

The first cut is the deepest: repeated interactions of coauthorship and academic productivity in Nobel laureate teams

Ho Fai Chan¹ · Ali Sina Önder² · Benno Torgler^{1,3}

Received: 3 March 2015 / Published online: 30 November 2015
© Akadémiai Kiadó, Budapest, Hungary 2015

Abstract Despite much in-depth investigation of factors influencing the coauthorship evolution in various scientific fields, our knowledge about how efficiency or creativity is linked to the longevity of collaborative relationships remains very limited. We explore what Nobel laureates' coauthorship patterns reveal about the nature of scientific collaborations looking at the intensity and success of scientific collaborations across fields and across laureates' collaborative lifecycles in physics, chemistry, and physiology/medicine. We find that more collaboration with the same researcher is actually no better for advancing creativity: publications produced early in a sequence of repeated collaborations with a given coauthor tend to be published better and cited more than papers that come later in the collaboration with the same coauthor. Our results indicate that scientific collaboration involves conceptual complementarities that may erode over a sequence of repeated interactions.

Keywords Innovation · Scientific collaboration · Team formation · Nobel laureates

Introduction

Dramatic changes in science over the past decades have increased task complexity, reshaping how scientists cooperate and turning science into a team effort (Katz and Martin 1997; Adams et al. 2005). In particular, the one-author-per-paper trend that dominated

✉ Benno Torgler
benno.torgler@qut.edu.au

¹ Queensland Behavioural Economics Group (QuBE), School of Economics and Finance, Queensland University of Technology, 2 George St, GPO Box 2434, Brisbane, QLD 4000, Australia

² Department of Economics, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany

³ Center for Research in Economics, Management and the Arts (CREMA), Zurich, Switzerland

science from the 1600s until around the 1920s decreased in the 1950s, was barely visible by the 1980s (Greene 2007), and has become a rarity in scientific journals today. For example, of the 700 reports published in *Nature* in the first 10 months of 2008, only six were single author papers (Whitfield 2008). Our understanding of such collaboration is informed by visualisation of collaborative patterns (Newman 2004) and an evolving understanding of the principles of team formation (Guimera et al. 2005; Milojević 2014), which provides useful insights into optimal team size. The emerging use by scientists of collaborative indexes to more effectively measure researchers' scientific impact (Stallings et al. 2013) also suggests that in the past few decades, single authors have performed worse than teams (Wuchty et al. 2007). Nevertheless, knowledge of how teams perform over time remains limited.

To help fill this void, we explore the productivity patterns of repeated scientific collaborations by Nobel laureates and their collaborators, thus ensuring a homogenous focus group of productive scientific “stars” with intellectual human capital of extraordinary scientific value. In particular, laureates are homogenous in their capacity to produce successful, innovative ideas and attract fairly able co-authors (Zuckerman 1996), which allows us to focus on team efficiency while holding team talent constant.

Data and descriptive analysis

Our dataset consists of 34,448 publications registered in Scopus (up to 2008) of 192 Nobel laureates who received the Nobel Prize in chemistry (56), physics (69), or physiology/medicine (67) between 1970 and 2000. The dataset includes 43,451 Nobel laureate coauthor pairs, for whose publications citation records are traceable up to 2014. The patterns of laureates' accumulation of coauthors are similar in different fields. Although most Nobel laureates cooperate with fewer than 160 different coauthors over their academic lifecycle, a few cooperate with over 1000 different coauthors. The long tails of the histograms (Fig. 4) somewhat reflect the fact that “hyper-authorships” tend to be the product of highly complex subfields such as biomedicine or high-energy physics (Cronin 2001).

Our first analysis explores the arrival of new coauthors and the intensity of coauthorship over Nobel laureates' academic lifecycle. Figure 1 shows the number of new coauthors that appear in laureates' publications at a given age. The patterns for the arrival of new coauthors are comparable in chemistry, physics, and physiology/medicine before laureates reach age 60. Age 60 marks the peak for arrival of new coauthors in chemistry and physics, although there seems to be no clear peak in physiology/medicine.

The intensity of coauthorship captures the number of total collaborations between a laureate and a given coauthor (Fig. 2). “Laureates' Age” corresponds to the laureates' age of first collaboration, with the vertical axis depicting the average number of collaborations between Nobel laureates and arriving coauthors (i.e., when collaboration begins) at a given age for the laureate. In chemistry and medicine, early collaborations tend to be more intense (albeit with a large variance). In physics, however, laureates' intensities of coauthorship tend to show a positive trend at younger ages but no clear peak is observed.

In addition, we refine the measure of collaborative intensity by taking into account the number of coauthors in each publication. For example, the level of collaboration intensity between coauthors on a publication with five coauthors may differ from the intensity experienced on a paper with two coauthors. We therefore utilize the *A*-index developed by

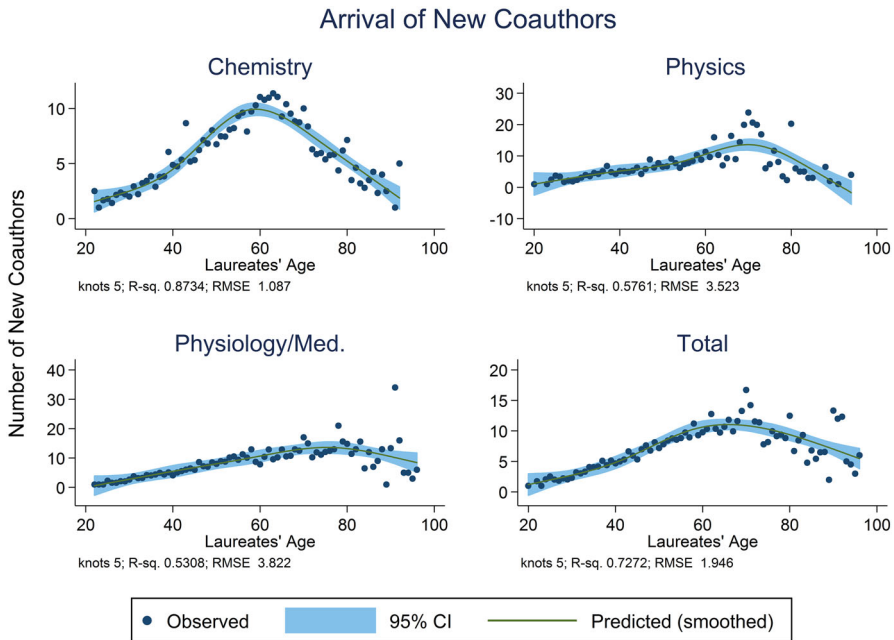


Fig. 1 Arrival of new coauthors by field. *Note:* Smoothed values are computed using restricted cubic spline

Stallings et al. (2013) to account for each coauthor’s share in each publication. The *A*-index provides an estimation of the individual contribution (the relative share of credit among coauthors). Computation of the *A*-index requires grouping of coauthors according to their relative contributions to the publication. The groups are then ranked by the level of contributions. For the authors in the *i*th rank group, the *A*-index is defined as:

$$A_i = \frac{1}{m} \sum_{j=i}^m \frac{1}{\sum_{k=1}^j c_k},$$

where *m* equals the total number of rank groups and *c_i* is the number of coauthors in the *i*th rank group with the same level of contribution. The *A*-index is thus bounded by 1. To assign the rankings to each author based on the respective level of contribution, we follow Stallings et al. (2013) and Biswal (2013) in assuming that the listing order of the authors implies the relative contribution; that is, we assume the last author to be the corresponding author who has the same level of contribution as the first author (both ranked first), while the ranks for the other coauthors are in increasing order based on their listing (decreasing level of contribution). Table 3 shows the *A*-index calculated under this assumption for up to ten coauthors, although the *A*-index captures only the individual contribution. Thus, to measure the contribution of each Nobel laureate-coauthor pair, we propose the following method to calculate the collaboration contribution for a co-author pair using the *A*-index of author *i* and *j*:

$$C_{ij} = (A_i + A_j) * \frac{A_i * A_j}{\left(\frac{A_i + A_j}{2}\right)^2},$$

where *C_{ij}* measures the co-contribution of author *i* and *j* with adjustment for the equity of the level of contribution between author *i* and *j*. The adjustment implies a larger

discounting factor for coauthor-pairs with higher inequity with respect to the level of contributions between authors i and j . Thus, the maximum value of C_{ij} is equal to the sum of A_i and A_j . We make this adjustment because we assume that the intensity of collaboration between the coauthor-pair who contributed equally is higher than pairs with unequal contribution given the same value of A_i plus A_j . Figures 5 and 6 show the weighted number of total collaborations between a laureate and a given coauthor assuming unequal and equal author contribution, respectively. The results resemble those in Fig. 2 where early collaborations (before age of 40) are more intense.

We choose arrival of new coauthors and intensity of collaboration to capture the dynamics of Nobel laureates' collaborations over their academic lifecycle because these reflect the social and academic norms in the respective fields. We assess the quality of such collaborations based on the number of citations received. For every laureate-coauthor pair that has published collaboratively in at least 4 distinct years, we calculate the average number of citations received by publications during first 2 years and last 2 years of collaboration. Figure 3 then plots the relationship between the two publication sets, with the average number of citations received by publications in first 2 years on the horizontal axis plotted against the average number of citations received by publications in last 2 years on the vertical axis (panel a). Panel b contains data restricted to laureate-coauthor pair that has published collaboratively in at least 7 distinct years. Data are plotted in the logarithmic scale. The red line represents the fitted values of a power law model between early and late citations ($y = ax^b$) and the green diagonal line indicates that late citations are equal to early citations (positively linear). The numbers of observations below and above the diagonal line are shown in the figure; the former (below the green) represents the number of coauthor-pairs where citations received by early publications are higher than citations

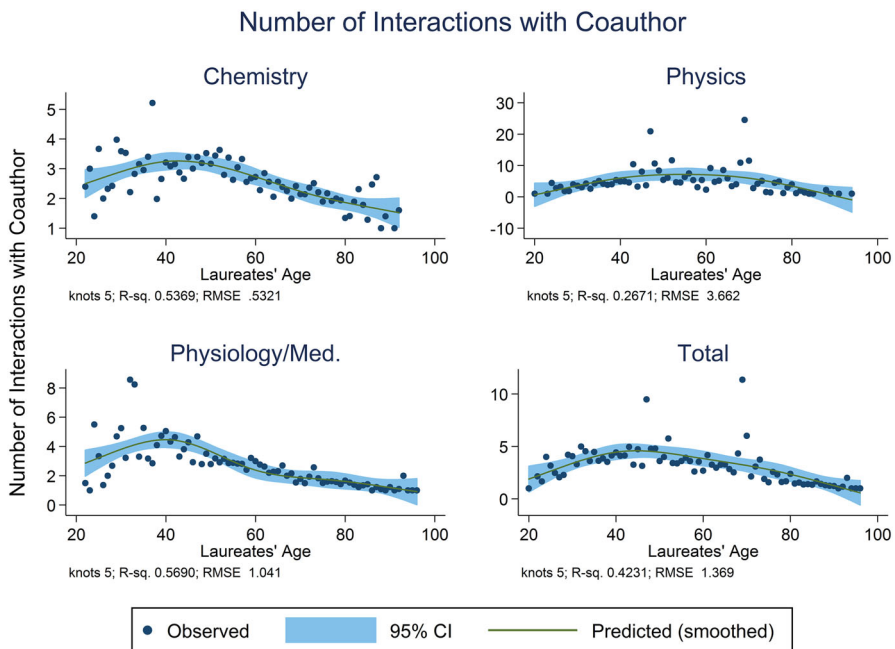


Fig. 2 Intensity of cooperation by field. *Note:* Smoothed values are computed using restricted cubic spline

for late publications, and vice versa for the latter. Results reveal that collaboration success is minimally dependent on pure luck: laureate and coauthor pairs that receive a high number of citations for their later publications are also those who receive a high number for their early publications (positive slope of the red line). Conversely, most collaborations that yield no highly cited publications early on tend to yield even fewer successful publications down the road.

The decay in citation success appears to be strongest in chemistry. The laureate co-author-pair ratio for early citations to late citations is equal to 1.245 (799/642, see panel a), which indicates that early collaborations are more successful. The ratio in physics and physiology/medicine are similar (1.184 and 1.229 respectively). It is clear from panel b (representing more long-term collaborations) that a greater number of observations lie

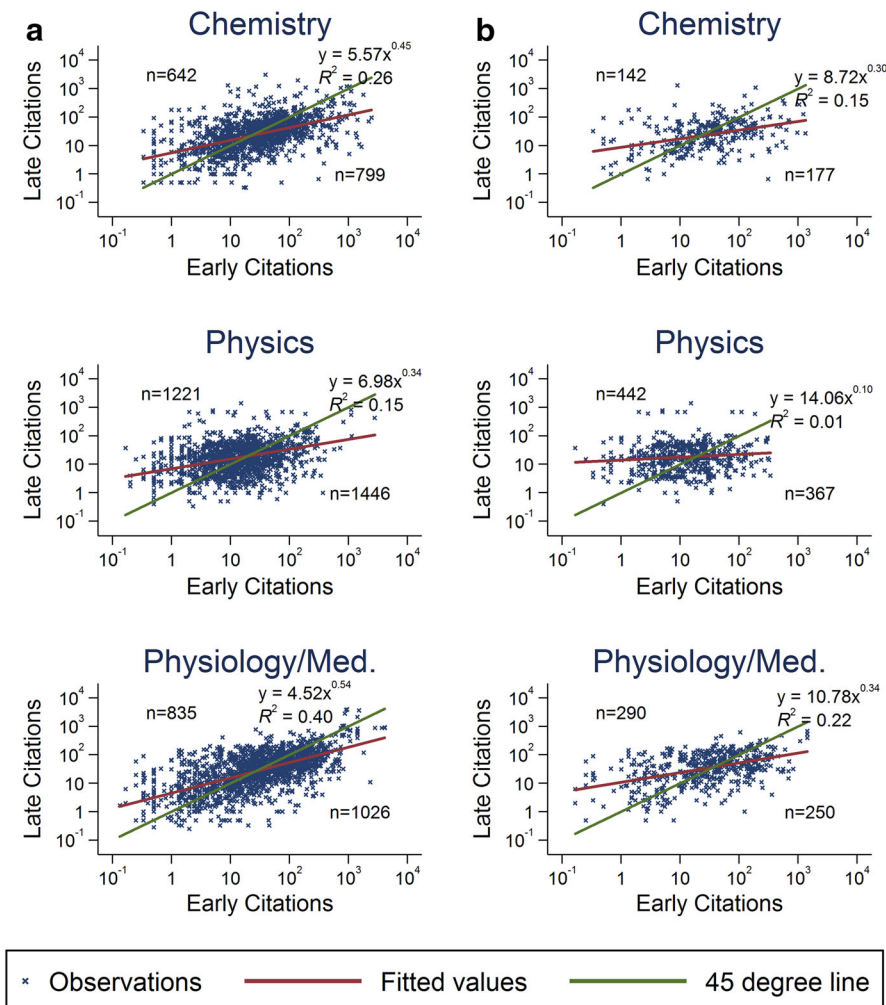


Fig. 3 Citations received by early and late collaborations of laureate-coauthor pairs. *Note:* The fitted values were obtained by linear least-square model, with the equation $\log_{10}(y) = a + b\log_{10}(x)$. Data are plotted in the logarithmic scale

below the diagonal line for chemistry (1.246) whereas more observations lie above the diagonal line in physics (ratio = 0.83) and physiology/medicine (ratio = 0.86), indicating that late publications are more successful. The differences in citation success in earlier versus later publications over the lifecycle of a given collaboration is greater in chemistry, perhaps because most chemistry research is done in a way to generate very specific data that are best published within a few high impact publications. Research in physics and physiology/medicine, on the other hand, generate rather more multidimensional data that sustain a large number of good ideas leading to several high impact publications, especially in highly complex research areas where experiments require a very costly setup.

The results for citations adjusted by collaboration contribution (citation counts multiplied by C_{ij}) are depicted in Figs. 7 and 8 (for unequal and equal contributions, respectively). While the positive correlation between citations received by early and late publications remains robust when accounting for collaboration contributions, the ratio is mostly above 1 (with the exception of physics), which indicates a greater citation success for early publications.

Our results are robust to our definition of early and late, and they hold when we define early and late interactions to cover all interactions that fall into the first half and the second half of the collaboration period, respectively. These results are reported in Fig. 9. For all disciplines, the ratio is above 1, indicating that the first period of collaboration is more successful than the second period. It is only for long-term collaborations (panel b) in physics that we observe the later period as more successful.

In order to investigate whether introduction of laureates who are still actively collaborating creates any bias in our analysis we differentiate between laureates who died before 2009 and those who are either still living or who died after 2009. The results are presented in Table 1, analyzing laureate-coauthor pairs that have published in at least 4 distinct years. We provide an overview of the ratio results, which (in line with our initial analysis) focus on raw citation counts, citations weighted for equal and unequal co-author contribution, and an alternative definition of early and late collaborations as in the previous paragraph. Overall, we can see that the ratio is mostly above one, indicating that early collaborations are more successful, which confirms the robustness of our initial results. The analysis of the deceased laureates further confirms the tendency in physics that later collaborations are more productive.

Two-stage estimation and discussion of results

In the first stage (see Table 4), to isolate the correlation between citations received for an article and the intensity of cooperation between that article's coauthors, we define journal quality as the journal's 2012 impact factor from the ISI Web of Knowledge 2012 Journal Citation Reports and regress this variable on paper characteristics in the first stage estimation to obtain prediction errors ($\hat{\mu}_{ij,h}$). In the second stage estimation, we regress citation count on the same explanatory variables as in the first stage but also on the predicted errors derived therein. The journal impact factor in the second step is thus the error obtained in the first, corresponding to the portion of journal impact factor not explained by the paper and collaboration characteristics. In this way, we separate the effects of journal quality on citation success from other explanatory variables.

The bases for these estimations are the following two specifications:

$$\text{Step 1 : } (Journal_Impact)_{ij,h} = f(\text{total collaboration, collaboration year, \#authors, laureate characteristics}) + \mu_{ij,h}$$

$$\text{Step 2 : } (Citations)_{ij,h} = f(\text{total collaboration, collaboration year, \#authors, laureate characteristics, } \hat{\mu}_{ij,h}) + \varepsilon_{ij,h}$$

We regress the journal impact factor for paper *h* of the laureate-coauthor pair *ij* on the total number of laureate (*i*) and coauthor (*j*) collaborations in our dataset (*total collaboration*), the year of appearance of that particular paper *h* in the life cycle of *ij* collaboration (in the first year, second year, or *n*th year of the collaboration), the total number of authors in publication *h*, and the Nobel laureate’s characteristics (field, age during publication, and individual fixed effects). To avoid collinearity between the total number of collaborations and the appearance number of a particular collaboration, we use indicator variables for various levels of total collaboration: 6–20, 21–40, 41–70, 71–110, and more than 110 (with between 1 and 5 as the reference group). Table 5 presents the descriptive statistics of the dependent and independent variables.

We focus on the marginal effects of repeated collaborations between laureate-coauthor pairs on citation success of their publications. In doing so, we must recognise that citations may be affected by the quality of the journal in which the article is published (e.g., due to increased visibility), or same variables affecting an article’s publication success may possibly be affecting also its citation success. Thus when citations are regressed on article’s

Table 1 Ratio of early to late citation success

	Living laureates or laureates deceased after 2009			Laureates deceased before 2009		
	<i>n</i> ₁	<i>n</i> ₂	Ratio	<i>n</i> ₁	<i>n</i> ₂	Ratio
Raw Citations counts (in line with Fig. 3)						
Chemistry	453	604	1.333	189	195	1.032
Physics	1063	1254	1.180	158	192	1.215
Phy./Med	700	899	1.284	135	127	0.941
Citations weighted for unequal coauthor contribution (in line with Fig. 7)						
Chemistry	386	671	1.738	172	212	1.233
Physics	1064	1253	1.178	209	141	0.675
Phy./Med	581	1018	1.752	123	139	1.130
Citations weighted for equal coauthor contribution (in line with Fig. 8)						
Chemistry	389	668	1.717	178	206	1.157
Physics	997	1320	1.324	145	205	1.414
Phy./Med	617	982	1.592	125	137	1.096
First and second half of whole publication set (in line with Fig. 9)						
Chemistry	394	663	1.683	181	203	1.122
Physics	1101	1216	1.104	215	135	0.628
Phy./Med	613	986	1.608	128	134	1.047

*n*₁ represents the number of observations with more citations in the later publications than in the earlier publications and *n*₂ represents the number of observations with more citations in the earlier publications than in the later ones. Ratio = *n*₂ divided by *n*₁

characteristics that include publishing journal quality (measured by impact factor), such quality will be highly correlated with other explanatory variables. This correlation could produce misleading outcomes because journal quality and citation of the article, rather than being independent, are determined by the same exogenous factors, including collaboration intensity. The citation success results show that the first four collaboration bins are all highly significant but negative (relative to the reference group of 5 or fewer collaborations), with only the fifth bin, the most extreme number of collaborations, being positive and insignificant (Table 2). Hence, all else being equal, and except for the extreme case of over 110 collaborations, the first cooperation sets tend to be more successful, leading to more citations per paper (between 16 and 48). Among laureates who won the prize while under 50, collaborations repeating more than 20 times have a positive and significant coefficient. For the laureates who won the prize after 50, the most successful papers are the early publications with the most intensive collaboration (over 110 repeated interactions). Most laureate-coauthor pairs collaborate over several years. The year (e.g., first, second, third ...) of the laureate-coauthor collaboration in which a particular publication occurs is captured by the variable *Collaboration Year* in Table 2. Square of the collaboration year is included to capture the non-linear productivity pattern over the life cycle of collaborations.

Table 2 Regression results for the 2SLS

	Overall	Age of the laureate		Chemistry	Physics	Phy./Med
		<50	≥50			
Collaborations 6–20	−15.65*** (2.215)	−11.95* (6.949)	−15.03*** (2.189)	21.42*** (4.424)	−13.70*** (2.329)	29.87*** (3.476)
Collaborations 21–40	−22.95*** (2.578)	35.80*** (8.654)	−20.45*** (2.611)	35.19*** (7.995)	−17.76*** (2.652)	82.50*** (5.024)
Collaborations 41–70	−45.76*** (3.133)	34.84*** (7.823)	−39.69*** (3.190)	42.37*** (11.007)	−20.99*** (2.771)	69.80*** (5.736)
Collaborations 71–110	−48.14*** (3.380)	43.74*** (9.877)	−42.66*** (3.415)	40.20*** (11.945)	−23.18*** (2.835)	92.54*** (7.378)
Collaborations >110	4.87 (3.173)	62.15*** (11.716)	4.15 (3.132)	26.97*** (7.636)	−1.68 (2.848)	76.61*** (6.531)
Collaboration year	−3.30*** (0.449)	−10.78*** (1.577)	−2.42*** (0.445)	−8.43*** (1.211)	−2.92*** (0.533)	−11.18*** (0.974)
Collaboration year ²	0.18*** (0.018)	0.34*** (0.058)	0.14*** (0.018)	0.28*** (0.041)	0.18*** (0.037)	0.38*** (0.037)
Journal quality	8.60*** (0.222)	9.04*** (0.478)	8.72*** (0.258)	11.70*** (0.682)	5.16*** (0.417)	3.62*** (0.175)
Number of authors	2.08*** (0.070)	6.25*** (0.325)	1.52*** (0.059)	−1.64 (1.184)	0.38*** (0.010)	22.33*** (0.511)
Laureate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	155,179	27,961	127,218	32,473	77,016	45,690
Adjusted R-squared	0.347	0.446	0.345	0.161	0.212	0.674

The table reports second stage coefficients where the dependent variable is an article's citation count. We run our analysis first for all Nobel laureates and then separately for laureates who won the prize while under or over the age of 50. We then conduct regressions for each field separately. Robust standard errors in parentheses

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

Long lasting collaborations are those that produce as good (or even better) cited publications during later years of collaboration as in the early years of it, and this is revealed by the non-linear marginal effect of the *collaboration year*. Non-linearity of citation success over the life cycle of a given collaboration captures an interesting relationship: although creativity and impact decays over the life cycle of many collaborations (most repeat over less than 4 distinct years), there are also some very long lasting collaborations that do not experience such a strong decay in productivity, hence the analysis should not be restricted to a strictly linear relationship between collaboration years and citation count.

Comparing our results for different fields, we find that although the total number of citations received by a paper in chemistry and medicine is strongly positively correlated with the total number of collaborations between the laureate and that particular coauthor, earlier papers in the collaboration sequence are expected to receive higher citations. In physics, on the other hand, total number of citations is strongly negatively correlated with total collaborations, except for collaborations that repeat more than 110 times, thus most citations are expected for papers from collaborations that repeat either less than 5 times or more than 110 times.

Our results suggest a “collaborative idea scarcity”, meaning that ideas that come early in the lifecycle of a collaboration between coauthors are on average the most innovative ones based on citation count. This further suggests that a collaboration may run out of creative ideas over time. What, then, are the most likely reasons for such a result? One explanation may be that the creativity of the original combination that generates new insights and breakthrough may emerge early rather than later during researchers’ collaboration. Likewise, efficient problem solving may emerge initially but become less relevant after success has been achieved. From then on, the pool of creative ideas seems to decrease. These views are somewhat supported by the evidence that success may be augmented by pairing high conventionality with novelty using atypical combinations (Uzzi et al. 2013) that themselves may be encouraged by novel interactions. It is also possible that highly innovative researchers such as Nobel laureates may be more critical of new collaborations and may only agree to those that seem to offer meritorious rigor. Moreover, receiving the Nobel Prize might have changed the perception of the laureates with existing or potential coauthors, and vice versa, hence changing the collaboration patterns and structure (Chan et al. 2015). On the other hand, collaborations may be chosen for reasons other than their effect on output or may, over time, transform into friendships, which decreases the pressure to collaborate productively (Hollis 2001). That is, cooperation can lead to an intellectual companionship that overcomes isolation, creating a personal relationship between the coauthors (Katz and Martin 1997). Thus, whereas a new collaboration can enhance diversity of perspective, a long-lasting collaboration may reduce diversity not only in perspective but also in expertise and experience.

It is also important to consider the type of research and the environment in which research is being produced. In-depth research of highly complicated topics requires the assembly of large research teams and may involve very high monetary costs due to specialized and highly technical equipment requirements. It is reasonable to expect that such highly complicated research will yield a continuous stream of data and several layers of complicated yet innovative and important results. Publication of such rich material may lead to a longer lifespan (in terms of publication count) of collaboration between researchers in that research team. Hence the collaboration lifespan depends on the complexity of research topic, however a reduction in diversity of ideas and creative perspective within the same research team is apparently no exception in this case as well.

Conclusion

One definite strength of new collaborations is that these are often characterised by a willingness to consider new ideas and/or adapt to novel approaches. In any collaboration—but particularly in science—trust is crucial to the sharing of ideas, models, data, or material of substantial scientific merit; and the scientific colleagues of Nobel laureates may be more willing to trust Nobelists in that regard. Hence, to benefit from the increasingly collaborative nature of scientific inquiry, researchers need a better understanding of what determines team success. The results reported here suggest that the advantages and costs of ongoing collaboration should be carefully weighed because, from a creativity viewpoint, collaborations have an expiration date, even for Nobel laureates.

Nobel laureates can be seen as (and they probably really are) researchers with evergreen research agenda and research ideas, and yet the impact of their collaboration with the same coauthor diminishes over the lifespan of such collaboration. This is an important lesson for all researchers: one should not underestimate the diminishing returns to collaboration due to stagnation and exhaustion. A crucial strategy for keeping one's research agenda evergreen is to keep one's coauthor pool evergreen.

Acknowledgments For advice and suggestions, thanks are due to two anonymous referees. We acknowledge financial support from the Australian Research Council (FT110100463).

Appendix

See Figs. 4, 5, 6, 7, 8, 9 and Tables 3, 4 and 5.

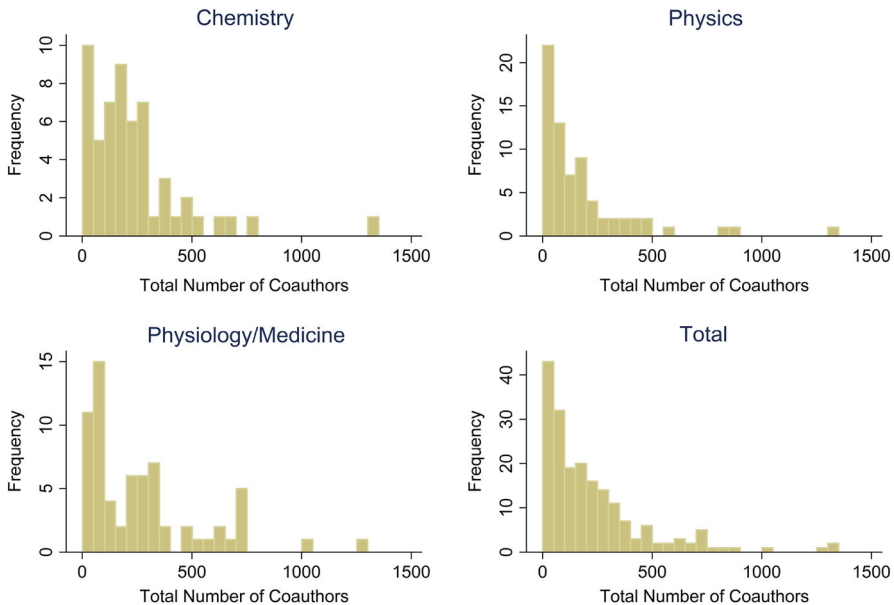


Fig. 4 Distribution of the total number of Nobel laureate coauthors. *Note:* Bin width = 50

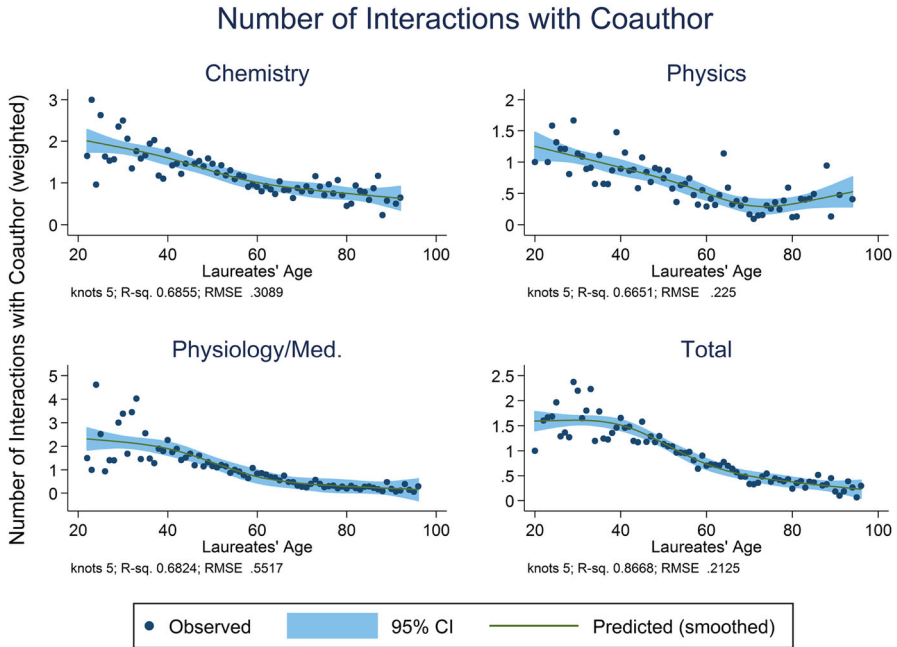


Fig. 5 Intensity of cooperation by field weighted by unequal co-author contribution

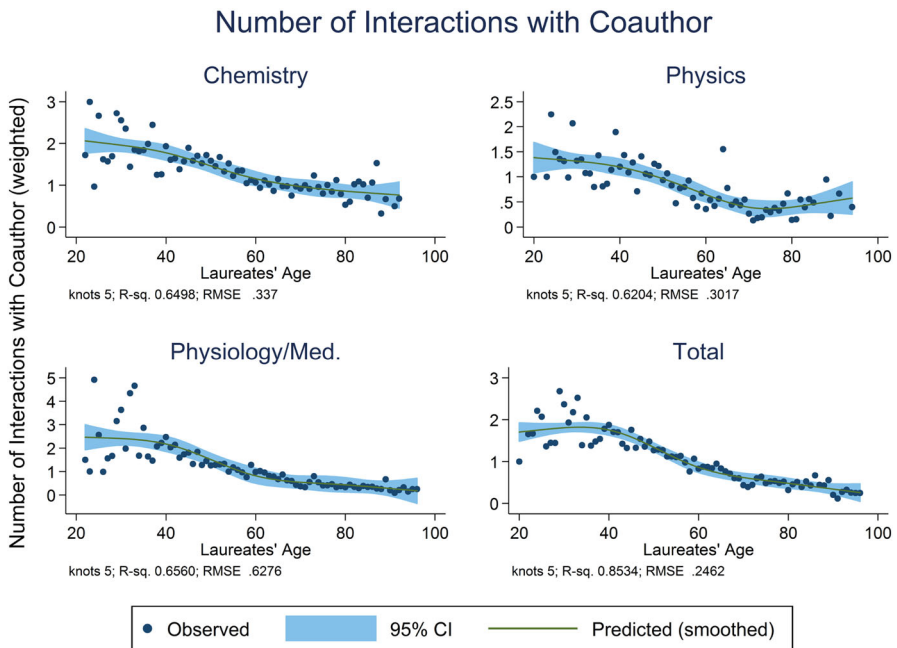


Fig. 6 Intensity of cooperation by field weighted by equal co-author contribution

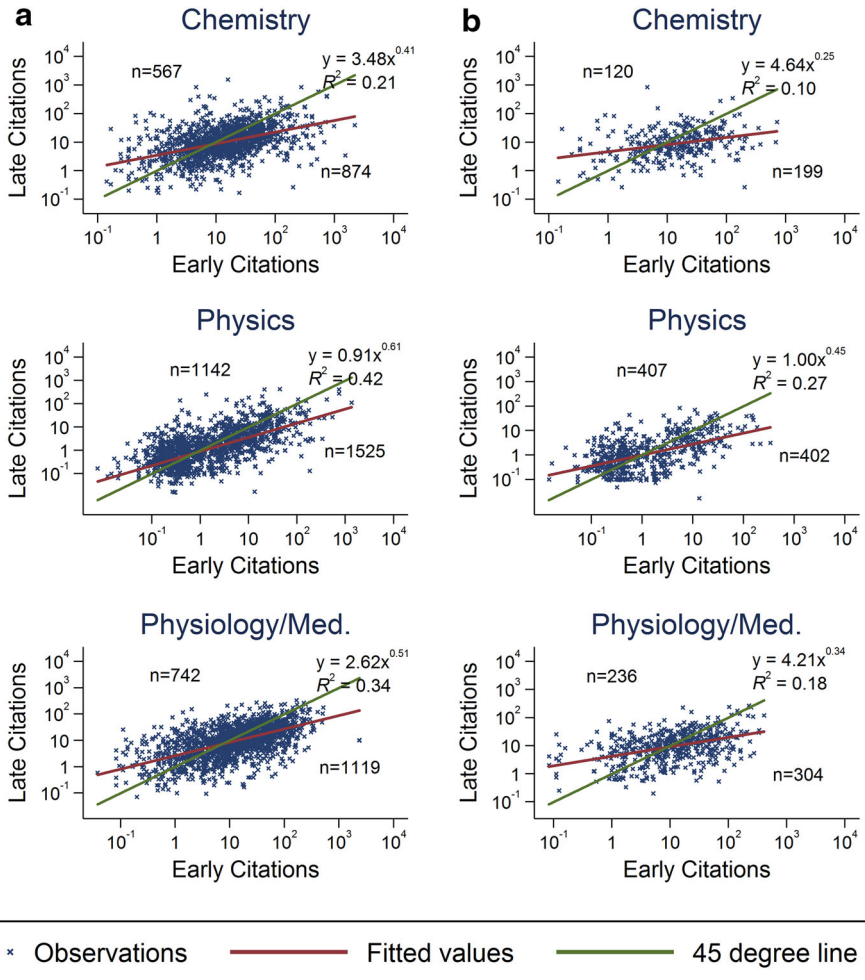


Fig. 7 Citations received by early and late collaborations of laureate-coauthor pairs weighted by unequal co-author contribution

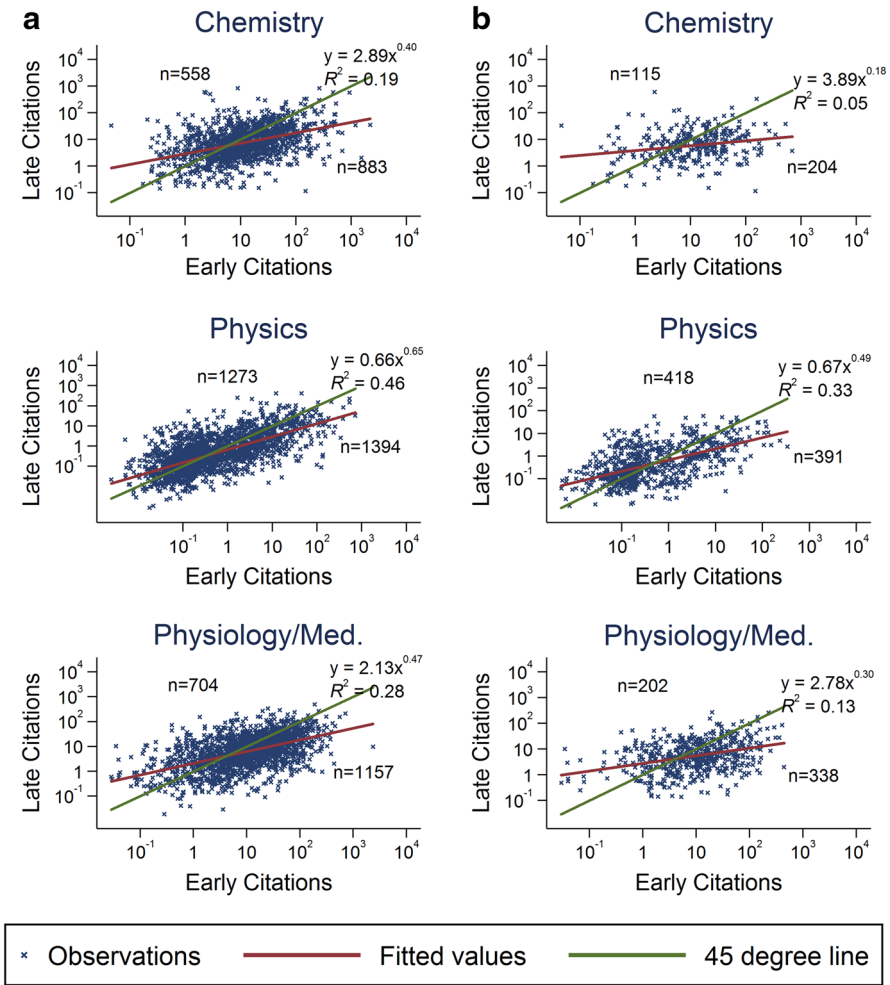


Fig. 8 Citations received by early and late collaborations of laureate-coauthor pairs weighted by equal co-author contribution

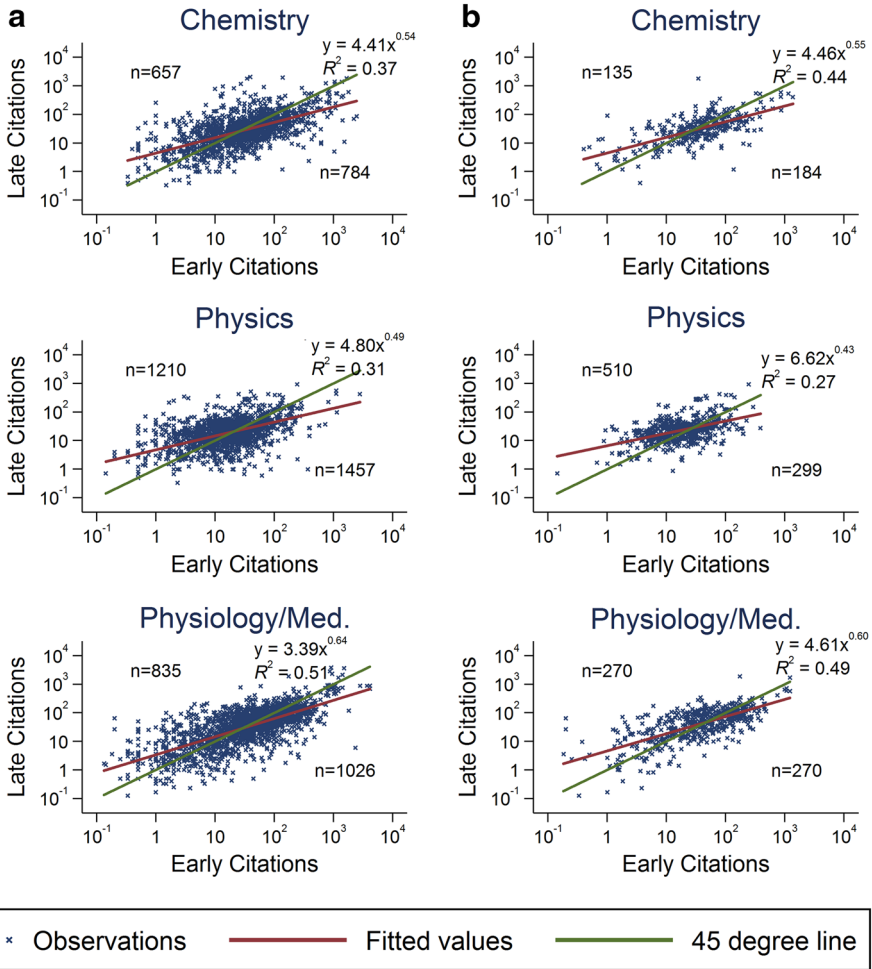


Fig. 9 Citations received by first and second half of all collaborations of laureate-coauthor pairs

Table 3 A-index for unequal co-author contributions

Coauthors' share	Author position									
	1	2	3	4	5	6	7	8	9	10
Number of authors										
1	1.000									
2	0.500	0.500								
3	0.417	0.167	0.417							
4	0.361	0.194	0.083	0.361						
5	0.321	0.196	0.113	0.050	0.321					
6	0.290	0.190	0.123	0.073	0.033	0.290				
7	0.265	0.182	0.127	0.085	0.052	0.024	0.265			

Table 3 continued

Coauthors' share	Author position									
	1	2	3	4	5	6	7	8	9	10
8	0.245	0.174	0.126	0.091	0.062	0.038	0.018	0.245		
9	0.229	0.166	0.124	0.093	0.068	0.047	0.030	0.014	0.229	
10	0.214	0.159	0.122	0.094	0.072	0.053	0.037	0.023	0.011	0.214

Table 4 First stage regression results for 2SLS

	Overall	Age of the laureate		Chemistry	Physics	Phy./Med
		<50	≥50			
Collaborations 6–20	0.15*** (0.055)	0.82*** (0.138)	−0.01 (0.060)	0.26*** (0.098)	0.29*** (0.055)	0.50*** (0.128)
Collaborations 21–40	0.30*** (0.065)	1.23*** (0.195)	0.16** (0.070)	0.72*** (0.192)	0.63*** (0.054)	0.11 (0.198)
Collaborations 41–70	0.34*** (0.075)	0.93*** (0.178)	0.28*** (0.082)	0.89*** (0.225)	0.62*** (0.056)	0.58** (0.277)
Collaborations 71–110	0.55*** (0.089)	1.34*** (0.208)	0.45*** (0.100)	0.67*** (0.257)	0.62*** (0.059)	2.15*** (0.359)
Collaborations >110	0.74*** (0.120)	1.49*** (0.301)	0.65*** (0.132)	1.23*** (0.245)	0.74*** (0.066)	1.16*** (0.319)
Collaboration year	−0.17*** (0.013)	−0.20*** (0.031)	−0.17*** (0.014)	−0.21*** (0.029)	−0.12*** (0.011)	−0.28*** (0.034)
Collaboration year ²	0.01*** (0.001)	0.01*** (0.001)	0.01*** (0.001)	0.01*** (0.001)	0.01*** (0.001)	0.01*** (0.001)
Number of authors	0.01*** (0.000)	0.02*** (0.002)	0.01*** (0.000)	0.03** (0.012)	0.01*** (0.000)	0.10*** (0.003)
Laureate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	155,179	27,961	127,218	32,473	77,016	45,690
Adjusted R-squared	0.248	0.390	0.184	0.121	0.303	0.195

First stage coefficients are being reported, where dependent variable is the impact factor of the journal where the article is published. Robust standard errors in parentheses

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

Table 5 Descriptive statistics of dependent and independent variables employed in 2SLS regression analysis

	Mean	SD	Min	Max
Citations received	79.47	395.41	0	8947
Journal quality (impact factor)	7.19	7.56	0.03	51.66
Collaborations 6–20	0.27	0.44	0	1
Collaborations 21–40	0.15	0.36	0	1
Collaborations 41–70	0.11	0.31	0	1
Collaborations 71–110	0.07	0.25	0	1
Collaborations >110	0.05	0.21	0	1
Collaboration year	3.57	3.75	1	36
Number of authors	50.71	64.70	2	181

References

- Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981–1999. *Research Policy*, *34*, 259–285.
- Biswal, A. K. (2013). An absolute index (*Ab-index*) to measure a researcher's useful contributions and productivity. *PLoS ONE*, *8*(12), e84334.
- Chan, H. F., Önder, A. S., & Torgler, B. (2015). Do Nobel laureates change their patterns of collaboration following the prize reception? *Scientometrics*, *105*(3), 2215–2235.
- Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, *52*(7), 558–569.
- Greene, M. (2007). The demise of the lone author. *Nature*, *450*, 1165.
- Guimera, R., Uzzi, B., Spiro, J., & Amaral, L. A. N. (2005). Team assembly mechanisms determine collaboration network structure and team performance. *Science*, *308*, 697–702.
- Hollis, A. (2001). Co-authorship and the output of academic economists. *Labour Economics*, *8*(4), 503–530.
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26*, 1–18.
- Milojević, S. (2014). Principles of scientific research team formation and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 3984–3989.
- Newman, M. E. J. (2004). Who is the best connected scientist? A study of scientific coauthorship networks. In E. Ben-Naim, H. Frauenfelder, & Z. Toroczkai (Eds.), *Complex networks* (pp. 337–370). Berlin: Springer.
- Stallings, J., Vance, E., Yang, J., Vannier, M. W., Liang, J., Pang, L., et al. (2013). Determining scientific impact using a collaboration index. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 9680–9685.
- Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*, 468–472.
- Whitfield, J. (2008). Group theory. *Nature*, *455*, 720–723.
- Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*, 1036–1039.
- Zuckerman, H. (1996). *Scientific Elite: Nobel laureates in the united states*. New Brunswick: Transaction Publishers.