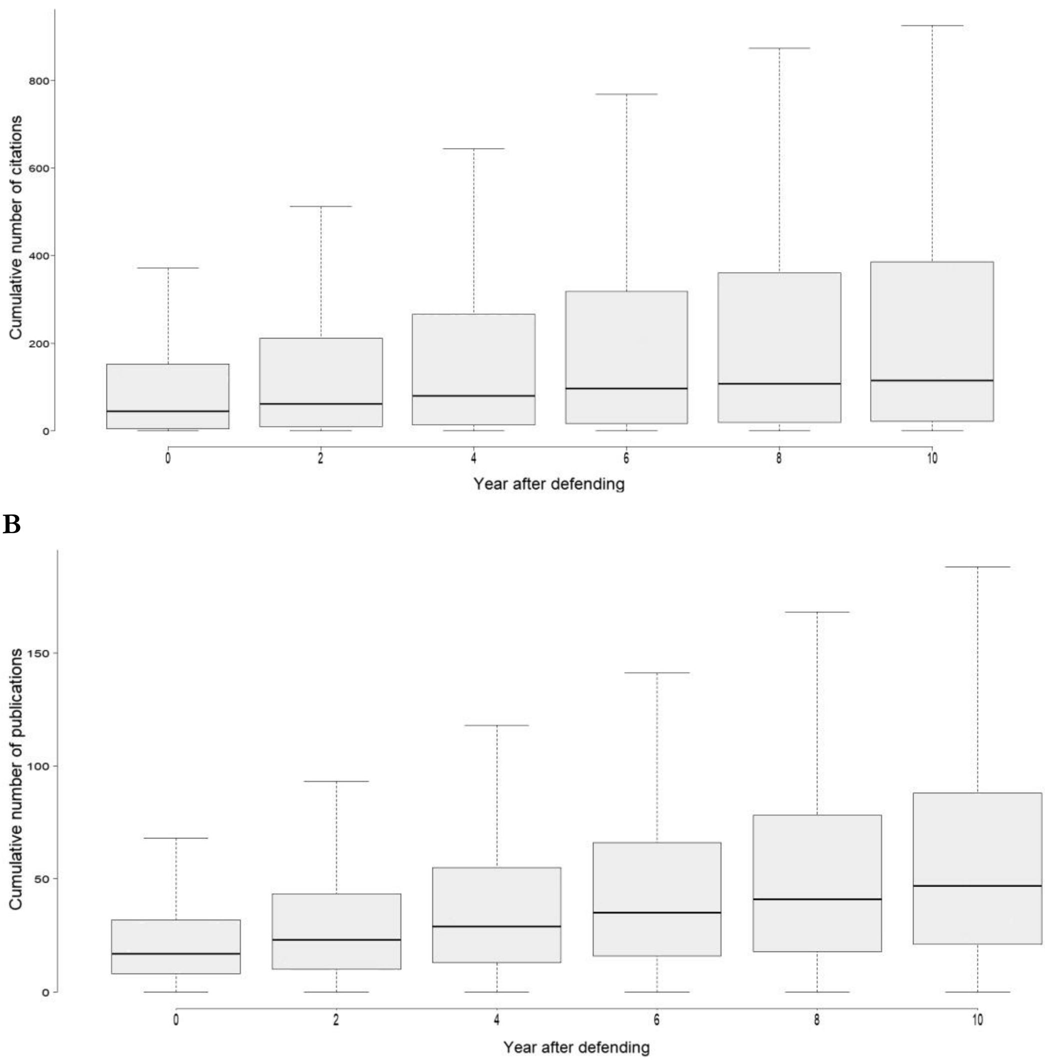


Figure 3: Citations and papers of PhD students

A. Cumulative number of citations by the Hungarian PhD holders (year of defense 1993–2010) in 0–10 years after defense. **B.** Cumulative number of papers by Hungarian PhD holders (year of defense 1993–2010) in 0-10 years after their defense.

In our case the MTMT database contains only raw citation number in the year of downloading, in our case in 2020. To handle this problem, we compared each PhD holder in two-year periods, in a cumulative way using the year of their PhD defense as a starting point. We compared PhD holders by their research field. The MTMT database contains document types as articles, books, others, but we were unable to distinguish research articles and review articles. So, in our study we did not consider the document types. *Figure 3A* shows the cumulative number of citations of the examined PhD holders, while *Figure 3B* shows the cumulative number of publications by the examined PhD holders.

Network variables

To answer our research questions, whether cohesive or diverse co-authorship network structure favours the success of a young researcher, we analyzed the weighted and dynamic ego-networks of PhD holders. Such networks were generated from the publication records. These ego-networks include the PhD student in the center (ego), to which co-authors (alters in the ego-network terminology) are connected to. Links are undirected but weighted by the number of co-authored papers. The networks are dynamic, such that we add new collaborators and new links to the ego-network of individual PhD students as new papers are published, but do not delete ties over the years. Since we have access to the publications of co-authors, the links between alters contain those publications that were not authored by the PhD student.

Cohesive networks are dense and include strong, high-bandwidth ties (Aral 2016). That is, co-authors are frequently publishing with each other. Such network structures are thought to capture an environment, in which shared work experience and developed trust facilitate learning from peers. In cohesive networks knowledge transfer is faster and more efficient such that the PhD student can learn complex knowledge easier (Reagens–McEvily 2003). On the contrary, diverse networks, in which co-authors have not worked with each other but with the PhD student, capture an environment that provides the student with diverse capabilities of co-authors. In such networks, innovation and novel combination is more likely (Burt 2001). In case the student can integrate distinct pieces of knowledge, diverse networks might help her/him to publish papers with high degree of novelty.

We used Burt (2000) constraint indicator that characterizes ego-networks in the cohesive-diverse continuum using the formula:

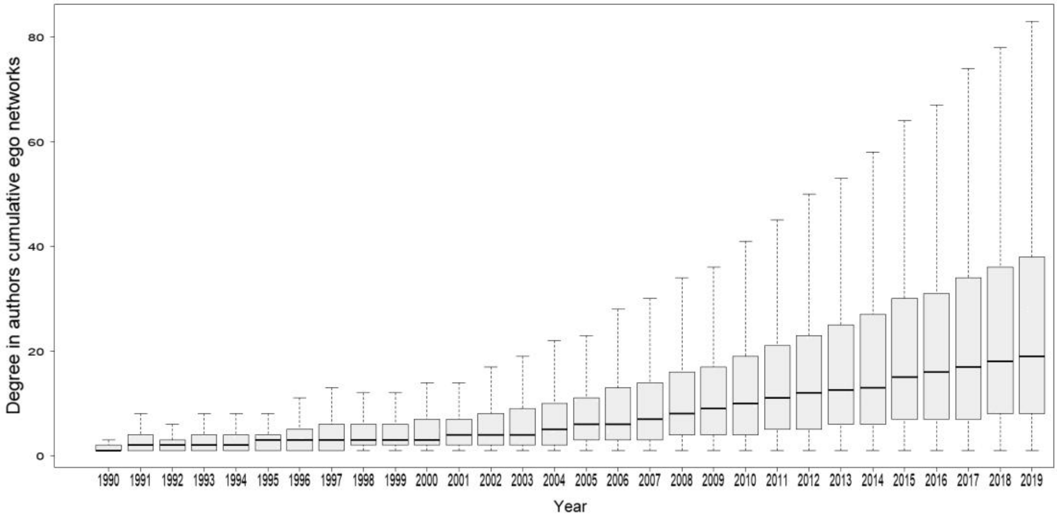
$$CON_i = \sum_j \left(p_{ij} + \sum_q p_{iq} p_{qj}\right)^2 ; q \neq i, j,$$

(Eq. 1)

where p_{ij} and p_{iq} is the number of papers that PhD student *i* has co-authored with colleagues *j* and q, and p_{qj} is the number of papers that *j* and *q* has co-authored without *i*. The indicators takes high values in case co-authors publish intensively together and low values are produced when co-authors do not publish together.

As the size of ego-networks grow, the probability that co-authors are connected might decrease, which has been often found in co-author networks (see for example Tóth–Lengyel 2021). Thus, one must consider the degree of PhD students as well that is their number of co-authors.

A



B

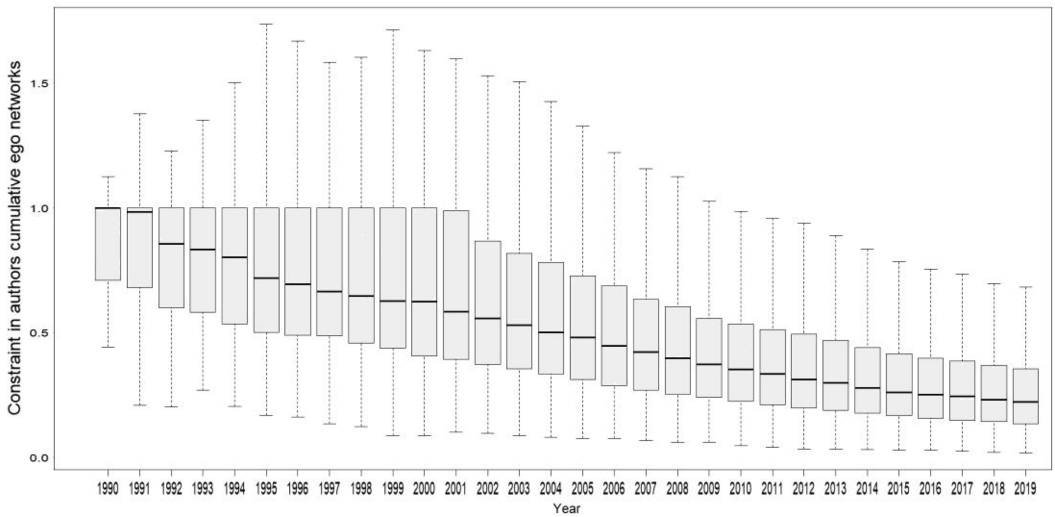


Figure 4: Degree and constraint of ego-networks over time
A. The distribution of degree in cumulative ego networks of the Hungarian PhD holders (year of defense 1993–2010) between 1990–2019. **B.** The distribution of constraint in cumulative ego networks of the Hungarian PhD holders (year of defense 1993–2010) between 1990–2019.

The distributions of degree and constraint are depicted in *Figure 4*. As expected, these two indicators change the opposite direction. The number of relations rise in time *(Figure 4A)* which is obvious because we used a cumulative ego network and did not erase former co-authorships. We can also see an increasement in the distribution of degree in time, which means that while some researchers could evolve their co-authorship networks after their PhD, others had a narrowed scientific network. The distribution of constraint lightly decreases *(Figure 4B),* and the median of constraint also falls in time. The cause is that the size of ego networks rise in time and those PhD holders who get more and more co-authors have also a more and more diverse collaboration network.

Table 1: Pearson correlation of network parameters of cumulative ego networks. 2 years (below diagonal) and 8 years (above diagonal) after Hungarian PhD holders defense (year of defense 1993–2010)

		1	2	3	4	5	6	7
Degree	1.	1	–0.26	–0.57	–0.44	–0.48	–0.42	–0.41
Betweenness centrality	2.	–0.18	1	0	0.26	0.37	0.14	0.03
Constraint	3.	–0.61	–0.06	1	0.36	0.44	0.71	0.70
Global clustering with ego	4.	–0.31	0.12	0.22	1	0.96	0.46	0.46
Global clustering without ego	5.	–0.37	0.22	0.30	0.93	1	0.68	0.68
Graph density with ego	6.	–0.45	0.08	0.71	0.40	0.58	1	0.94
Graph density without ego	7.	–0.44	–0.02	0.70	0.35	0.58	0.96	1

We calculated further measures that might be also used to characterize cohesion and diversity in ego-networks. Betweenness centrality quantifies diversity in the network of PhD students by measuring the number of shortest paths in the network that go through the ego. The higher betweenness centrality of the ego the more diversity in the network. Global clustering quantifies the fraction of closed triangles in the network among all possible triangles, while network density measures the fraction of observed ties among all possible ties with the ego. The higher these measures the higher cohesion in the ego-network.

Table 1 reports Pearson correlation coefficients between network parameters at the second and eight year after PhD defense. As expected, we find a negative correlation between degree and all other network indices. Constraint is strongly correlated with network density. We have run alternative regression specifications with the network measures in Table but only found significant results for constraint.

Regression framework

Our data enables us to capture impact of publications as a snapshot in 2020 by the total number of citations received until then. This allows for cross-sectional specification, in which we can compare students who finished in the same year and consider publications that they produced until a certain year after defense. This way, we can avoid the problem that earlier publications have more time to collect citations.

To answer the question whether cohesive co-authorship networks of PhD students during their studies help their future success, we estimate the number of accumulated citations ($CIT_{i,t+8}$) of student i to the paper that she or he published until the 8th (t+8) year following defense at year t with the following equation:

$$CIT_{i,t+8} = \alpha + \beta_1 CIT_{i,t+2} + \beta_2 PAP_{i,t+2} + \beta_3\left(PAP_{i,t+8} - PAP_{i,t+2}\right) + \beta_4 DEG_{i,t+2} + \beta_5 CON_{i,t+2} + \theta_i + t_i + \varepsilon_i, \quad \text{(Eq. 2)}$$

where $CIT_{i,t+2}$ denotes citations to papers published until the second year after defense, $PAP_{i,t+2}$ and $PAP_{i,t+8}$ are papers published until the second and eight year after defense, $DEG_{i,t+2}$ is the degree, and $CON_{i,t+2}$ is the constraint measure of the student's co-author network, θ_i is scientific area-specific fixed-effect, t_i is year dummies and ε_i is the error term.

We used linear regression models (OLS specification) to the citation number at 8 years after defending with explanatory variables as degree, constraint and the cumulative number of papers and citations at 2 years after defending. As fix effect we use years and research fields of doctoral schools. These latter refer to 54 categories of research fields defined by the National Accreditation Committee: exactly one research field has been assigned to each doctoral school. In our case, success is determined solely with the raw citation number as of downloading time in 2020. All variables are log-transformed.

Results

Table 2 reports results of an OLS regression of estimating Eq.2. In columns 1-3, we estimate citations to papers that were published until the 8th year following defense with variables that capture publications and co-authorship until the 2nd year following defense. We introduce variables in a stepwise manner such that a baseline model is run in column 1 and networks variables are introduced in columns 2 and 3. Throughout the models, we found a very strong positive correlation between CIT_{t+2} and CIT_{t+8} that is a trivial relation but has importance in our empirical exercise. Because citations are collected for all publications in 2020, CIT_{t+8} includes CIT_{t+2}. However, the very high correlations also mean that at most of the citations at the end of the early career stage are received to the publications that were published during or closely after PhD studies. PAP_{t+2} is negatively correlated while ΔPAP is positively correlated with the dependent variable. These findings suggest that due to accumulation of citations, the best strategy for PhD students is to produce few but high impact papers that will help them to collect citations in their early career.

Table 2: Estimates for Citations 8 year after defense, OLS regressions with year and scientific field fixed effects and robust standard errors

	(1)	(2)	(3)
CIT (log)	0.899***	0.894***	0.893***
	(0.009)	(0.009)	(0.009)
PAP (log)	–0.211***	–0.217***	–0.226***
	(0.017)	(0.019)	(0.020)
Δ PAP (log)	0.317***	0.312***	0.313***
	(0.010)	(0.011)	(0.011)
DEG (log)		0.043**	0.087***
		(0.018)	(0.028)
CON (log)			0.244**
			(0.121)
Constant	2.327***	2.324***	2.141***
	(0.374)	(0.372)	(0.383)
N	2,061	1,948	1,948
R2	0.919	0.917	0.918

Note: * p<0.1, ** p<0.05, *** p<0.01

In column 2, we introduce Degree that leaves correlations of other covariates almost unchanged. *DEG* is positively correlated with CIT_{t+8} suggesting that the number of co-authors facilitates citations. Note that there might be various mechanisms at play: citations might grow with the number of co-authors because they can also cite the paper or spread the word, and alternatively, the project and the PhD student can gain from working with and learning from many collaborators.

Constraint is positively correlated with CIT_{t+8}. Controlling for *DEG*, the number of publications, and including year and scientific field dummies, *CON* quantifies the extent to which co-authors of the PhD student have collaborated in publications that are published until the second year after the defense of the student. Our finding suggests that such cohesive ego-networks are beneficial for PhD students. Because we also control for the citations to papers, this finding confirm that PhD students benefit the most from working in cohesive collaboration networks because these create efficient learning environments.

Correlations of independent variables indicate that the models are not violated by multicollinearity. The highest value of the Pearson correlation coefficients is $\varrho=0.41$ between DEG and PAP_{t+2}. We document the correlation between *DEG* and *CON* in *Table 1* ($\varrho=0.61$), but the inclusion of these variables together are conceptually motivated as we describe before. Further, the inclusion of *CON* in Model 3 does not substantially influence the coefficient of *DEG*.

Table 3: Estimates for Citations 8 year after defense by scientific areas, OLS regressions with year and scientific field fixed effects and robust standard errors

	Sciences	Life Sciences	Engineering	Social Sciences
	(1)	(2)	(3)	(4)
CIT_t2 (log)	0.919***	0.870***	0.917***	0.910***
	(0.020)	(0.017)	(0.034)	(0.026)
PAP (log)	–0.294***	–0.220***	–0.222***	–0.267***
	(0.046)	(0.033)	(0.072)	(0.058)
Δ PAP (log)	0.357***	0.296***	0.255***	0.388***
	(0.023)	(0.018)	(0.034)	(0.033)
DEG (log)	0.068	0.162**	0.191*	0.039
	(0.063)	(0.054)	(0.110)	(0.075)
CON (log)	0.173	0.697***	0.654	0.060
	(0.255)	(0.274)	(0.401)	(0.303)
Constant	2.149***	2.324***	1.124***	0.434
	(0.438)	(0.372)	(0.422)	(0.536)
N	437	1,948	155	279
R2	0.919	0.917	0.942	0.910

Note: * p<0.1, ** p<0.05, *** p<0.01

In *Table 3,* we report full regression models decomposed into four big scientific areas such as Science, Life Science, Engineering, Social Science. To achieve these scientific areas, we have grouped the 54 scientific fields. We found that *DEG* and *CON* is significant for the Life Science subsample while *DEG* only is weakly significant in Engineering. Thus, cohesive research environment is important for Life Science students and less for students in other fields.

Conclusion

In this study we examined the success of students who defended theses in Hungarian doctoral schools between 1990 and 2010 by looking at their publications records and accumulated citations in 2019. Our bibliometric database contains the PhD students' publications and their co-authors' publications as well between 1990 and 2019. We analyzed whether cohesive or diverse co-author network structure gives a better chance to a young researcher to stand out in terms of citations eight years after defense. Linear regression models suggest that those students who participate in cohesive collaboration networks, receive significantly more citations at the end of their career. This result highlights the need for strong collaborations and effective learning environment during doctoral studies. However, our results regarding the structure of co-author networks are specific to Life Science students. Thus, cohesion is mostly important in areas where new knowledge is produced in teamwork.

The present paper contributes to a growing literature, in which studies try to determine factors that support the future success of young researchers. Li and co-authors (2019) demonstrate that those students who publish with top scientists had a greater chance to be more successful 20 years later. Moreover, this effect is more important in the case of PhD students affiliated with a less prestigious PhD school. Sarigöl and colleagues (2014) illustrate a similar phenomenon: a paper gets more citations if its' authors are central in the large co-author network of their field. We add to this discussion by studying the co-author ego-networks of PhD students. Our findings confirms that the structure of the group collaboration matters for the future academic career of students.

We also find that those students are more successful, measured in citations, who focus on few papers. These results are robust across all large scientific fields. By concentrative efforts into a small number of publications, the students are able to achieve higher quality papers that might be accepted to better journals. Because citations typically demand several years to accumulate, students need high-impact papers already at the beginning of their career to stand out later when they are at the end of the early-career stage. This can help them in research proposals and thus facilitate academic career on the long run as well.

References

Aral, S. (2016): The future of weak ties. *American Journal of Sociology*, 121/6, 1931–1939.

Aral, S. – Van Alstyne, M. (2011): The diversity-bandwidth trade-off. *American Journal of Sociology,* 117/1, 90–171.

Beaver, D. D. (2001): Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics*, 52/3, 365–377.

Bol, T. – De Vaan, M. – Van de Rijt, A. (2018): The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115/19, 4887–4890.

Burt, R. S. (2009): *Structural holes: The social structure of competition*. Harvard University Press.

Burt, R. S. (2000): The network structure of social capital. *Research in organizational behavior*, 22, 345–423.

Burt, R. S. (2001): *Structural Holes versus Network Closure as Social Capital in Social Capital: Theory and Research.* Edited by Nan Lin, Karen S. Cook and R. S. Burt. New York: Aldine de Gruyter, 31–56.

Clauset, A. – Larremore, D. B. – Sinatra, R. (2017): Data-driven predictions in the science of science. *Science*, 355/6324, 477–480.

De Vaan, M. – Stark, D. – Vedres, B. (2015): Game changer: The topology of creativity. *American Journal of Sociology*, 120/4, 1144–1194.

De Solla Price, D. J. – Beaver, D. (1966): Collaboration in an invisible college. *American Psychologist*, 21/11, 1011–1018.

Fortunato, S. – Bergstrom, C. T. – Börner, K. – Evans, J. A. – Helbing, D. – Milojević, S. – Vespignani, A. (2018): Science of science. *Science*, 359/6379, eaao0185.

Glänzel, W. – Schubert, A. (2004): Analysing scientific networks through co-authorship. In Moed, H. F. – Glänzel, W. – Schmoch, U. (eds.): *Handbook of quantitative science and technology research.* Dordrecht: Springer, 257–276.

Hood, W. – Wilson, C. (2001): The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*, 52/2, 291–314.

Juhász, S. – Tóth, G. – Lengyel, B. (2020): Brokering the core and the periphery: Creative success and collaboration networks in the film industry. *Plos one*, 15/2, e0229436.

Katz, J. S. – Martin, B. R. (1997): What is research collaboration?. *Research Policy,* 26/1, 1–18.

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

Luukkonen, T. – Tijssen, R. – Persson, O. – Sivertsen, G. (1993): The measurement of international scientific collaboration. *Scientometrics*, 28/1, 15–36.

Ma, Y. – Mukherjee, S. – Uzzi, B. (2020): Mentorship and protégé success in STEM fields. *Proceedings of the National Academy of Sciences*, 117/25, 14077–14083.

Mali, F. – Kronegger, L. – Doreian, P. – Ferligoj, A. (2012): Dynamic scientific co-authorship networks. In Scharnhorst, A. – Börner, K. – Van den Besselaar, P. (eds.): *Models of science dynamics*. Berlin–Heidelberg: Springer, 195–232.

Mukherjee, S. – Huang, Y. – Neidhardt, J. – Uzzi, B. – Contractor, N. (2019): Prior shared success predicts victory in team competitions. *Nature Human Behaviour*, 3/1, 74–81.

Melin, G. – Persson, O. (1996): Studying research collaboration using co-authorships. *Scientometrics*, 36/3, 363–377.

Reagans, R. – McEvily, B. (2003): Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, 48/2, 240–267.

Sarigöl, E. – Pfitzner, R. – Scholtes, I. – Garas, A. – Schweitzer, F. (2014): Predicting scientific success based on coauthorship networks. *EPJ Data Science*, 3, 1–16.

Sekara, V., Deville, P. – Ahnert, S. E. – Barabási, A. L. – Sinatra, R. – Lehmann, S. (2018): The chaperone effect in scientific publishing. *Proceedings of the National Academy of Sciences*, 115/50, 12603–12607.

Sinatra, R. – Wang, D. – Deville, P. – Song, C. – Barabási, A. L. (2016): Quantifying the evolution of individual scientific impact. *Science*, 354, 6312.

Sonnenwald, D. H. (2007): Scientific collaboration. In B. Cronin (ed.): *Annual Review of Information Science and Technology*, Vol. 41, Medford (NJ): Information Today, 643–681.

Tortoriello, M. – Reagans, R. – McEvily, B. (2012): Bridging the knowledge gap: The influence of strong ties, network cohesion, and network range on the

transfer of knowledge between organizational units. *Organization Science*, 23/4, 1024–1039.
Tóth, G. – Lengyel, B. (2019): Inter-firm inventor mobility and the role of co-inventor networks in producing high–impact innovation. *The Journal of Technology Transfer*, 1–21.
Van Balen, B. – Van Arensbergen, P. – Van Der Weijden, I. – Van Den Besselaar, P. (2012): Determinants of success in academic careers. *Higher Education Policy,* 25/3, 313–334.
Vedres, B. (2017): Forbidden triads and creative success in jazz: The Miles Davis factor. *Applied Network Science,* 2/1, 1–25.
Waltman, L. – Calero‐ Medina, C. – Kosten, J. – Noyons, E. C. – Tijssen, R. J. – Van Eck, N. J. – Wouters, P. (2012): The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American society for information science and technology*, 63/12, 2419–2432.
Wang, Y. – Jones, B. F. –Wang, D. (2019): Early-career setback and future career impact. *Nature Communications*, 10/1, 1–10.
Wuchty, S. – Jones, B. F. – Uzzi, B. (2007): The increasing dominance of teams in production of knowledge. *Science*, 316/5827, 1036–1039.
Zeng, A. – Fan, Y. – Di, Z. – Wang, Y. – Havlin, S. (2021): Fresh teams are associated with original and multidisciplinary research. *Nature Human Behaviour*, 1–9.
Ziman, J. (1994): *Prometheus Bound. Science in a Dynamic Steady State*. Cambridge, Cambridge University Press.