Detrimental Collaborations in Creative Work: Evidence from Economics

Keyvan Vakili
London Business School
kvakili@london.edu

Florenta Teodoridis
USC Marshall School of Business
florenta.teodoridis@marshall.usc.edu

Michaël Bikard
INSEAD
michael.bikard@insead.edu

Abstract: Prior research on collaboration and creativity has mostly assumed that individuals choose to collaborate because collaboration positively contributes to output quality. In this paper, we argue that collaboration conceals individual contributions, and that the presence of a collaboration credit premium—when the sum of fractional credits allocated to each collaborator exceeds 100%—might motivate individuals to collaborate even when their collaboration hurts output quality. We test our argument on a sample of economists in academia. To estimate the causal effect of collaboration, we take advantage of the norm of alphabetical ordering of authors on scientific articles published in economics journals. This norm means that economists whose family name begins with a letter from the beginning of the alphabet receive systematically more credit for collaborative work than economists whose family name begins with a letter from the end of the alphabet. Using this systematic difference as an instrument for collaborative behavior, we show that economists sometimes choose to collaborate even in cases where this choice decreases output quality. Collaboration can therefore create a misalignment between the incentives of creative workers and the prospects of the project.

INTRODUCTION

Collaboration has become ubiquitous in creative work, but the reasons why individuals collaborate remain poorly understood. In particular, a common assumption in prior literature is that individuals choose to collaborate because collaboration contributes positively to output quality (Wuchty et al. 2007, Singh and Fleming 2010). However, in deciding whether to collaborate, creative workers consider not only output quality but also the amount of credit that they will receive (Merton 1968). And, the two objectives might not always align.

The goals of maximizing output quality, on one hand, and individual credit, on the other, could diverge if credit is not tightly linked to contribution. When individuals work alone, their contribution is observable and credit allocation is straightforward. However, collaborative work obfuscates individual contributions, potentially leading to the emergence of a collaboration credit premium—that is, a situation in which the share of credit for collaborative output sums up to more than 100% (Bikard, Murray, and Gans 2015; Kay, Proudfoot, and Larrick 2018). In the presence of a credit premium, each collaborator, on average, receives credit that is greater than their contribution.

Building on this core insight, we posit that individuals may sometimes rationally choose to collaborate on a project even when collaboration negatively affects the quality of the output. Assuming that creative workers are primarily motivated by maximizing their share of individual credit, they are likely to seek collaboration even where it might not be optimal for the project, as long as the expected benefits from the collaboration credit premium more than compensate for the project-level loss in output quality. Collaboration can, therefore, cause a decoupling between individual incentives to maximize credit and the goal of maximizing output quality.

Testing this argument empirically is challenging, however. Individuals do not randomly choose to collaborate. It is therefore difficult to estimate how much of the association between collaboration and output quality might be driven by selection versus treatment mechanisms. Moreover, we are interested in estimating the causal impact of collaborations for a subset of collaborations only: those formed specifically as a result of the pull exerted by the collaboration credit premium.

We address these challenges in a sample of academic scientists in the field of economics. Specifically, we exploit a collaboration norm among researchers in this field: the alphabetical ordering of authors' names on publications. Because of this norm, economists whose family name begins with a letter toward the end of the alphabet engage less in collaboration because they receive less credit for their collaborative publications. Indeed, these individuals are less likely to be promoted in their first academic job or to win prestigious awards than individuals whose family name begins with a letter toward the beginning of the alphabet (Einav and Yariv 2006). This norm essentially creates a variance in the gap between contribution and credit for collaborative work, and hence in individuals' collaborative behavior, such that individuals with family names toward the end of the alphabet collaborate less frequently than those with family names toward the beginning of the alphabet. We use the alphabetical order of individuals' family name as an instrument for propensity to collaborate to capture the causal impact of collaboration on output quality. Importantly, the instrument captures the variance in propensity to collaborate that is driven by the variance in the level of credit premium that individuals receive from their collaborations.

Our results first confirm that alphabetical ordering has a sizable effect on economists' propensity to collaborate. Individuals whose family names begin with letters toward the end of the alphabet are significantly less likely to collaborate on their publications than their colleagues whose family names begin with letters toward the beginning of the alphabet. Next, using the alphabetical rank of individuals' family name as an instrument for propensity to collaborate, we show that collaboration has a significant negative effect on the quality of certain publications in our sample. Our findings suggest that collaborations formed because of their associated credit premium have an average negative effect on output quality.

Our study makes two main contributions. First, we investigate a previously implicit assumption in the literature on collaboration and creativity. We show that what is good for individual workers is not necessarily good for their projects, and vice versa. Second, our results highlight a negative aspect of collaboration in creative work that has been omitted in previous studies on this topic. Collaboration can be challenging, not only because it induces free-riding, coordination costs, conflict, or groupthink, but also because it can create a decoupling of the goals of project owners from those of creative workers. Researchers and practitioners who ignore these dynamics might have an overly optimistic view of the impact of collaboration on creative performance.

COLLABORATION AND THE CREDIT PREMIUM

Collaboration as a Choice

In many organizations, including in academia, individuals involved in creative tasks are free to decide whether to collaborate or not. This decision involves a number of important trade-offs (Taylor and Greve 2006; Bikard, Murray, and Gans 2015; Deichmann and Jensen 2018).

The benefits of collaboration are considerable. By bringing together individuals with different knowledge, opinions, skills, and resources, collaboration enables cross-fertilization and recombination of ideas (Hargadon and Sutton 1997; Toh and Polidoro 2013; Choudhury and Haas 2018). It also enables specialization and a productive division of labor whereby individuals focus on what they can do best, thereby potentially increasing productivity for all (Jones 2009). Finally, collaboration can provide additional advantages such as learning (Dodgson 1993), making it easier to identify and filter out bad ideas (Singh and Fleming 2010), increasing the number of communication channels and the legitimacy for new ideas (Reagans and Zuckerman 2001; Cattani and Ferriani 2008), and providing a safe space to discuss original thoughts (Edmondson 1999).

There are also costs associated with collaboration. A large stream of work in psychology has shown that collaboration can lead to social loafing and free-riding (Latané, Williams, and Harkins

1979; Albanese and Van Fleet 1985). Moreover, an increase in the number of individuals involved in a project naturally increases the number of communication linkages among individuals nonlinearly. This, in turn, might increase the amount of time and effort it takes to keep all members informed of the project's status and expand the variety of interpretations based on the same information. As a consequence, the time spent on the project might increase to accommodate the need to reconcile and integrate information at the team level (Lawrence and Lorsch 1967; Dougherty 1992; Perlow 1999; Heath and Staudenmayer 2000; Porac et al. 2004; Cummings and Kiesler 2007). Collaboration also introduces the possibility of conflict among collaborators, which can potentially harm team productivity (O'Dell 1968; Brewer and Kramer 1986; Cattani et al. 2013; Goncalo et al. 2015). Finally, several studies on groupthink show that collaborators on a project may engage in dysfunctional or irrational decision-making to minimize conflicts and maintain coherence in the team (I. L Janis 1972).

In balancing these benefits and drawbacks, creative workers are likely to consider not only what is best from a project standpoint but also what is in their personal interest. This is important because collaborators are not always rewarded equally for joint work. Merton (1968), for example, argues that, independent of their contribution, high-status individuals on scientific teams are likely to accrue more credit for their paper than their lesser-known co-authors. Conversely, lower-status individuals or those belonging to minority populations might receive relatively low recognition even if their actual contribution is considerable (e.g., Heilman and Haynes 2005). Thus, individual recognition for collaborative work does not always reflect one's effective contribution. The net benefit from collaboration appears very different for different people—even those working on the same project.

Considering these complex trade-offs, one might wonder whether the benefits of collaboration exceed the costs on average. The growth in collaboration among creative workers (de Solla Price 1965; Wuchty, Jones, and Uzzi 2007; Leahey 2016) might be interpreted as evidence of a net positive effect. Wuchty, Jones, and Uzzi (2007) examine 19.9 million scientific publications over 5 decades as well as 2.1 million patents and document a large-scale shift toward collective work in a vast array of fields ranging from the social sciences to engineering, including patenting. Furthermore, collaborative work seems to be of higher quality, on average, than solo work (see also Singh and Fleming 2010). These results indicate that creative workers increasingly find value in collaboration. However, these findings do not necessarily mean that collaboration always has a positive impact on output quality. In this paper, we argue that collaboration can sometimes benefit collaborators while hurting output quality.

Individuals versus Projects

Although prior studies have explored the implications of collaboration separately from a project and from an individual standpoint, they have not examined the relationship between the two. This omission would be inconsequential if the two goals were always aligned. However, it could be significant if collaboration can be simultaneously good for individuals and bad for the project, or vice versa. In this scenario, the decoupling between individual- and project-level incentives might lead individuals to choose to collaborate (or to work alone) when it is not optimal from a project standpoint.

In this paper, we posit that individual- and project-level incentives in creative work are not always aligned. We build our argument on three premises. First, credit received for collaborative work is generally disconnected from actual contribution. This results from the fact that each person's contribution is difficult to evaluate. In reality, collaborators themselves—who can presumably observe their own contribution—misjudge its significance systematically. A large body of work in psychology has shown that individuals tend to overestimate their contribution to collaborative work in a number of settings, from married couples sharing housework to co-workers in firms and athletes on sports teams (e.g., Ross and Sicoly 1979; Schroeder, Caruso, and Epley 2016). If individuals themselves routinely misjudge their own contributions, evaluations from outsiders are likely to be particularly inexact, since they generally have much less information about each person's work.

Second, the individual shares of credit allocated for collective work can sum to more than 100%. Kay, Proudfoot, and Larrick (2018) show that observers systematically overestimate the creative skills of individuals in firms when their teammates are not visible. In line with this finding, Bikard, Murray, and Gans (2015) study collaboration choices and credit allocation among MIT scientists and find evidence for the existence of a collaboration credit premium—that scientists receive on average more than the fractional share of credit for their joint work (see also Freeman, Ganguli, and Murciano-Goroff (2014)). Taken together, these studies suggest that incentives to collaborate might be particularly strong especially when different collaborators are evaluated by different people who might not know every team member.

Third, the choice of whether to collaborate depends on the amount of credit individuals expect from the audience. Clearly, taste for collaboration varies from one person to the next, and so do individual skills at working jointly with other people. On average, however, we assume that creative workers will decide to collaborate more when they anticipate receiving more credit for joint work. Note that this assumption is consistent with prior studies not only in science (Bikard, Murray, and Gans 2015) but also in firms (Lee and Puranam 2017; Deichmann and Jensen 2018).

These three premises suggest that individuals might sometimes collaborate more than would be desirable from a project standpoint. When individuals know that they will be rewarded above and beyond their fair contribution to a project, they are likely to seek collaboration even where the costs of joint work exceed the benefits to the project. In other words, individuals might decide to collaborate even in cases where collaboration worsens the output quality.

To illustrate how such detrimental collaborations might occur, consider a situation where two R&D workers in a biotech company face a choice of whether to work independently or together. The two will be rewarded for their work by their respective managers, who do not know the other person. For the sake of argument, let's assume that both employees are equally capable but that coordination costs and disagreements between the two mean that, after controlling for the effort, the jointly produced work is of lower quality than the work they could accomplish separately. Thus, from a project standpoint, it is desirable that they not collaborate. However, if each person knows that they will receive more than 50% of the credit for the joint work, they might still decide to work together as long as the collaboration credit premium compensates for the expected loss in output quality.

Note that we do not suggest that all collaborations have a detrimental effect on output quality. Rather, we argue that certain collaborations with a negative effect on output quality might occur because individuals could get more credit for their work when they collaborate than when they work alone.

METHODS

Empirical Strategy

We test our predictions by examining the collaborative choices of economists in academia. This context provides several important advantages. First, we can observe both the evolution of collaborative choices and the quality of collaborative output using publication data. The list of authors on each paper helps us identify the set of collaborators on each project. The number of citations to each publication provides a convenient—though imperfect—proxy for output quality. Both measures have been used extensively in prior studies. Assuming academics care primarily about maximizing their credit on the portfolio of their publications, it is then possible to impute the level of credit premium by observing the combination of their collaboration choices and the quality of their publications. This is at the core of the method used by Bikard et al. (2015) to estimate the level of collaboration credit premium among a population of MIT science and engineering faculty members. We use a similar approach, explained in more detail below, to confirm the presence of a credit premium in our sample as a prerequisite to testing our prediction.

Second, in the field of economics, the strong norm of alphabetical ordering of the authors on publications enables us to estimate the causal impact of collaborations formed as a result of credit premium on output quality. Estimating the causal effect of collaboration is rife with endogeneity challenges. For example, a positive association between collaboration and output quality might stem from the fact that promising projects attract more collaborators than less-promising ones. Also, managers in R&D departments might ask individuals to work alone on incremental tasks but to work in teams on more impactful projects (Toh and Polidoro 2013). These selection processes complicate the task of the empiricist.

To address these concerns, we use an instrumental variable method. This empirical strategy targets the endogeneity issue by engaging an instrument variable that is correlated with the resulting project quality but only through its impact on the propensity to collaborate. Our choice of instrument is directly informed by past research showing that economists whose family names begin with letters toward the end of the alphabet receive less credit for their collaborative papers than their peers whose family names begin with letters toward the beginning of the alphabet. Therefore, we expect that researchers with less-favorable alphabetical ranks will collaborate less frequently at the margin. Put differently, individuals with more-favorable alphabetical ranks may engage in more collaborations because of the higher levels of credit premium they experience when collaborating.

Using the alphabetical rank of individuals as an instrumental variable presents two important advantages. First, it satisfies the three conditions for a valid instrument: (1) alphabetical ranking of last names has a causal effect on economists' propensity to engage in collaboration (i.e., the first stage), (2) the alphabetical rank is unlikely to have an effect on the quality of the work through any channels other than individuals' propensity to collaborate (i.e., the exclusion restriction), and (3) the alphabetical rank is as good as randomly assigned relative to any endogeneity concern that might confound the casual impact of collaboration on project quality (i.e., the independence assumption). In other words, it is reasonable to assume that people do not select their last names in expectation of benefiting from a career as researchers in economics.

Second, as we show below, individuals' alphabetical rank influences the level of collaboration credit premium that economists experience. This is important because our hypothesis is concerned with capturing the effect of collaboration in those cases where collaboration is driven by the variance in the level of collaboration credit premium.

While the field of economics in academia meets all the criteria for testing our theory, our empirical focus on this setting does not mean that detrimental collaborations driven by a collaboration credit premium do not occur in other settings. The decoupling between contribution and credit is inevitable in any creative context such as firm R&D, artistic production, or consulting. In these contexts, individual contributions are hard to measure. The creative process is usually iterative, and the value of one's input may not necessarily be correlated with the amount of effort one exerts. Moreover, in many nonacademic settings, just as in academia, collaborators are evaluated by different audiences, increasing the chance that evaluators do not know all the team members, therefore leading to the emergence of a collaboration credit premium (Bikard, Murray, and Gans 2015; Kay, Proudfoot, and Larrick 2018). For example, R&D teams may include individuals from different divisions or regional offices that are evaluated by their respective managers. Collaborating artists often receive individual reputational awards from dispersed audiences. Even collaborating firms can be recognized and rewarded differently by their stakeholders for the same output. In these and similar contexts, the independent evaluation of collaborators based on their collaborative output while their contribution is not observable can give rise to a collaboration credit premium and, hence, create incentives to engage in detrimental collaborations.

Sample

We build our sample by first collecting the full list of 21,905 PhD graduates from 12 top economics departments in the United States between 1990 and 2015 from the ProQuest Dissertation & Theses database. We then construct the career history of each graduate since graduation using online resources such as university websites, LinkedIn, and company pages. Next, we exclude all individuals who did not pursue a career in academia. For each remaining individual, we specifically focus on the period before tenure. We do so because the amount of credit attributed to pre-tenure academics by their audience is an important input in the tenure decision, one of the most important career steps in academia. In contrast, we expect tenured faculty to be relatively less concerned with credit given their lifetime appointment. Note also that a sample of tenured faculty is also likely to suffer from selection bias because collaboration choices affect whether individuals receive tenure. In other words, we are concerned that a sample that includes tenured individuals would capture researchers who advanced in their career because they took advantage of the credit premium phenomenon. The presence of this group would downward-bias the effects we are interested in evaluating.

Our final sample includes 1,164 pre-tenure economists in academia. Next, we extract the full record of each individual's publication portfolio from Scopus, a comprehensive bibliographic database maintained by Elsevier that documents worldwide academic publications across a variety of domains. We also collect key information about the institutions and departments at which these individuals worked during each year in our observation window, since the beginning of their PhD program.

Figure 1 shows the distribution of individuals across the institutions from which they graduated and the years of their graduation, the distribution of their publication rates per year, the alphabetical distribution of their family names, and the distribution of the number of authors on all papers in our sample. Table 1 shows summary statistics for the key variables in our sample. On average, individuals in our sample graduated from their PhD program in 2004; published 0.42 papers per year, of which 61% involved collaborators; and received 14 citations per year. Twenty-five percent of the individuals in our sample are women. Note that while the majority of researchers in our sample work in an economics department, some work in departments that cover a mix of economics and other fields such as health economics, law and economics, and economics and policy. In our empirical analysis, we control for the type of department in which each researcher works.

—Figure 1 and Table 1 about here—

Empirical Approach

We test our hypothesis in two steps. To establish the basis for our instrumental variable analysis, we first provide evidence for the presence of a credit premium in our sample and confirm that individuals whose family names begin with a letter toward the beginning of the alphabet experience higher levels of credit premium than their peers whose family names begin with letters toward the end of the alphabet. Next, using the variation in credit premium induced by the norm of alphabetical ordering as an instrument, we estimate the impact of credit-driven collaborations on the quality of output.

Evidence of Collaboration Credit Premium

It is difficult to estimate the level of credit allocation for each instance of collaboration in our sample, since the individual contribution of collaborators and the amount of credit attributed to each by their respective audiences are unobservable. To circumvent these challenges, we use an indirect method of estimating the collaboration credit premium developed by Bikard et al. (2015). The method assumes that (1) researchers, on average, choose a level of collaboration that is expected to maximize their allocated credit; and (2) the credit allocated to each author on a paper amounts to a fraction of the citations to that paper. The latter assumption simply suggests that the credit for the paper must be shared among collaborators, albeit in a way that might sum up to more or less than 100%.

The intuition behind the method is that individuals will increase their collaboration rate if they think that such an increase will increase their fractional credit. Similarly, they will reduce their collaboration rate if they believe that doing so will increase the fractional credit allocated to them for their papers. Therefore, at the population level, we should see neither a positive nor a negative relationship between the average fractional credit allocated to individuals and their collaboration rate. Any significant positive or negative relationship suggests that individuals can still increase their allocated credit by increasing or decreasing their collaboration rates. Based on this idea, one can examine a range of hypothetical levels of credit premium and identify the latent level of credit premium for which the relationship between the fractional credit and the collaboration rate becomes zero. The online appendix explains the method in more detail and provides all the estimations conducted to uncover the level of credit premium in our sample.

Our estimates suggest that for a team of two authors on a paper in our sample, each author receives on average 79% of the citations to the paper as individual credit, 29 more percentage points of credit than in the no-premium scenario.[1] The magnitude of the effect is slightly larger than that reported in Bikard et al. (2015) for their sample of MIT science and engineering faculty members.

Our instrumental variable method relies on the assumption that individuals whose family names begin with letters from the beginning of the alphabet should experience a higher level of collaboration credit premium than their peers whose family names begin with letters from the end of the alphabet. To test this assumption, we repeat our estimation of credit premium separately for individuals whose family names begin with letters toward the first half of the alphabet (A to M) and for those whose family names begin with letters toward the second half of the alphabet (N to Z) (see the online appendix for more details). Our estimations suggest that the former group receives on average 82% of the citations to their papers as individual credit, whereas the latter group receives only about 68% of the citations to their papers, a markedly lower level of credit premium.

RESULTS

We test whether the presence of a credit premium can lead to the formation of detrimental collaborations: that is, collaborations that might maximize the credit allocated to collaborators but harm the quality of their output. To demonstrate the necessity of employing our instrumental variable technique, we begin by replicating the effect documented in past research that has shown a positive

[1] Note that the level of collaboration credit premium varies with the number of collaborators. In the appendix, we estimate the level of credit premium as a function of the number of collaborators on each paper. The reported level of credit premium here is for the median number of collaborators in our sample, which is two.

relationship between collaboration and output quality (e.g., Wuchty, Jones, and Uzzi 2007). These results are based on correlational studies and hence do not separate selection and treatment effects. We similarly find a significant positive relationship between collaboration and output quality in our sample (Table 2).

—Table 2 about here—

Next, we repeat this estimation using our instrumental variable strategy. We use a two-stage least-squares model (2SLS) procedure at the individual-year level, with standard errors clustered at both institution and year levels.

In the first stage, we estimate the impact of the alphabetical rank of family name on the number of collaborative papers individuals produce in a given year. Based on our estimations in the previous section, we expect individuals with more-favorable alphabetical ranks to engage in more collaborations due to the higher levels of credit premium they receive.

In the second stage, controlling for the number of papers produced, we test whether an increase in the number of collaborative papers in a year increases or decreases the total number of citations to an individual's papers in that year. We construct our dependent variable as the log of citations plus 1 to account for the skewed distribution of the variable. We control for the individual's number of papers produced each year to account for the potential effect of that individual's collaboration on productivity independent of its effect on output quality. A positive effect of collaboration on citations suggests that collaboration undertaken because of a credit premium, on average, increases the quality of a paper—that is, the number of citations to the paper. In contrast, a negative effect suggests that collaboration undertaken in expectation of a credit premium, on average, decreases quality.

Because our instrument—alphabetical rank—is fixed at the individual level, we cannot include individual fixed effects in our estimations. Instead, we control for a wide range of time-variant and time-invariant characteristics of individuals in all regressions. In particular, we control for the total number of papers produced by each individual in a given year to ensure that the impact of collaboration on citations is not driven by changes in productivity. Further, researchers with stronger records may care less about the credit attributed to them, and hence their collaboration choice may be less sensitive to their alphabetical rank. Therefore, we control for researcher's publication record by including the cumulative number of papers that each researcher had produced and the cumulative number of citations to those papers by $t - 1$. Moreover, research suggests that women receive less credit for their work than men on average (Sarsons 2017). Thus, we also control for researcher gender. We further control for starting and finishing years of individuals' PhD program, the institution from which they graduated, the institution at which they worked at time *t*, and the type of department at which they worked at time *t*, to ensure that the comparison is between individuals from the same cohort facing similar institutional norms and expectations. One might be also concerned that the distribution of alphabetical rank is systematically correlated with individuals' ethnicity, which may positively or negatively affect their propensity to collaborate or the quality of their work. We confirm that Asian family names in our sample are more likely to begin with letters from the second half of the alphabet. Thus, we also control for whether an individual has an Asian family name or not. We also control for the logged (1 plus the) number of papers for which authors are not listed alphabetically for each individual-year observation. Such papers may signal collaboration with non-economists or publication in non-economics outlets, both of which may affect citation outcomes for the researcher.

Table 3 shows our first-stage results. Consistent with our expectation, the alphabetical rank of individuals has a statistically significant and economically meaningful effect on their choice of collaboration. The estimates suggest that one drop in alphabetical rank is equivalent to a 0.002-percentage-point drop in the yearly number of collaborative papers. Given that the average number of collaborative papers per year is 0.24, the estimate suggests that a one-standard-deviation decrease in one's alphabetical rank leads to an approximately 10% decline in one's propensity to collaborate. Because we cluster the standard errors, the first-stage F-statistic does not provide a reliable source for testing the weakness of our instrument. Instead, we report the Kleibergen-Paap Wald F-statistic.

Comparing our Kleibergen-Paap Wald F-statistic of 27.349 with the critical values of the Stock-Yogo weak identification test suggests that the relative bias of our estimation is smaller than 5% for a 5%-level Wald test (Stock and Yogo 2005).

—Table 3 about here—

Table 4 shows the results for the second stage. In line with our hypothesis, the estimates suggest that the number of collaborative papers in a year has a significant negative effect on the number of citations received by an author in that year. More specifically, switching from a solo paper to a collaborative paper, motivated by the collaboration credit premium, leads to an average 63% decline in the number of citations per year. Given that authors receive 14 citations per year on average, the effect is equivalent to a loss of approximately nine citations per year due to collaborations that were exogenously driven by the variance in credit premium allocated to individuals. The 95% confidence interval for the effect size denotes a range of 20% to 83% decline in the number of yearly citations.

—Table 4 about here—

These results confirm our prediction that a collaboration credit premium can lead to the formation of collaborations that hurt output quality. Note that we cannot generalize our estimated effect to the whole sample or more broadly to the field of economics. In other words, we do not (and cannot) claim that all collaborations in the field of economics have a negative average effect on output quality. Rather, consistent with our hypothesis, our results show that the credit-premium-driven increase in collaboration levels in our sample has a negative treatment effect on output quality.

As robustness, in Table A2 in the appendix, we report the first-stage results for the tenured researchers in our sample. As expected, the estimated effect of alphabetical rank is positive, close to zero, and nonsignificant at the 10% level (p-value = 0.415). This suggests that tenured faculty in our sample are less sensitive to the skewed allocation of a credit premium in collaborative efforts. In Tables A3 and A4, we report the second-stage results based on a two-stage generalized method of moments

(GMM) estimation and on a limited-information maximum likelihood method. All results remain robust to the effect reported in our main estimation models.

DISCUSSION AND CONCLUSION

Creative workers in firms and in academia can often choose whether to collaborate or to work alone. This choice is an important one, not only because it influences the quality of the work output but also because it affects the amount of credit that creative workers receive. Prior research has implicitly assumed that these two goals, project quality and personal credit, are aligned—in other words, that what is good from a project standpoint is also good for individuals, and vice versa. We challenge this assumption and argue that collaborative work can lead to a decoupling of individual- and project-level incentives.

Our argument rests on the idea that credit allocation in collaborative work is difficult because individual contributions are generally unobservable. Team members might share credit in a way that sums up to more than 100%, especially where individuals are evaluated by different audiences (Bikard et al. 2015; Kay et al. 2018). This collaboration credit premium is likely to create a strong pull toward collaboration. Under these circumstances, individuals might choose to collaborate even if it effectively reduces output quality, as long as the credit premium more than offsets the negative impact of collaboration on quality.

We hypothesized and found empirical support for the proposition that, to gain extra credit, researchers in economics decide to collaborate even when collaboration is detrimental to output quality. We focused on economics because its norm of alphabetical ordering provides an opportunity to address the usual endogeneity concerns with estimating the causal effect of collaboration on the quality of output. However, our theoretical claims apply to other creative settings where individual contributions to collaborative work are difficult to observe, where individuals can choose whether or not to collaborate, and where one can expect the existence of a collaboration credit premium.

While the case of economists in academia provides a useful setting in which to test our theoretical predictions, some limitations remain. First, although economics is a prominent academic discipline, we study only one setting and are unable to assess how the phenomenon of detrimental collaboration varies from one setting to the next. In particular, it is possible that economists react more strongly to credit-related incentives than individuals in other creative fields. After all, economists are known to be more sensitive to economic incentives than the rest of the population (Marwell and Ames 1981; Carter and Irons 1991).

Second, our study is limited by the fact that our measures of credit allocation are indirect. Our empirical strategy relies on individuals at the end of the alphabet differing from people from the beginning of the alphabet by the fact that they receive less credit for collaborative work but are being otherwise comparable in other dimensions (at least, those that are not orthogonal to output quality). Yet it is possible that those two groups differ in ways that are unobservable to us. For example, one could imagine that economists toward the end of the alphabet are more likely to leave the field since they know that they are being discriminated against. We investigated this potential selection process by comparing the distribution of alphabetical ordering at the PhD and at the faculty levels and found no meaningful differences. Still, some economists might act strategically to avoid alphabetical discrimination, for example by changing their last name or by leaving the field.

Our study makes several theoretical contributions. First, prior research on collaboration and creativity has taken two distinct approaches. A large stream of literature has examined the implications of collaboration for output quality (Hargadon and Sutton 1997; Taylor and Greve 2006). Others have focused on individual incentives to collaborate (Merton 1968; Sarsons 2017). An implicit assumption in both streams of work is that the goals of maximizing output quality and individual credit are aligned. Our paper makes this assumption explicit and shows that it is not always warranted. Collaboration can lead to a decoupling of individual- and project-level incentives. The quest for individual credit is not always aligned with the pursuit of output quality.

Second, our study also contributes to a large body of literature emphasizing the mostly positive implications of collaboration in creative work. Many scholars have been optimistic about the spread of collaboration, noting that collaboration tends to be correlated with higher-quality output (Wuchty, Jones, and Uzzi 2007; Singh and Fleming 2010). We contribute to this stream of work by showing that this positive correlation might be partially driven by individuals collaborating on more-promising ideas rather than ideas leading to better outcomes because of collaboration. In our data, we replicate the finding that articles written by collaborating individuals tend to be of better quality on average. However, when we use alphabetical name ordering to identify the treatment effect of collaboration, we find that the latter is negative among collaborations formed because of the variances in credit premium. Our findings therefore suggest that prior studies' optimism regarding the causal impact of collaboration on output quality might be overstated because of lack of attention to selection biases.

Our study also has implications for the organization of R&D in firms. Collaboration has become ubiquitous, not only in academia but also in industry. And although this organization of innovative work can be advantageous, it also brings important challenges. Our findings highlight that the decoupling of individual incentives from project performance is particularly problematic from a managerial standpoint in contexts where employees have some say in their choice to collaborate. Firms that ignore these challenges might not only fail to reward individuals who contribute a lot but might also end up rewarding individuals who chose a work structure that benefits themselves more than the project. In particular, our findings highlight the risks of over-rewarding collaborative work. While it might not be easy to eliminate detrimental collaborations driven by the credit premium, closer attention to the division of labor and to each person's incentives can help companies better monitor the performance of individual team members and potentially intervene to mitigate the costs of collaboration to the extent possible.

We do not suggest that all collaborations have a negative impact on output quality or that all creative workers exhibit opportunistic behavior when engaging in collaboration. There is extensive evidence on the benefit of and need for collaboration, as discussed earlier in this paper. In other words, the take-away must not be to impose blanket policies that deter or excessively police collaboration behavior. Rather, we hope that our findings inspire a more nuanced evaluation of collaboration incentives, ensuring that workers are rewarded fairly, whether or not they choose to collaborate.

This paper is a first step in exploring the relationship between individual- and project-level incentives in collaborative creative work. In particular, we show that individuals can sometimes benefit from collaborating even when it reduces the quality of their creative output. The importance of furthering this line of research should not be understated. As an organization of work, collaboration is becoming more and more prevalent, and there is no sign that this trend is slowing. By highlighting the fact that collaboration can create a gap between what benefits a project and what benefits the individuals working on that project, we hope that our study enhances our understanding of the drivers of creative performance as a collective enterprise.

REFERENCES

Albanese, Robert, and David D. Van Fleet. 1985. "The Free Riding Tendency in Organizations." *Scandinavian Journal of Management Studies* 2 (2): 121–36. https://doi.org/10.1016/0281–7527(85)90003–9.

Bikard, Michaël, Fiona Murray, and Joshua Gans. 2015. "Exploring Tradeoffs in the Organization of Scientific Work: Collaboration and Scientific Reward." *Management Science* 61 (7): 1473–95.

Brewer, Marilynn B., and Roderick M. Kramer. 1986. "Choice Behavior in Social Dilemmas: Effects of Social Identity, Group Size, and Decision Framing." *Journal of Personality and Social Psychology* 50 (3): 543.

Carter, John R., and Michael D. Irons. 1991. "Are Economists Different, and If So, Why?" *Journal of Economic Perspectives* 5 (2): 171–77. https://doi.org/10.1257/jep.5.2.171.

Cattani, Gino, and Simone Ferriani. 2008. "A Core/Periphery Perspective on Individual Creative Performance: Social Networks and Cinematic Achievements in the Hollywood Film Industry." *Organization Science* 19 (6): 824–44. https://doi.org/10.1287/orsc.1070.0350.

Cattani, Gino, Simone Ferriani, Marcello M. Mariani, and Stefano Mengoli. 2013. "Tackling the 'Galácticos' Effect: Team Familiarity and the Performance of Star-Studded Projects." *Industrial and Corporate Change* 22 (6): 1629–62. https://doi.org/10.1093/icc/dtt001.

Choudhury, Prithwiraj, and Martine R. Haas. 2018. "Scope versus Speed: Team Diversity, Leader Experience, and Patenting Outcomes for Firms." *Strategic Management Journal* 39 (4): 977–1002. https://doi.org/10.1002/smj.2753.

Cummings, Jonathon N., and Sara Kiesler. 2007. "Coordination Costs and Project Outcomes in Multi-University Collaborations." *Research Policy* 36 (10): 1620–34.

Deichmann, Dirk, and Michael Jensen. 2018. "I Can Do That Alone . . . or Not? How Idea Generators Juggle between the Pros and Cons of Teamwork." *Strategic Management Journal* 39 (2): 458–75. https://doi.org/10.1002/smj.2696.

Dodgson, Mark. 1993. “Learning, Trust, and Technological Collaboration.” *Human Relations* 46 (1): 77–95. https://doi.org/10.1177/001872679304600106.
Dougherty, Deborah. 1992. “Interpretive Barriers to Successful Product Innovation in Large Firms.” *Organization Science* 3 (2): 179–202.
Edmondson, Amy. 1999. “Psychological Safety and Learning Behavior in Work Teams.” *Administrative Science Quarterly* 44 (2): 350–83.
Einav, Liran, and Leeat Yariv. 2006. “What’s in a Surname? The Effects of Surname Initials on Academic Success.” *Journal of Economic Perspectives* 20 (1): 175–87. https://doi.org/10.1257/089533006776526085.
Freeman, Richard B., Ina Ganguli, and Raviv Murciano-Goroff. 2014. “Why and Wherefore of Increased Scientific Collaboration.” *The Changing Frontier: Rethinking Science and Innovation Policy,* August, 17–48.
Goncalo, Jack A., Jennifer A. Chatman, Michelle M. Duguid, and Jessica A. Kennedy. 2015. “Creativity from Constraint? How the Political Correctness Norm Influences Creativity in Mixed-Sex Work Groups.” *Administrative Science Quarterly* 60 (1): 1–30. https://doi.org/10.1177/0001839214563975.
Hargadon, Andrew, and Robert Sutton. 1997. “Technology Brokering and Innovation in a Product Development Firm.” *Administrative Science Quarterly* 42 (4): 716–49. https://doi.org/10.2307/2393655.
Heath, Chip, and Nancy Staudenmayer. 2000. “Coordination Neglect: How Lay Theories of Organizing Complicate Coordination in Organizations.” *Research in Organizational Behavior* 22 (January): 153–91. https://doi.org/10.1016/S0191–3085(00)22005–4.
Heilman, Madeline, and Michelle Haynes. 2005. “No Credit Where Credit Is Due.” *Journal of Applied Psychology* 90 (5): 905–16.
Janis, I. L. 1972. *Victims of Groupthink.* Houghton Mifflin Boston.
Janis, Irving L. 1971. “Groupthink.” *Psychology Today* 5 (6): 43–46.
———. 2008. “Groupthink.” *IEEE Engineering Management Review* 36 (1): 36.
Janis, Irving Lester, and Irving Lester Janis. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes.* Vol. 349. Houghton Mifflin Boston.
Jones, Benjamin. 2009. “The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder?” *Review of Economic Studies* 76 (1): 283–317.
Karau, S. J, and K. D Williams. 1993. “Social Loafing: A Meta-Analytic Review and Theoretical Integration.” *Journal of Personality and Social Psychology* 65: 681–681.
Kay, Min B., Devon Proudfoot, and Richard P. Larrick. 2018. “There’s No Team in I: How Observers Perceive Individual Creativity in a Team Setting.” *Journal of Applied Psychology* 103 (4): 432–42. https://doi.org/10.1037/apl0000270.
Latané, Bibb, Kipling Williams, and Stephen Harkins. 1979. “Many Hands Make Light the Work: The Causes and Consequences of Social Loafing.” *Journal of Personality and Social Psychology* 37 (6): 822–32. https://doi.org/10.1037/0022–3514.37.6.822.
Lawrence, Paul R., and Jay W. Lorsch. 1967. “Differentiation and Integration in Complex Organizations.” *Administrative Science Quarterly* 12 (1): 1–47. https://doi.org/10.2307/2391211.
Leahey, Erin. 2016. “From Sole Investigator to Team Scientist: Trends in the Practice and Study of Research Collaboration.” *Annual Review of Sociology* 42 (1): 81–100. https://doi.org/10.1146/annurev-soc-081715–074219.
Lee, Sunkee, and Phanish Puranam. 2017. “Incentive Redesign and Collaboration in Organizations: Evidence from a Natural Experiment.” *Strategic Management Journal* 38 (12): 2333–52. https://doi.org/10.1002/smj.2685.
Marwell, Gerald, and Ruth E. Ames. 1981. “Economists Free Ride, Does Anyone Else?: Experiments on the Provision of Public Goods, IV.” *Journal of Public Economics* 15 (3): 295–310. https://doi.org/10.1016/0047–2727(81)90013-X.
Merton, Robert K. 1968. “The Matthew Effect in Science The Reward and Communication Systems of Science Are Considered.” *Science* 159 (3810): 56.
O’Dell, Jerry W. 1968. “Group Size and Emotional Interaction.” *Journal of Personality and Social Psychology* 8 (1, Pt.1): 75–78. https://doi.org/10.1037/h0025326.

Perlow, Leslie A. 1999. “The Time Famine: Toward a Sociology of Work Time.” *Administrative Science Quarterly* 44 (1): 57–81.

Porac, Joseph F., James B. Wade, Harald M. Fischer, Joyce Brown, Alaina Kanfer, and Geoffrey Bowker. 2004. “Human Capital Heterogeneity, Collaborative Relationships, and Publication Patterns in a Multidisciplinary Scientific Alliance: A Comparative Case Study of Two Scientific Teams.” *Research Policy* 33 (4): 661–78. https://doi.org/10.1016/j.respol.2004.01.007.

Reagans, Ray, and Ezra W. Zuckerman. 2001. “Networks, Diversity, and Productivity: The Social Capital of Corporate R&D Teams.” *Organization Science* 12 (4): 502–17. https://doi.org/10.2307/3085985.

Ross, Michael, and Fiore Sicoly. 1979. “Egocentric Biases in Availability and Attribution.” *Journal of Personality and Social Psychology* 37 (3): 322–36. https://doi.org/10.1037/0022–3514.37.3.322.

Sarsons, Heather. 2017. “Recognition for Group Work: Gender Differences in Academia.” *American Economic Review: Papers and Proceedings* 107 (5): 141–45.

Schroeder, Juliana, Eugene M. Caruso, and Nicholas Epley. 2016. “Many Hands Make Overlooked Work: Over-Claiming of Responsibility Increases with Group Size.” *Journal of Experimental Psychology: Applied* 22 (2): 238–46. https://doi.org/10.1037/xap0000080.

Singh, Jasjit, and Lee Fleming. 2010. “Lone Inventors as Sources of Breakthroughs: Myth or Reality?” *Management Science* 56 (1): 41–56.

Solla Price, D. J. de. 1965. *Little Science, Big Science . . . and Beyond.* Columbia University Press.

Stock, James H., and Motohiro Yogo. 2005. “Testing for Weak Instruments in Linear IV Regression.” *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80.

Taylor, A., and H. R. Greve. 2006. “Superman or the Fantastic Four? Knowledge Combination and Experience in Innovative Teams.” *Academy of Management Journal* 49 (4): 723.

Toh, Puay Khoon, and Francisco Polidoro. 2013. “A Competition-Based Explanation of Collaborative Invention within the Firm.” *Strategic Management Journal* 34 (10): 1186–1208. https://doi.org/10.1002/smj.2059.

Whyte, Glen. 1989. “Groupthink Reconsidered.” *Academy of Management Review* 14 (1): 40–56. https://doi.org/10.5465/amr.1989.4279001.

Wuchty, Stefan, Benjamin Jones, and Brian Uzzi. 2007. “The Increasing Dominance of Teams in Production of Knowledge.” *Science* 316 (5827): 1036.

Figure 1. The distributional characteristics of the final sample.

Figure 1a

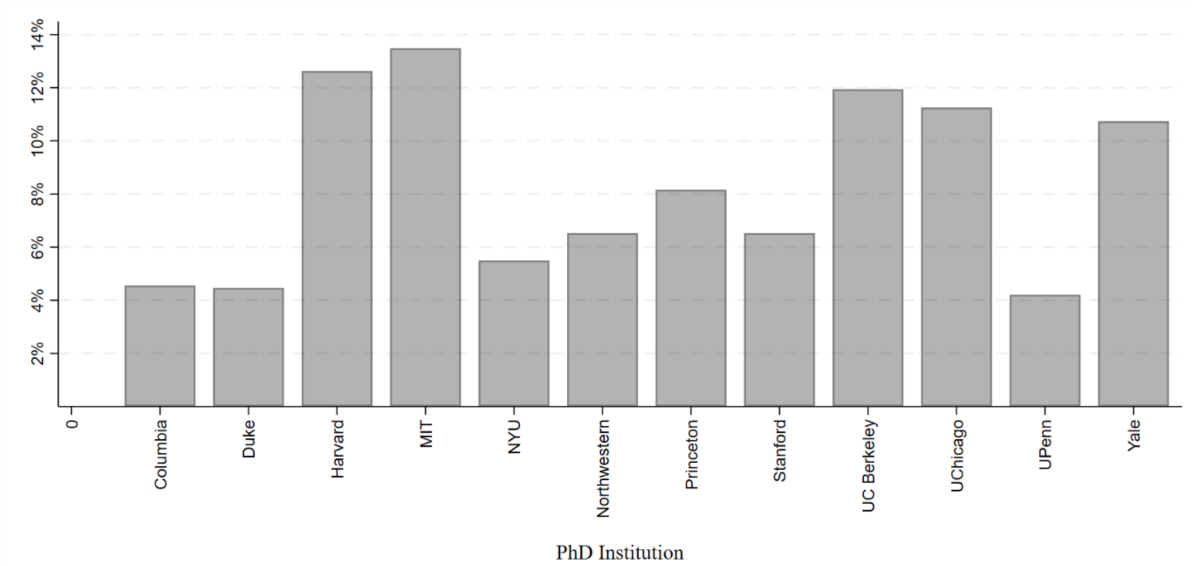


Figure 1b

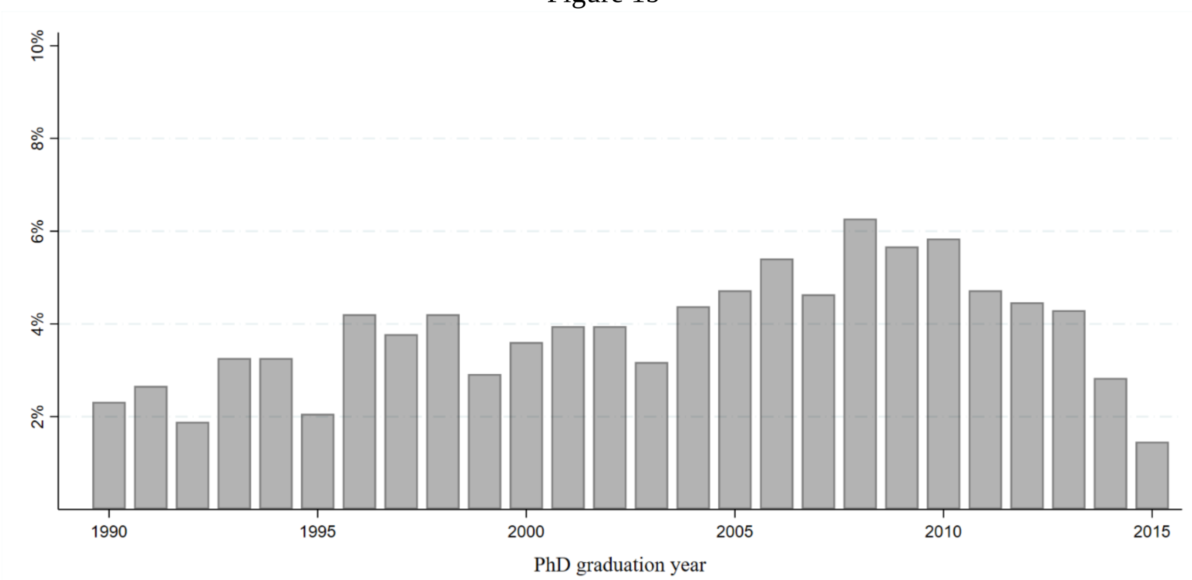


Figure 1c

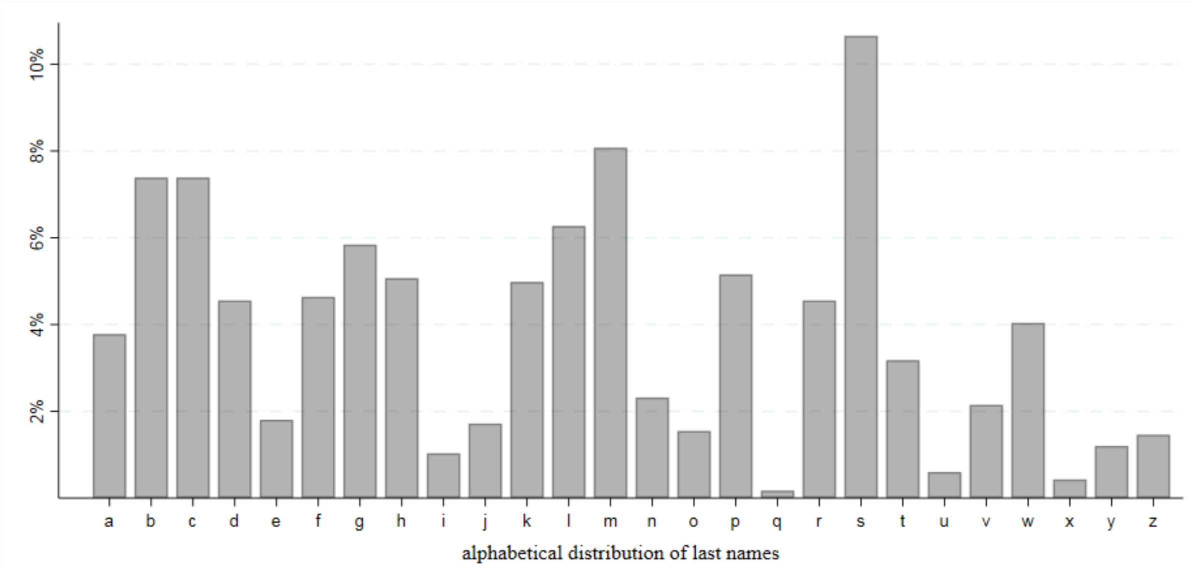


Figure 1d

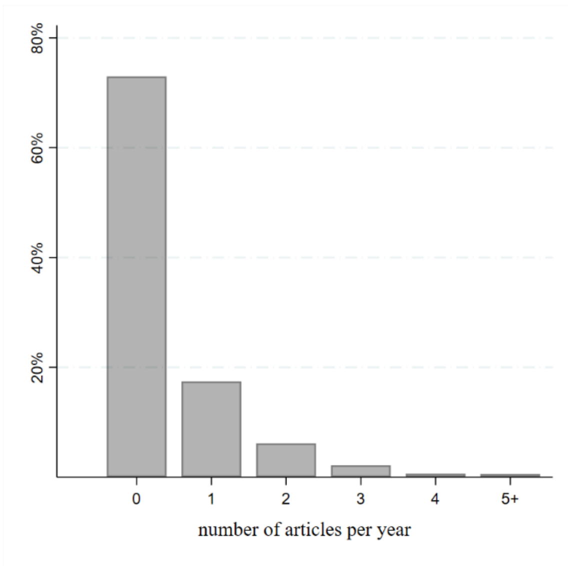
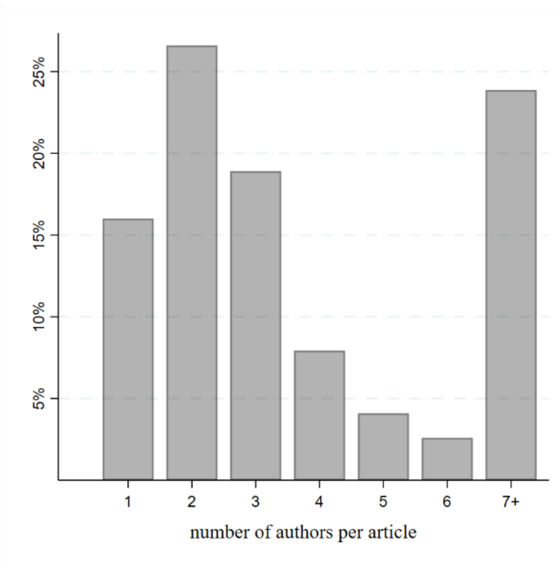


Figure 1e



Note. Figures 1a through 1e show the distribution of individuals across the institutions from which they graduated (N=1,164), the years of their graduation (N=1,164), the alphabetical distribution of their family names (N=1,164), their publication rates per year (N=16,260), and the distribution of the number of authors on all papers in our sample, respectively (N=16,260).

Table 1. Summary Statistics

	Obs.	Mean	St. Dev.	Min	Max
Alphabetical rank (individual level)	1,164	11.588	6.991	0	26
PhD graduation year (individual level)	1,164	2003.556	6.846	1990	2015
Female (individual level)	1,164	0.250	0.433	0	1
Asian family name (individual level)	1,164	0.299	0.458	0	1
Number of articles (individual-year level)	16,260	0.425	0.888	0	14
Collaborative articles (individual-year level)	16,260	0.240	0.667	0	14
Number of citations (individual-year level)	16,260	14.030	73.030	0	2,715
Cumulative count of articles (individual-year level)	16,260	1.678	3.479	0	84
Cumulative count of citations (individual-year level)	16,260	61.700	228.971	0	6,236
Number of articles not in alphabetical order	16,260	0.105	0.579	0	13

Table 2. The Relationship between Collaboration (Non-instrumented) and the Output Quality

Model:	**OLS**
DV:	**ln(citations+1)**
Count of collaborative papers	0.086*
	(0.047)
Ln(number of articles+1)	2.726***
	(0.151)
Ln(cumulative number of articles+1)	-0.449***
	(0.061)
Ln(cumulative number of citations+1)	0.189***
	(0.019)
Female	-0.003
	(0.019)
Asian family name	-0.016
	(0.016)
Number of articles not in alphabetical order	-0.153***
	(0.036)
Additional controls: year dummies, current institution dummies, current department dummies, starting year of PhD dummies, finishing year of PhD dummies	Yes
Obs.	16,260
F-statistics	132.06
Centered R^2	0.614

Notes. The analysis is at the individual-year level. The estimation is based on OLS regression with robust standard errors dual-clustered at the institution and year levels.
*** $p < 0.01$, ** $p < 0.05$.

Table 3. The Effect of Alphabetical Rank of Scientists on Their Propensity to Collaborate (First-Stage Results of the Instrumental Variable Method)

Model:	**OLS**
DV:	**Count of collaborative papers**
Alphabetical rank	-0.002***
	(0.000)
Ln(number of articles+1)	0.837***
	(0.050)
Ln(cumulative number of articles+1)	0.056***
	(0.014)
Ln(cumulative number of citations+1)	-0.002
	(0.004)
Female	-0.026***
	(0.008)
Asian family name	0.008
	(0.007)
Number of articles not in alphabetical order	0.312***
	(0.058)
Additional controls: year dummies, current institution dummies, current department dummies, starting year of PhD dummies, finishing year of PhD dummies	Yes
Obs.	16,260
F-statistics	114.46
Adjusted R^2	0.535
Kleibergen-Paap Wald F-statistic	27.349

Notes. The analysis is at the individual-year level. The estimation is based on OLS regression with robust standard errors dual-clustered at the institution and year levels. The critical value of the Stock-Yogo weak identification test for a bias smaller than 5% for a 5% level Wald test is 16.38.

*** $p < 0.01$.

Table 4. The Impact of Collaboration (Instrumented) on the Output Quality (Second-Stage Results of the Instrumental Variable Method)

Model:	**OLS**
DV:	**ln(citations+1)**
Count of collaborative papers (instrumented)	-0.998**
	(0.383)
Ln(number of articles+1)	3.632***
	(0.335)
Ln(cumulative number of articles+1)	-0.390***
	(0.068)
Ln(cumulative number of citations+1)	0.187***
	(0.019)
Female	-0.030
	(0.026)
Asian family name	-0.008
	(0.021)
Number of articles not in alphabetical order	0.184
	(0.143)
Additional controls: year dummies, current institution dummies, current department dummies, starting year of PhD dummies, finishing year of PhD dummies	Yes
Obs.	16,260
F-statistics	132.06
Centered R^2	0.614

Note. The analysis is at the individual-year level. The estimation is based on OLS regression with robust standard errors dual-clustered at the institution and year levels.
*** $p < 0.01$, ** $p < 0.05$.

Appendix

for

Detrimental Collaborations in Creative Work: Evidence from Economics

Estimating Collaboration Credit Premium

To estimate the level of credit allocation for each instance of collaboration in our sample, we use the method developed by Bikard et al. (2015). The method assumes that (1) scientists, on average, choose a level of collaboration that is expected to maximize their allocated credit; and (2) the credit allocated to each author on a paper amounts to a fraction of the citations to that paper. The latter assumption simply suggests that the credit for the paper must be shared among collaborators, albeit in a way that might sum up to more or less than 100%.

Following Bikard et al. (2015), we define $\alpha(N)$ as the average fraction of citations attributed to each author listed on a paper, where N is the total number of authors. In the presence of credit premium, $\alpha(N)$ would be larger than $\frac{1}{N}$, and in the presence of credit penalty $\alpha(N)$ would be smaller than $\frac{1}{N}$. In other words, instead of authors each on average obtaining one Nth of the citations to a paper as their credit, they could receive $\alpha(N) > \frac{1}{N}$ of the citations (credit premium) or $\alpha(N) < \frac{1}{N}$ (credit penalty). Thus, to test for the presence of average credit premium in our sample, we need to identify $\alpha(N)$ and compare its value with $\frac{1}{N}$. To do so, we again rely on the Bikard et al. (2015) arguments. Their approach observes that if the presumed level of $\alpha(N)$ reflects the true value of credit allocation in the sample, regressing the calculated fractional credit for each author of a paper on the paper's number of authors should produce an effect of zero. This intuition is driven by the first assumption behind the method—that individuals, on average, are aware of the level of collaboration that maximizes the credit attributed to them. Thus, if individuals optimize, we should see neither a positive nor a negative effect at the population level. A positive effect means that higher levels of collaboration increase the allocated credit, in which case the optimal response is to collaborate more. A negative effect means that the level of collaboration has a negative impact on the allocated credit, in which case the optimal response is to collaborate less. Note that the method does not require every researcher to have a precise estimate of their level of credit allocation. Rather, it requires that, on average, researchers understand whether increasing or decreasing the number of collaborators would increase or decrease their individual credit.

Based on this intuition, we can then try different levels of $\alpha(N)$ to construct the individual credit as the dependent variable and check which one generates an estimated effect of zero for the number of collaborators in the said regression. Including individual, department, and institution-year fixed effects in the regression further controls for idiosyncratic characteristics of individuals, their departments, and their institutions that could affect their collaborative choices. We use the following regression:

$$\ln(1 + Frac\_Credit_{it}) = \beta_0 + \beta_1 NAuthors_{it} + \mu_i + \varphi_{it} + \gamma_{it}.\delta_t + \varepsilon_{it}$$

where the dependent variable is the natural log of 1 plus the sum of fractional credits for papers produced by individual i in year t. The i's fractional credit on each paper is calculated as $\alpha(N)$ multiplied by the number of citations received by the paper. Following the standard procedure in the field, we use log normalization to address the heavily skewed citation rates in our sample. $NAuthors_{it}$ is the average number of authors listed on individual i's publications in year t. μ_i, φ_{it}, and $\gamma_i.\delta_t$ each represent individual, department, and institution-year fixed effects, respectively. We use an ordinary least squares (OLS) regression with robust standard errors clustered at the individual level for the estimation.

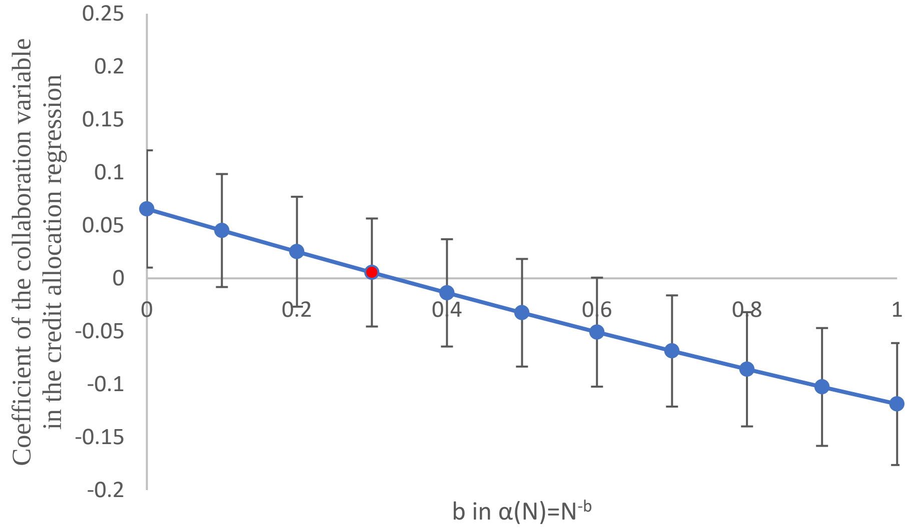
To construct different levels of $\alpha(N)$ in a systematic way, following Bikard et al. (2015), we define $\alpha(N)$ as N^{-b} where N is the number of authors and b ranges from 0 to 1. Thus, for $b = 0$, $\alpha(N)$

equals 1 which represents the maximum amount of credit premium possible—that is, authors each receive 100% of the number of citations to their papers as their fractional credit. When $b = 1$, $\alpha(N)$ equals $\frac{1}{N}$, which represents the case of zero credit premium—that is, authors each receive on average one Nth of the number of citations to their papers as fractional credit. Thus, values of $0 < b < 1$ represent different levels of credit premium in this range, and values of b > 1 represent different levels of credit penalty.

Following this procedure, we ran the above-described regression for all values of b between 0 and 1 with increments of 0.1. Since we find evidence for collaboration credit premium within this range of values for b, there is no need to extend our testing to values of b that are greater than 1. Table A1 below shows the regression results for each value. Figure A1 depicts the estimated value of β_1 (along with its respective 95% confidence interval) with respect to different values of b. The results suggest that the estimated β_1 is closest to zero for $b = 0.34$, which suggests a credit premium level equal to $N^{-0.34}$. This means that for a team of two authors on a paper, each author receives roughly 79% of the citations to the paper as individual credit, an additional 29 percentage points of credit relative to the no-premium scenario. The magnitude of the effect is similar to that reported in Bikard et al.'s (2015) estimated credit premium of $N^{-0.48}$ for their sample of MIT science and engineering faculty members, suggesting that the level of collaboration credit premium among the economists in our sample is slightly larger. Figure A2 replicates Figure A1 for tenured faculty, showing that tenured faculty also experience a similar level of collaboration credit premium in our sample.

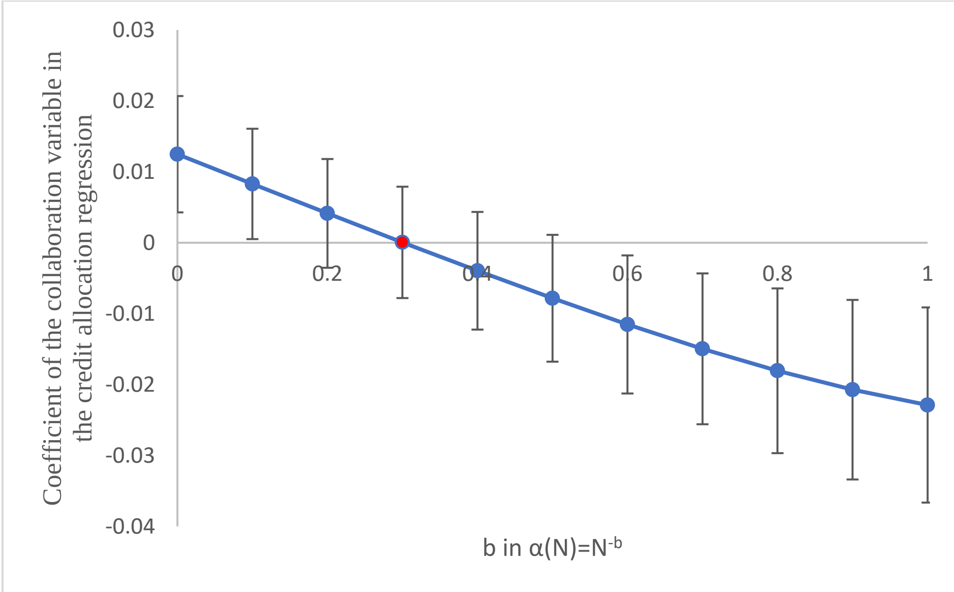
The norm of alphabetical ordering on economics articles means that individuals with a family name that begins with a letter from the beginning of the alphabet should experience a higher collaboration credit premium than their peers whose family names begin with letters from the end of the alphabet. We repeat our estimation of credit premium separately for individuals whose family names begin with letters toward the first half of the alphabet (A to M) and for those whose family names begin with letters toward the second half of the alphabet (N to Z). Figure A3 shows the results for the two groups, suggesting that the former group enjoys a higher level of credit premium ($\alpha(N) = N^{-0.29}$) than the latter ($\alpha(N) = N^{-0.55}$).

Figure A1: Imputing the level of credit allocated to each co-author from scientists' collaboration choices



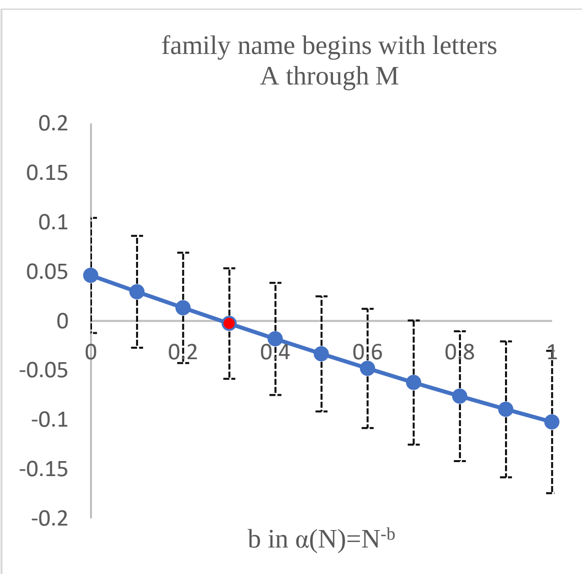
Note: *α(N)* indicates the level of fractional credit allocated to each author on a publication in our sample and is defined as N^{-b} where *N* is the number of authors and *b* ranges from 0 to 1 (with increments of 0.1). The graph shows the relationship between different assumed levels of *b* and collaboration rate. The *α(N)* for which this relationship is neither positive nor negative reveals the average level of credit allocated to each co-author in our sample. Table A1 in the appendix shows the regression results for each value of *b*.

Figure A2: Replicating figure A1 for the sample of tenured faculty



Note: this graph replicates Figure 2 in the main article for the sample of tenured faculty. *α(N)* indicates the level of fractional credit allocated to each tenured author on a publication in our sample and is defined as N^{-b} where *N* is the number of authors and *b* ranges from 0 to 1 (with increments of 0.1). The graph shows the relationship between different assumed levels of *b* and collaboration rate. The *α(N)* for which this relationship is neither positive nor negative reveals the average level of credit allocated to each co-author in our sample.

Figure 3: Comparing the level of credit premium for scientists whose family name begins with letters A through M versus those whose family name begins with letters N through Z



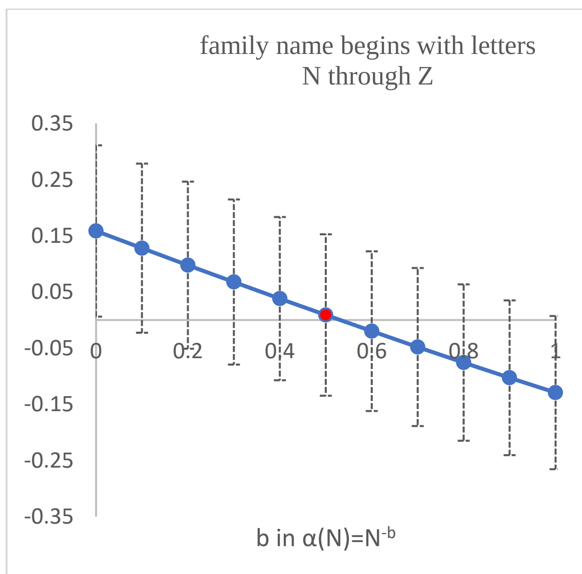


Table A1: Regression results for imputing the level of credit allocated to each co-author from scientists' collaboration choices

Model:	**OLS**										
Dependent variable:	**Fraction credit (yearly citations × N^{-b}) allocated to each co-author if**										
	b=0	**b=0.1**	**b=0.2**	**b=0.3**	**b=0.4**	**b=0.5**	**b=0.6**	**b=0.7**	**b=0.8**	**b=0.9**	**b=1**
Number of authors	0.065**	0.045*	0.025	0.005	-0.014	-0.032	-0.051*	-0.069**	-0.086***	-0.102***	-0.119***
	(0.028)	(0.027)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.027)	(0.028)	(0.029)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
institution-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650
Adjusted R-Squared	0.648	0.647	0.646	0.646	0.645	0.644	0.644	0.644	0.643	0.643	0.643

Note: The table presents the results used for creating Figure 2 in the main paper. The dependent variable in each column is the level of fractional credit allocated to each author in our sample and is defined as the total number of citations to that author's publications in any given year multiplied by N^{-b}, where N is the average number of authors on those publications and b is the parameter of interest that determines the level of fractional credit allocation. Each column shows the same regression for different values of b ranging from 0 to 1 with increments of 0.1.

*** p<0.01, ** p<0.05, * p<0.1

Table A2: The relationship between collaboration (non-instrumented) and the output quality

Model:	OLS
DV:	ln(citations+1)
Count of collaborative papers	0.086*
	(0.047)
Ln(number of articles+1)	2.726***
	(0.151)
Ln(cumulative number of articles+1)	-0.449***
	(0.061)
Ln(cumulative number of citations+1)	0.189***
	(0.019)
Female	-0.003
	(0.019)
Asian family name	-0.016
	(0.016)
Number of articles not in alphabetical order	-0.153***
	(0.036)
Additional controls: year dummies, current institution dummies, current department dummies, starting year of PhD dummies, finishing year of PhD dummies	Yes
Obs.	16,260
F-statistics	132.06
Centered R^2	0.614

Note: The analysis is at the individual-year level. The estimation is based on OLS regression with robust standard errors dual clustered at the institution and year levels. *** p<0.01, ** p<0.05

Table A3: The first-stage results of the instrumental variable method for tenured faculty

Model:	OLS
DV:	Count of collaborative papers
Alphabetical rank	-0.002
	(0.002)
Ln(number of articles+1)	1.332***
	(0.051)
Ln(cumulative number of articles+1)	0.057*
	(0.033)
Ln(cumulative number of citations+1)	0.021
	(0.012)
Female	-0.005
	(0.046)
Asian family name	0.068
	(0.040)
Number of articles not in alphabetical order	0.189***
	(0.037)
Additional controls: year dummies, current institution dummies, current department dummies, starting year of PhD dummies, finishing year of PhD dummies	Yes
Obs.	6,434
F-statistics	249.76
Adjusted R^2	0.600

Note: This table replicates the results in Table 2 for tenured faculty in our sample. The analysis is at the individual-year level. The estimation is based on OLS regression with robust standard errors dual clustered at the institution and year levels.
*** p<0.01, * p<0.1

Table A4: The second-stage results of the instrumental variable method based on a two-stage generalized method of moments (GMM)

Model:	**OLS**
DV:	**ln(citations+1)**
Count of collaborative papers (instrumented)	-0.998**
	(0.383)
Ln(number of articles+1)	3.632***
	(0.335)
Ln(cumulative number of articles+1)	-0.390***
	(0.068)
Ln(cumulative number of citations+1)	0.187***
	(0.019)
Female	-0.030
	(0.026)
Asian family name	-0.008
	(0.021)
Number of articles not in alphabetical order	0.184
	(0.143)
Additional controls: year dummies, current institution dummies, current department dummies, starting year of PhD dummies, finishing year of PhD dummies	Yes
Obs.	16,260
F-statistics	132.06
Centered R^2	0.614

Note: The analysis is at the individual-year level. The estimation is based on OLS regression with robust standard errors dual clustered at the institution and year levels. *** p<0.01, ** p<0.05

Table A5: The second-stage results of the instrumental variable method based on a limited-information maximum likelihood method

Model:	OLS
DV:	**ln(citations+1)**
Count of collaborative papers (instrumented)	-0.998**
	(0.383)
Ln(number of articles+1)	3.632***
	(0.335)
Ln(cumulative number of articles+1)	-0.390***
	(0.068)
Ln(cumulative number of citations+1)	0.187***
	(0.019)
Female	-0.030
	(0.026)
Asian family name	-0.008
	(0.021)
Number of articles not in alphabetical order	0.184
	(0.143)
Additional controls: year dummies, current institution dummies, current department dummies, starting year of PhD dummies, finishing year of PhD dummies	Yes
Obs.	16,260
F-statistics	132.06
Centered R^2	0.614

Note: The analysis is at the individual-year level. The estimation is based on OLS regression with robust standard errors dual clustered at the institution and year levels. *** p<0.01, ** p<0.05