Credit History: The Changing Nature of Scientific Credit

Joshua S. Gans* and Fiona Murray**

PRELIMINARY – DRAFT ONLY

19th July 2013

* Rotman School of Management, University of Toronto and NBER.
** Sloan School of Management, MIT and NBER

I. Introduction

We are interested in the changing institutional conditions confronting scientists engaged in cumulative innovation since Vannevar Bush famously evoked the image of an Endless Frontier of cumulative scientific progress. In seeking to explain how the organization of scientific knowledge production has changed, scholars have focused on two critical factors: the vast range of new technologies that have enabled scientific progress (Agrawal) and the increasing burden of knowledge needed for cumulative progress at the frontier (Jones). We propose a third factor grounded in the formal institutions and norms of science that has previously been overlooked: the assessment and allocation of credit. The institutions of credit are central to the incentive system of open science (Merton 19xx, Dasgupta and David 1998). While not as easy to observe as the large pieces of new equipment that fill laboratories today, and harder to conceive than the notion that contributing something today means knowing more than those who contributed fifty years ago, we argue that credit has also changed in a range of critical dimensions in the years after the "Endless Frontier."

In this essay we focus briefly on the two traditional arguments made regarding the changing organization of science. We then develop our perspective on the core organizational choices made by scientists as a way of motivating the central importance of scientific credit in the ways in which knowledge production is organized. This lays the groundwork for us to elaborate our "credit history" – how the institutions and norms of credit have changed over the past fifty years. We do this by exploring three debates that have animated the scientific community over the past fifty years. To provide more specificity, we then focus on the specific case study of synthetic organic chemistry. Building our general and case study insights, we then elaborate a formal model that places credit allocation alongside the changing technical costs and knowledge burden of research to explore the relative importance of these three factors and made predications for how science is likely to change going forward.

II. Credit and the Organization of Science

A. Credit in the Sociology and Economics of Science

The post-Vannevar Bush years saw a growing interest among scholars from disciplines spanning the history, economics, sociology and philosophy of science in understanding science, particularly the production and use of scientific knowledge. Pioneers including de Solla Price (1963), Garfield and Merton were among the leading scholars to explore not only the content but the social and institutional nature of scientific progress at the frontier. In doing so, they often built on the long tradition of bibliometrics (sometimes referred to as scientometrics) thus focusing their attention on the observable outputs of scientific research: publications (see Nalimov and Mulchenko 1969, Pritchard 1969 and later van Raan 1988 for background on the role of bibliometrics in science). As Price described, they "turn[ed] the tools of science on science itself" (Price 1963, p. XX) using publications and the citations that they received from follow-on researchers to systematically examine and describe the production, dissemination and use of scientific knowledge.

With regard to the organizational choices made by scientists, publications provide some important tracers of their decisions at least in terms of the number and nature of collaborators: publication authorship is at least a partial reflection of those who contributed to a project. Certainly, when compared to many other arenas of knowledge production (for example, inside corporations), publications provide a significant window into research choices because of the centrality of disclosure and openness to the scientific community (see Gans, Murray and Stern 2012). Nonetheless, while publications provide critical empirical data on scientific progress and are important for our understanding of knowledge disclosure, their analysis can obscure the meaningful organizational choices made by scientists and the ways in which these have changed over the past several decades. Prior to these disclosure choices, scientists face a range of organizational choices that we place front and center of our analysis of the transformation of knowledge production in the United States since Vannevar Bush.

B. Organizational Choices and Credit

A useful way to understand scientists' organizational choices is to build on the conceptual approach developed by Green and Scotchmer (1995) for cumulative innovation in the private sector and Aghion et al. (2008) in the context of scientific research. From this perspective, we understand research to follow a particular "line" which sets an intellectual trajectory for progress. With respect to organizational arrangements, we argue that researchers make three critical choices: First they must determine a sequence of cumulative projects that follow along the line they are pursuing; i.e., they set a particular intellectual trajectory and map out two or more projects along that line. Second, they must determine the optimal way to approach these projects with respect to collaborative choices. Third, they must determine their disclosure choices for these projects. Taken together, these three choices lead to three distinctive outcomes:

- **Integration** – under an organizational strategy we call integration, scientists may choose to undertake all projects in a line themselves and only then publish.
- **Collaboration** – under a collaboration strategy, they bring in collaborators (from other laboratories presumably with complementary skills) to complete the projects in a line and publish a paper with co-authors.
- **Publication** – rather than collaborate or integrate a scientist may choose to publish individual line stages and then simply wait to be cited (in the market for ideas) by follow on research further down the line.

Our organizational perspective highlights the central role of credit in the organization of science. Specifically, a scientist choosing among these options must consider the cost of pursuing each project along the line and the time it will take follow-on researchers to accumulate the relevant specialized knowledge – the traditional factors thought to shape organizational choices from one laboratory to another. Credit is also a key factor. The argument is as follows: selecting whether to integrate, collaborate or publish and rely on the citation market depends at least in part on the ways in which scientists' believe that they will be rewarded for each of these alternatives. On the one hand, integration provides significant credit for a substantial amount of research progress along the line, but this must be balanced against the potential costs of acquiring the specialized knowledge and accessing relevant technology. In contrast, the attractiveness of collaboration depends upon how credit is allocated between the scientist and his collaborators (see Bikard, Gans and Murray 2013 for an elaboration of this issue). Lastly, citation markets provide an alternative form of credit – in the form of citation and acknowledgement that may itself be valued by researchers and those who evaluate them. The tradeoffs among these organizational choices emphasize the importance of credit as a countervailing set of benefits that balance the technical costs of pursuing particular organizational strategies along a given research line.

While not described in the context of organizational choices and tradeoffs, distinguished sociologist of science Robert Merton underscored the central importance of institutions of credit and the informal norms regarding credit which he describes as one of the "psychosocial processes affect[ing] the allocation of rewards to scientists for their contributions- an allocation which in turn affects the flow of ideas and findings through the communication networks of science" (p. 56). Best known for his observation of the Matthew Effect (Merton 1968) - the notion that credit is differentially allocated to those who are already famous and have had prior credit bestowed upon them, Merton describes credit as being potentially "mis-allocated" under some conditions arguing that "Eminent scientists get proportionately great credit for their contributions to science while relatively unknown scientists tend to get disproportionately little credit for comparable contributions." (Merton 1968, 443) This underscores credit's importance, and the potential for its changing allocation, for scientific organization.

Building on Merton's work, other sociologists of science have explored the importance of credit in the careers of scientists, showing for example, that highly productive scientists are more significantly rewarded for their work (Glaser, Crane). In addition, since the pioneering work of Hagstrom, it has been recognized that citations provide an additional reward to researchers as part of an exchange relationship that places credit and recognition is placed at the center of a system by which scientists disclosure and information is provided in return for credit in the form of citations (see also Murray 2010 for an exploration of this process in the context of patent rights). Cole and Cole have been more elaborate in their description of the different types of rewards scientists' accrue: professorships in leading departments, honorific titles, and wide citation being among the most salient.

The role of credit in the organization of science came to the attention of economists when Partha Dasgupta and Paul David published their influential 1994 paper "Towards a New Economics of Science". In it, they argued for the need to re-examine the organizational structures as well as the institutions and policies of science. They suggest that "it is a system that remains an intricate and rather delicate piece of social and institutional machinery" (p. 489) and emphasize the importance of the norms and general "institutions" governing the production of knowledge. While drawing on the sociology of science, their approach has been to examine some of the consequences of particular economic characteristics of knowledge production: the costs of monitoring effort, the indivisibility of some projects, and the potential for selective (and non) disclosure. They too focus on the reward system as a central element in science and argue that "an individual's reputation for 'contributions' acknowledged within his or her collegiate reference groups is the fundamental 'currency' in the reward structure that governs the community of academic scientists" (p. 498). This reference to the notion of currency highlights the notion of credit and currencies of credit as being central not only to the norms of science but also its economic foundations. From the perspective of a social planner, the importance of priority in credit speeds up discovery along research lines and ensures their disclosure. From the perspective of a scientist organizing to pursue a given research line issues of credit allocation shape their organizational choices.

The notion that credit is a critical factor driving the organization of science is in counterpoint to prior approaches taken in the academic literature. Traditionally, scholars have examined two main determinants of the organization of research: the technology of knowledge production and the burden of knowledge. With regards to the influence of technological change, a significant body of knowledge has argued for the important (albeit complex) role of new technologies in facilitating the pursuit of scientific progress (see Mokyr 2002). From Boyle's air pump (Shapin and Schaffer 1986) to Volta's pile (Pancaldi 2003), new technologies have enabled scientists to pursue more complex and distinctive research lines.

In the post-Vannevar Bush period, this has been of particular importance in areas such as biology and physics (Knorr-Cetina 1999). In Biology, the invention and automation of Polymerase Chain Reaction (PCR), DNA sequencing and DNA synthesis have, among other technologies, opened new biological research lines. Likewise, in physics, the development of new, more powerful telescopes and massive particle colliders (each with their attendant computing power) have enabled the exploration of new knowledge frontiers while at the same time changing the lives of physicists, and their ways of collaborating and organizing of research (Galison 1997). Beyond the specific technologies of knowledge production, recent work (e.g. Agrawal and Goldfarb 2008) has highlighted the ways in which the coming of the Internet has shaped the organization of research and the extent and nature of collaboration versus integration. In particular, the data (in engineering) show that faculty in middle-ranking universities have seen the greatest organizational change, becoming more likely to be engaged in top-tier collaborations than prior to Bitnet introduction.

An alternative, or perhaps more accurately a complementary, perspective on the changing organization of science is articulated by Ben Jones who outlined the importance of the burden of knowledge on researcher's organizational choices. His line of argument focuses on the growing length of scientific training as scientists seek to accumulate an ever growing body of knowledge in order to make contributions at the frontier. As a corollary to the increasing requirement for training, scientists are accordingly becoming narrower in their expertise and more highly specialized – an effect he refers to as the death of the Renaissance man (Jones 2008). According to the burden of knowledge argument, the combined need for more and more specialized knowledge leads researchers into pursuing their chosen research lines through higher levels of collaboration. In support of this perspective, Jones presents compelling data on the rise in number of authors on scientific publication across all fields (Wutchy et al. 2007). However, this approach, by examining only the outputs of researcher rather than their underlying organizational choices, potentially overlooks the ways in which credit allocation must, necessarily, have changed over the past fifty years, in order to account for the benefits of collaboration (Bikard, Gans and Murray 2013). Alternatively, if credit has *not* changed to account for the growing need to collaborate, or the larger scientific communities within which evaluation of contributions is more challenging, then the tradeoffs between collaboration and integration will be driven by both burden of knowledge and credit considerations.

C. Credit History

While conceptual in our exposition, the three organizational alternatives we have laid out – integration, collaboration and publication – reflect the very real choices made by scientists. As our discussion above highlights, these choices are contoured not only by changing technology of production and changes in the knowledge burden. They are also shaped by changes in the central reward mechanism of science – credit allocation. Unlike technologies that are readily observable, or even the knowledge burden – which can be imputed at least in part from educational requirements – the changing nature of credit and its allocation is more difficult to trace and demonstrate. We aim to do so by developing a short history of the debates around credit and the scientific reward system that animate scientists themselves – our so-called credit history. This history is basically told from the perspective of scientists themselves as captured in the editorial pages of their major journals – *Science* and *Nature* among others as well as the *Chronicle of Higher Education (CHE)*. While not comprehensive, our explication of the discussions in the scientific community that have raged as individuals continue to debate the changing nature and organization of science and the role of credit in these changes provides us with a window into the challenges that scientists confront as they wrestle with their own, autonomous (but fragmented) credit system.

Several debates are relevant to the link between credit and the organization of scientific knowledge production: authorship conventions (including ordering and ghost authorship) which link to tradeoffs around credit and collaboration, salami slicing which speaks to the role of credit in choices of integration versus citation markets, and, lastly, the role of quantitative metrics (including publication numbers and citation counts) in tenure decisions which speak to the overall set of organizational choices. Other debates such as citation conventions and the Matthew Effect are also relevant.

The first critical issue is the link between research organization (and production), authorship and credit. Gaeta[1] (1999) has coined the term authorship "Law and order" to connote not only the rules of authorship but also the specific role of ordering of authorship (see Gans and Murray 2013 for a more comprehensive theoretical treatment of author ordering and credit). A vivid example of the challenges associated with these changing rules is described in a 1981 article: "*The fellowship application for the American College of Physicians asks a candidate to list per-cent participation in studies in which he is a listed author. Though seemingly a workable solution, the accuracy of the resulting judgments has been called into question. In at least one instance, when a whole research team applied for fellowships, their total participation came to 300 percent*" (Broad 1981, p. 1138). This suggests that while some genuine changes in collaboration may be taking place driven by specialization and technology, gratuitous authorship as

---

[1] http://www.ncbi.nlm.nih.gov/pubmed/10230981

well as fragmentation of results (as we discuss below) were increasing from the 1970s onwards. Not only an issue for credit, authorship conventions – rather than reflections of changing research organization – also raise questions of responsibility and liability for research findings and for the potential "false science" including fraud, that might have taken place (see Furman, Jensen and Murray 2001 and Azoulay et al. 2013 for a broader analysis of retractions in this context). Emerging authorship norms place an increased burden on the reward system of science, particularly in the evaluation of young faculty through key career milestones. Over a fifty-year period when the number of researchers (and their specialization) has increased, the evaluation of individual biographies (i.e. published contributions) is complex. Even in 1981, an editor of the *New England Journal of Medicine* notes that "You have to know the journals, and what impact they have.... You have to know the institutions, the people, the meetings.... It's a ticklish matter" (quoted in Broad 1981, p 1139). Today, evaluative choices for promotion, tenure, grant making and a wealth of other forms of scientific credit rely on publishing records with those records increasingly murky. Likewise, scientists themselves must make organizational choices over collaboration, integration or publishing in the shadow of a complex credit allocation process: one that is beset with error and indeterminacy over credit and authorship norms.

A second major theme that emerges in our credit history takes the colourful label "salami slicing" and emphasizes a debate that has emerged among scientists regarding the "size" of the least publishable unit (LPU) along a research line. This debate emphasizes the organizational choices between integration on the one hand and publication (of a smaller slice of the research line) with follow on citation by other researchers (or by self-citation by the research team). It focuses on the question of whether or not to publish the small step embodied in project 1 and wait for citation by another researcher pursuing project 2 (or pursing self-citation) or to complete projects 1 and 2 before publishing a larger slice of research thus making a larger contribution.

There is clear statistical evidence, from the 1970s onwards, to support the claim that publication length in shortening: particularly in the biological sciences, papers have shortened at significant rates. While not conclusive evidence of the rise of LPUs, anecdotal evidence supports the publishing dilemma of young faculty and the link between LPUs and credit. One Dean of Science described the dilemma in an article in *CHE* as follows*:* "*In order to appear to have more publications on their CVs, young scholars are often advised to break their research down into pieces and publish those pieces in multiple articles -- i.e., LPU's.... Having a couple of LPU's will ensure that the bean counters cannot assail her record. We both know that there are those among us who would easily ignore her aggressive pursuit of grants and a single brilliant paper in Cell if her four years here did not include the magic two papers.*"[2]

Far from being a new issue, discussions over LPUs and the link between publication strategies and credit can be traced back through the editorial pages of *Science* at least to 1981 (Broad 1981). In a provocative article, the careers of young scientists in the 1980s who typically had between 50 and 100 publications at the time of promotion were contrasted with a scientist from the late 1950s such as James Watson who, when being evaluated by peers (only a decade after Vannevar Bush), had only 18 papers (albeit

[2] http://chronicle.com/article/In-Defense-of-the-Least/44761

one that described the structure of DNA!). Broad noted the emergence of the LPU and argued that "*the increases stem not from a sharp rise in productivity but rather from changes in the way people publish. Co-authorship is on the rise, as is multiple publication of the same data.*" (p. 1137) He also notes the challenge for credit allocation questioning how in combination, LPU practices and changing co-authorship obfuscate the effort made by young scholars making evaluation and credit much more challenging.

More recently, in 2005, the journal *Nature Materials* explored the impact on the sustainability of scientific publishing of what it referred to as "fragmenting single coherent bodies of research into as many publications as possible — the practice of scientific salami slicing."[3] This has implications for the effectiveness and capacity of scientists to engage in meaningful peer review. It also leads to a much greater likelihood that publications will be plagiarized (at least in part), be overlapping or in other ways cross the boundary into false science (see Azoulay et al 2013). The process also has consequences for credit allocation: As the editors of *Nature Materials* outlined, poor practices associated with salami slicing "deny referees and editors the opportunity of assessing the true extent of its contribution to the broader body of research" raising the question of credit allocation for researchers selecting between integration and publishing.

D. Credit in Synthetic Organic Chemistry [Incomplete]

Our credit history paints the changing institutional norms around credit and its link to organizational choices in broad strokes. Before moving to a more formal model of

[3] http://www.nature.com/nmat/journal/v4/n1/full/nmat1305.html

the role of credit in the organization of knowledge production, we provide a more detailed case study of the changing role of credit in one area of science – synthetic organic chemistry (SOC). Not as widely explored as fields in the life sciences, this branch of chemistry (typically unloved by college pre-meds) emphasizes the total synthesis of complex organic (carbon backbone) molecules of medicinal, industrial or intellectual interest through an (often long) sequence of chemical reactions. It is a field that has garnered its share of Chemistry Nobel Prizes for individuals such as R.B. Woodward (1965) and E.J. Corey (1990) and has remained at the frontier of both pure knowledge production and useful application (Stokes 1987). As Corey himself described in his Nobel Prize Banquet Speech: "*Chemical synthesis is uniquely positioned at the heart of chemistry, the central science. Its impact on our lives and society is all pervasive. For instance, most of today's medicines are synthetic and the majority of tomorrow's will be conceived and produced by synthetic chemists. To the field of synthetic chemistry belongs an array of responsibilities which are crucial for the future of mankind, not only with regard to the health and needs of our society, but also for the attainment of a deep understanding of matter, chemical change and life.*"[4]

For our purposes, SOC is a particularly interesting arena of knowledge production because, unlike some other areas, it has not been the setting for significant changes in the technology of knowledge production. Nor has the burden of knowledge exploded dramatically. As a result it provides a case in which these factors are held constant and

[4] http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1990/corey-speech.html

the changing credit history becomes salient and central to the organizational choices of researchers. [To be completed]

III. Formal Model

We argue that it is the combined elements of technology, knowledge burden and credit institutions that shape the observable outputs of science – the number of publications, the number of authors, the rate of progress and its direction. More importantly, they inform the underlying organization of science: the decisions made by scientists about "laboratory life" particularly as it pertains to any particular research line the laboratory may be pursuing. With this in mind, we provide a formal model of the drivers of the choice of (optimal) organizational form for cumulative science and, in particular, how this choice is driven by mechanisms and institutions to allocate credit to individual scientists for their role in cumulative knowledge creation.

A. Basic Set-Up

We consider an environment whereby knowledge is created by cumulative scientific research. Specifically, we focus on a pioneer scientist's (*P*) decisions with respect to an initial scientific project, 1. Following Green and Scotchmer (1995) (and also Bresnahan (2011)), we assume that the cost to the scientist of pursuing 1 is c_1. The stand-alone (social) value of project 1 is x.

A follow-on project that builds on 1, project 2, may also be possible and can be conducted by *P* or another follow-on scientist, *F*. For a scientist, i, with in-depth knowledge of project 1, the probability that they perceive the project 2 opportunity is p_i.

To acquire the necessary in-depth knowledge of project 1 costs a scientist, C_i, so long as they have access to project 1's knowledge in the first place. The idea here is that, while project 1 knowledge may be disclosed (say through publication or communication), understanding it in a way that leads to follow-on research takes additional effort (for instance, by undertaking replication studies). If project 2 is possible, it comes with an expected quality, y, and research cost, c_{2i}. The interesting cases emerge when $c_{2F} \leq c_{2P}$ so we will assume that this is the case in what follows.

There are three ways research into projects 1 and 2 can be organized. First, under *integration*, P conducts both projects before publishing their results under sole authorship. Second, under *collaboration*, P collaborates with another scientist, F, over both projects. In this situation, each focuses on the project they can do at the least cost but pool their skills and communication in understanding the implications of project 1 for the project 2 opportunity. Third, under *publication*, P publishes their project 1 results and then F conducts research into project 2, citing back to P's initial contribution. Under both collaboration and publication, the market awards P and F with attribution regarding each scientist's contributions. A key focus here will be on how that attribution takes place.

B. Integration

In this option, P pursues both projects. Of key importance is that the entire quality of research, should it take place, is attributed to P. Thus, P's expected payoff is:

$$v_P^{Int} = \max\left[x - c_1 + \max\left[p_F(y - c_{2P}) - C_P, 0\right], 0\right] \tag{1}$$

This is also the social surplus generated from this research. Note that it is entirely possible that project 1 has no stand-alone value (i.e., $x = 0$) and its value rests solely on its ability to lead to research in project 2.

C. Collaboration

Under collaboration, P identifies F ex ante, and they choose to pursue both projects jointly.[5] The first consequence of this is that the costs of understanding the implication of project 1 for project 2 are shared across scientists. To this end, we assume that these costs are C_{PF} and can be allocated to P and F through internal bargaining with P's share being s. Similarly, we assume that the consequent probability that project 2's opportunity is perceived is p_{PF}.

The second consequence of collaboration is that co-authorship is formally given to both P and F on projects 1 and 2. Of course, one can imagine a scenario whereby this is only done with respect to project 2 but, as is explained below, this does not necessarily lead to different conclusions regarding whether collaboration is chosen. If both projects are successful, the research quality of their collaborative effort is $x + y$. However, the market – comprised of scientific peers – will award each with personal attribution of that output. We assume that the attribution going to scientist i is α_i when the market has some basis (perhaps on the basis of other knowledge of roles or reputation) to assign differential weights to each scientist. Otherwise we assume that attribution has to be the same in equilibrium; i.e., $\alpha_P = \alpha_F = \alpha$. Importantly, we make no assumption that

---

[5] There is an issue associated with whether F can be simply identified or not. As we note below, publication can work without this condition.

$\alpha_P + \alpha_F = 1$. Indeed, this could be greater than 1 although the social benefit from the projects can be no greater than $x + y$.

Under collaboration, the expected payoffs to each scientist are:

$$v_P^{Col} = \max\left[\alpha_P x - c_1 + \max\left[p_{PF}\alpha_P y - sC_{PF}, 0\right], 0\right] \tag{2}$$

$$v_F^{Col} = \max\left[\alpha_F x + \max\left[p_{PF}(\alpha_F y - c_{2F}) - (1-s)C_{PF}, 0\right], 0\right] \tag{3}$$

This reflects the notion that P has the lowest cost associated with conducting project 1 and that P and F choose the scientist with the lowest cost to conduct project 2. The allocation of the costs, C_{PF}, is assumed here to be determined internally. To keep things simple, it will be assumed that all of the internal bargaining rests with P and so s is the minimal amount (if it exists) that will ensure that F collaborates.

To see what s will be let's assume that it is jointly profitable for project 1 to be investigated and, individually profitable for project 2 to proceed (i.e., $\alpha_F y \geq c_{2F}$). In this case, from (3), the minimal s that allows F to earn a positive return is:

$$\hat{s} = 1 - \tfrac{\alpha_F x + p_{PF}(\alpha_F y - c_{2F})}{C_{PF}} \tag{4}$$

Thus, for P, its expected return is:

$$v_P^{Col} = (\alpha_P + \alpha_F)(x + p_{PF} y) - c_1 - C_{PF} - p_{PF} c_{2F} \tag{5}$$

Note that the social surplus from collaboration is:

$$\max\left[x - c_1 + \max\left[p_{PF}(y - c_{2F}) - C_{PF}, 0\right], 0\right] \tag{6}$$

Importantly, while the market can potentially award *P* and *F* a 'free lunch' if $\alpha_P + \alpha_F > 1$, the social surplus we will focus on only involves the 'real' variables.

D. Publication

The final organizational option for cumulative science is for *P* to research and publish the results of project 1 and then for another scientist, *F*, to investigate this project outcome and potentially perceive and research on project 2. For *F*, should *P* publish their research from project 1, they will have a choice as to whether to conduct an in-depth investigation of that research and, if that provides an opportunity, research and publish project 2. It is assumed that if project 2 is published that that involves a citation to *P*'s research in project 1. The market will then partially attribute some of project 2 to *P* as a share β_P of y and attribute β_F of y to *F*.

Given this, *F*'s expected payoff following a publication by *P* is:

$$v_F^{Pub} = \max\left[p_F(\beta_F y - c_{2F}) - C_F, 0 \right] \tag{7}$$

If *F*'s expected payoff is zero, we assume here that publication is infeasible as, if they are given the choice, *P* would prefer integration to publication. However, if *F*'s expected payoff is positive and so research into project 2 goes ahead, *P*'s expected payoff is:

$$v_P^{Pub} = \max\left[x - c_1 + p_F \beta_P y, 0 \right] \tag{8}$$

In this case, social surplus from publication is:

$$\max\left[x - c_1 + p_F(y - c_{2F}) - C_F, 0 \right] \tag{9}$$

Again, we assume that even if $\beta_P + \beta_F \neq 1$ the social surplus from project 2 if it is successful is y.

E. Equilibrium Choices

We now turn to consider P's organizational choice for research. Clearly, so long as $v_F^{Pub} > 0$, P will choose $\max\{v_P^{Int}, v_P^{Col}, v_P^{Pub}\}$. One issue of interest is whether these organizational choices will lead to one or both projects being undertaken. In that case, it is clear that the choice is restricted between those options where both projects are likely to be undertaken as an additional project cannot harm either scientist. For this reason, we focus here on the second issue of interest: when all organizational choices can feasibly induce research on projects, what determines which one is chosen? Thus, we assume that (i) $x - c_1 + p_F(y - c_{2P}) \geq C_P$ (both projects are undertaken under integration); (ii) $(\alpha_P + \alpha_F)(x + p_{PF}y) - c_1 - c_{2F} \geq C_{PF}$ (both projects are undertaken under collaboration); and (iii) $x - c_1 + p_F \beta_P y \geq 0$ and $p_F(\beta_F y - c_{2F}) \geq C_F$ (both projects are undertaken under publication). Thus,

$$v_P^{Int} = x - c_1 + p_P(y - c_{2P}) - C_P \tag{10}$$

$$v_P^{Col} = (\alpha_P + \alpha_F)(x + p_{PF}y) - c_1 - C_{PF} - p_{PF}c_{2F} \tag{11}$$

$$v_P^{Pub} = x - c_1 + p_F \beta_P y \tag{12}$$

$$v_F^{Pub} = p_F(\beta_F y - c_{2F}) - C_F \tag{13}$$

It is instructive to consider the pairwise choices between these organizational forms.

Collaboration versus integration: Collaboration will be chosen by *P* if:

$$v_P^{Col} \geq v_P^{Int} \Rightarrow (\alpha_P + \alpha_F - 1)x + \left((\alpha_P + \alpha_F)p_{PF} - p_P\right)y \geq p_{PF}c_{2F} - p_P c_{2P} + C_{PF} - C_P \quad (14)$$

By contrast, collaboration is socially superior to integration if:

$$\begin{aligned} & x + p_{PF}(y - c_{2F}) - C_{PF} \geq x - c_1 + p_P(y - c_{2P}) - C_P \\ & \Rightarrow (p_{PF} - p_P)y \geq p_{PF}c_{2F} - p_P c_{2P} + C_{PF} - C_P \end{aligned} \quad (15)$$

It is clear that the social choice and *P*'s choice will coincide if and only if $\alpha_P + \alpha_F = 1$. Interestingly, this stands in contrast to an often heard informal argument that $\alpha_P + \alpha_F > 1$ either in actuality or normatively. The informal argument suggests that collaboration is difficult to achieve and involves private costs that exceed the social costs of collaboration. Based on this model where collaboration allows knowledge transfer costs (C_{PF}) to be shared between both scientists, overweighting the reward to collaboration would encourage too much collaboration.

Publication versus integration: Publication will be chosen by *P* if:

$$v_P^{Pub} \geq v_P^{Int} \Rightarrow (p_P - p_F\beta_P)y \leq p_P c_{2P} + C_P \quad (16)$$

In this case, publication is socially preferable to integration if:

$$\begin{aligned} & x - c_1 + p_F(y - c_{2F}) - C_F \geq x - c_1 + p_P(y - c_{2P}) - C_P \\ & \Rightarrow (p_P - p_F)y + p_F c_{2F} + C_F \leq p_P c_{2P} + C_P \end{aligned} \quad (17)$$

Note, however, that because we assume that, under publication, *F* will choose to conduct the project 2, $p_F\beta_F y \geq p_F c_{2F} + C_F$. Thus, if publication is chosen by *P* we know that:

$$\begin{aligned} & (p_P - p_F\beta_P)y - p_P c_{2P} - C_P \leq p_F\beta_F y - p_F c_{2F} - C_F \\ & \Rightarrow (p_P - p_F(\beta_P + \beta_F))y + p_F c_{2F} + C_F \leq p_P c_{2P} + C_P \end{aligned} \quad (18)$$

This is a necessary condition for publication to be chosen by *P*. Thus, if $\beta_P + \beta_F \leq 1$, if publication is chosen in equilibrium then it is socially preferable to integration. However, if $\beta_P + \beta_F > 1$, it is possible that publication will be chosen in equilibrium when it is not socially preferable to integration. Specifically, (18) may hold when (17) does not.

Publication versus collaboration: Publication will be chosen by *P* over collaboration if:

$$v_P^{Pub} \geq v_P^{Col} \Rightarrow (1-\alpha_P-\alpha_F)x + ((\alpha_P+\alpha_F)p_{PF} - p_F\beta_P)y \leq p_{PF}c_{2F} + C_{PF} \tag{19}$$

In this case, publication is socially preferable to collaboration if:

$$(p_{PF} - p_F)y + p_F c_{2F} + C_F \leq p_{PF}c_{2F} + C_{PF} \tag{20}$$

As above recall that, $p_F\beta_F y \geq p_F c_{2F} + C_F$. Thus, if publication is chosen in equilibrium by *P*, then

$$\begin{aligned} &(1-\alpha_P-\alpha_F)x + (\alpha_P+\alpha_F)p_{PF}y - p_F\beta_P y \leq p_{PF}c_{2F} + C_F \\ &\Rightarrow (1-\alpha_P-\alpha_F)x + (\alpha_P+\alpha_F)p_{PF}y - p_F\beta_P y \underbrace{- p_F\beta_F y + p_F c_{2F} + C_F}_{<0} \leq p_{PF}c_{2F} + C_F \end{aligned} \tag{21}$$

This is a necessary condition for publication to be chosen by *P*. Thus, if $\alpha_F + \alpha_P = 1$ and $\beta_P + \beta_F \leq 1$ then, if publication is chosen by *P*, it will also be socially preferred to collaboration. However, if $\beta_P + \beta_F > 1$, then it is possible that publication will be chosen in equilibrium when it is not socially preferable to collaboration. Specifically, (21) may hold even when (20) does not hold.

F. Optimal Attribution

The above analysis suggests that setting $\alpha_F + \alpha_P = 1$ and $\beta_P + \beta_F = 1$ may have some desirable properties. However, it also demonstrated that socially sub-optimal outcomes can arise involving each of the three evaluated organizational choices. Given this, here we consider what levels of $(\alpha_F, \alpha_P, \beta_F, \beta_P)$ might generate a socially optimal outcome. The idea is to imagine that these parameters were chosen by a planner and to evaluate their properties to look for mirrors in terms of what actually happens in science. To this end, we follow Green and Scotchmer (1995) by considering situations where the follow-on scientist, *F*, may earn no surplus as a convenient means of avoiding having to deal with the potential range of parameters that may give rise to a socially optimal outcome. The idea here being that since *P*'s choice determines the outcome, it makes sense to ensure that as much of the surplus goes to *P* as possible.

Proposition 1. *There exists setting* $\alpha_F + \alpha_P = 1$ *and* $\beta_P + \beta_F = 1$ *such that P's equilibrium choice is socially optimal.*

The proposition is easily proved by finding the $\hat{\beta}_F$ such that $v_F^{Pub} = 0$ and letting $\hat{\beta}_P = 1 - \hat{\beta}_F$ and substituting $\alpha_F + \alpha_P = 1$ so that:

$$v_P^{Col} \geq v_P^{Int} \Rightarrow (p_{PF} - p_P) y \geq p_{PF} c_{2F} - p_P c_{2P} + C_{PF} - C_P$$

$$v_P^{Pub} \geq v_P^{Int} \Rightarrow (p_P - p_F) y + p_F c_{2F} + C_F \leq p_P c_{2P} + C_P$$

$$v_P^{Pub} \geq v_P^{Col} \Rightarrow (p_{PF} - p_F) y + p_F c_{2F} + C_F \leq p_{PF} c_{2F} + C_{PF}$$

These are identical to conditions (15), (17) and (20). This demonstrates that the choices between each of the organizational forms are driven by the same conditions as the socially optimal choices.

Specifically, note that:

$$\left(\hat{\beta}_P, \hat{\beta}_F\right) = \left(\tfrac{p_F(y-c_{2F})-C_F}{p_F y}, \tfrac{p_F c_{2F}+C_F}{p_F y}\right) \tag{22}$$

while so long as they sum to 1 it is arbitrary what α_F and α_P are. The reason for this is quite intuitive. With collaboration, we allowed *P* and *F* to negotiate cost sharing but given the structure this allowed them to transfer utility. Thus, the decision was driven wholly by the joint market reward to collaboration rather than the precise division. No such instrument existed for publication and hence, the market rewards needed to determine division as well as overall value in order to generate a socially optimal outcome.

It is useful to consider how Proposition 1 might change if, in fact, *s* (the share of costs accruing to *P* under collaboration) was fixed and non-negotiable.

Proposition 2. *When s is fixed, there exists setting* $\alpha_F + \alpha_P = 1$ *and* $\beta_P + \beta_F = 1$ *such that P's equilibrium choice is socially optimal.*

The proof proceeds using the same method as Proposition 1. In this case, the range of β_i remains as in (22) above. By contrast, the market weights for collaboration become:

$$\left(\hat{\alpha}_P, \hat{\alpha}_F\right) = \left(\tfrac{x+p_{PF}(y-c_{2F})-(1-s)C_{PF}}{x+p_{PF}y}, \tfrac{p_{PF}c_{2F}+(1-s)C_{PF}}{x+p_{PF}y}\right) \tag{23}$$

In this case, it can easily be demonstrated that the payoffs realized are the same as in Proposition 1.

G. Remark on Green and Scotchmer

As noted earlier, the underlying structure of cumulative science is based on the Green and Scotchmer (1995) model and notation for cumulative innovation. The contexts and issues dealt with are very distinct, however. Nonetheless, it is useful to note that the publication and associated citation plays a similar role to ex post licensing in their model except that the key parameters are market determined rather than determined through bilateral negotiation. That said, Green and Scotchmer (1995) do bring some of those factors into play when they consider how a planner might set patent length as well as antitrust policy. Nonetheless, this model builds on their insight that, in order to obtain good outcomes for cumulative innovation, pioneers need to have a stake in follow-on innovation. In science as opposed to commerce, this comes from the market allocation of credit attributed to each scientist as opposed to licensing and other factors that allow profit to be transferred from one innovator to another. In so doing, we believe this demonstrates the importance of these scientific institutions in fostering cumulative knowledge creation.

IV. Some Implications

Having constructed a model of the optimal organizational choice for cumulative science, we now turn to consider a number of issues we highlighted earlier to discuss what insight the model gives us into these. We must emphasize that this analysis is suggestive rather than conclusive. Specifically, we do not know what determines the allocation of credit in science. Our formal model tells us what that allocation will look like if it is optimal but, in fact, the processes by which these actually arise have likely been changing over time and are subject to various degrees of informational limitations that will lead to allocation being an inferred outcome rather than a precise outcome.

A. The Matthew Effect

The Matthew Effect is a staple of the sociology of science literature. It has many forms but broadly speaking it says that the more famous scientists (in terms of their past achievements or positions at elite institutions) receive more credit both in citations but also in terms of kudos in collaborative projects. The issue is whether such credit is proportionate to their actual contribution (which may be high for the same reasons they are famous) or disproportionate to that contribution. While there are many informal theories that discuss the potential drivers and plausibility of the Matthew Effect in science there is, to our knowledge, no formal model that derives the Matthew Effect in its disproportionate form as an equilibrium phenomenon.

It is instructive, therefore, to discuss this in the context of the model presented here. Specifically, if the market weights to collaboration and publication are determined optimally as analyzed in the previous sub-section, what does this say about the Matthew Effect? To consider this, let us focus on the pioneer scientist (P). There are several parameters that relate to P's ability to contribute to project 2. There is p_P, the probability that P has an insight that perceives the opportunity of project 2. There is C_P, the costs associated with understanding project 1 that will generate that insight. There is c_{P2}, P's costs associated with carrying out project 2. Finally, in a situation where it is exogenous, there is s, P's contribution to insight in a collaborative venture. The first three of these factors only play a real role when P chooses to pursue both projects themselves. For that reason, they do not drive the market weights P receives in collaboration or publication. The only time a contributive driver for P impacts on a market weight is for, s, when it is exogenous in a collaborative venture. In this case, if s is higher, P will receive more of the weight in kudos associated with collaboration. The conclusion here is that for a pioneer scientist, the Matthew Effect in its proportionate form does not appear to be a clear prediction of this model.

When it comes to a famous scientist's role as a follow-on researcher (*F*), there is more impact. If that scientist finds it less costly to engage in project 2 research (c_{F2}), then the market weights *F* receives under publication and collaboration both rise. Otherwise, drivers that are specific to *F* only impact on the weight a scientist receives under publication; specifically, the higher are p_F and C_F, the diminished is the market kudos flowing back to *P* for a citation. Note here, however, that while it is often said that the Matthew Effect works to provide a famous scientist with more citations here it is operating to deny kudos flowing back to earlier researchers.

It is worth also remarking on the role of *x*, which is the value of project 1 that, here, could only be generated by *P*. One could imagine that the market may perceive *x* as higher for a more famous scientist. In this case, under publication, *P* will clearly appropriate all of the value of *x* and it does not impact at all on the market weights given based on project 2. Under collaboration, it appears that the market would weight *P*'s contribution more highly because of its assessment of *x*. However, while this does change market weights, in actuality, the total kudos (i.e., the weight times the research quality of the projects) never awards any kudos for *x* to *P*. Hence, a change in *x* is neutral when it comes to the Matthew Effect (properly measured).

That said, the market weights for publication and collaboration are, in reality, given by market assessments of the underlying drivers as in (22) and (23). If, because of fame, these assessments are distorted upwards for one scientist, this may have an impact on the relative choices of organizational forms. In particular, a market bias in favour of *P* that is understood by both scientists may render project 2 under publication infeasible for *F* (as their expected surplus was zero based on real variables and with a diminished market weight it will be negative). This would rule out publication is a choice for *P*. In addition, unless they can internally negotiate taking into account such biases, a diminished weight for *F* will render their participation in collaboration infeasible. Thus, a disproportionate weight on a famous *P* would have the effect of pushing organizational choice outcomes towards integration. Ironically, if this were done, the Matthew Effect would not be observable at all as it only arises under non-integrated organizational forms.

The end result here is that the model can tell us some things about the Matthew Effect and its likely form but is incomplete as a means of fully understanding it and explaining its implications as generated by real world experience and data. Of course, this just goes to highlight that a fuller theoretical examination is something that would be a valuable avenue for future research.

B. Collaborative bias

As noted above, it is sometimes claimed that the market weights on a collaborative research project are greater than the weights that would arise if that project were not collaborative. This goes beyond the potential higher quality of such projects (CITE) and to whether, in fact, the market does and should divide the quality of collaborative project by the number of authors when assigning attribution of credit. In this sub-section, we explore what an inflated assessment of $\alpha_P + \alpha_F$ would do to the market weights assigned to publication and to the overall organizational form chosen.

As it turns out, if $\alpha_P + \alpha_F$ is set exogenously (assumed here to be > 1), then only one market weight remains to be determined: β_P. This is because β_F is determined as the value that leaves $v_P^{Pub} = 0$ and so is unchanged from Propositions 1 and 2. The issue becomes that β_P must do two things. It must continue to balance *P*'s incentives with respect to publication versus integration. And it must now re-balance *P*'s incentives with respect to publication and collaboration. It is easy to see that that latter task means that the optimal attribution, $\hat{\beta}_P$, will lie above the levels in Propositions 1 and 2. Thus, as a result of a bias to market weights on collaboration, not only will see observe socially sub-optimal over-collaboration but also over-publication as well. It also suggests that a market bias towards collaboration may itself be a driver of a Matthew Effect in publication and citation. The point here is that these decisions interact and so any analysis of patterns must take into account all of the organizational form options facing scientists.

C. Quantitative measures

Digitization has had the consequence of bringing quantitative measures of impact to science. This is particularly the case with citation counts that are both easily accessible and manipulable. What is lost, however, is a deeper use of citations. What degree of attribution made up a citation? What was the true impact of prior research on follow-on research?

At the same time, increased geographic dispersion, amongst other things, likely means that the roles of individual scientists in collaborative endeavours have become increasingly blurred. Propositions 1 and 2 both suggested that optimal attribution would depend on factors specific to the project but importantly specific to the scientists themselves and their roles (as pioneer and follower respectively). These factors may be less known in more recent science compared to that in the past.

What impact might these trends have on scientific credit and the choice of organizational form for cumulative science? The challenge here is to consider how this might have a systematic bias into these choices as the notion that such choices may not be socially optimal is rather trivial in this context. To that end, we focus here on the new prominence given to citation counts. Basically, the value of a given research output is increasingly measured by the number of citations it receives above and beyond other factors. Within the context of our model here that means that there a systematic increase in β_P if there is follow-on research while there is no necessary trade-off between the level of β_P and the level of β_F.

The direct impact of this, as predicted by the model, would appear to be an increasing bias towards publication as an organizational choice. However, this comes at the expense of collaboration and integration that may be more efficient. It also suggests that scientists will try to increase opportunities for citation through practices such as salami slicing of papers. Nonetheless, beyond these remarks, it is important to realize that the model cannot be taken too far as it does not consider hybrid options such as self-citation and publication followed by collaboration.

V. Conclusions and Future Directions

[To be done]

VI. References [Incomplete]

Adams, J.D. et al., 2005. Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981-1999. *Research Policy*, 34(3), pp.259–285.

Adams, J.D. & Griliches, Z., 1998. Research productivity in a system of universities. *Annales d'Economie et de Statistique*, pp.127–162.

Allen, T.J., 1978. *Managing the flow of technology*, MIT press Cambridge, MA.

Amabile, T.M. et al., 1996. Assessing the Work Environment for Creativity. *The Academy of Management Journal*, 39(5), pp.1154–1184.

Ancona, D.G. & Caldwell, D.F., 1992. Bridging the Boundary: External Activity and Performance in Organizational Teams. *Administrative Science Quarterly*, 37(4), pp.634–665.

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

Beaver, D.B. & Rosen, R., 1978. Studies in scientific collaboration. *Scientometrics*, 1(1), pp.65–84.

Becker, G.S. & Murphy, K.M., 1992. The Division of Labor, Coordination Costs, and Knowledge. *The Quarterly Journal of Economics*, 107(4), pp.1137–1160.

Biagioli, M., 2003. Rights or Rewards. In *Scientific Authorship: credit and intellectual property in science*. pp. 253–279.

Bikard, M., F. Murray & J.S. Gans 2013.

Bordons, M. et al., 1996. Local, domestic and international scientific collaboration in biomedical research. *Scientometrics*, 37(2), pp.279–295.

Bornmann, L. & Daniel, H.-D., 2008. What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), pp.45–80.

Broad, William. (1981). The Publishing Game: Getting more for Less. Science. Mar. 13, 1981, Vol. 211, Issue 4487, p1137-1139

Burt, R.S., 2004. Structural Holes and Good Ideas. *American Journal of Sociology*, 110(2), pp.349–399.

Cattani, G. & Ferriani, S., 2008. A Core/Periphery Perspective on Individual Creative Performance: Social Networks and Cinematic Achievements in the Hollywood Film Industry. *Organization Science*, 19(6), pp.824–844.

Cawkell, A.E., 1976. Citations, obsolescence, enduring articles, and multiple authorships. *Journal of Documentation*, 32(1).

Cronin, B., Shaw, D. & La Barre, K., 2004. Visible, Less Visible, and Invisible Work: Patterns of Collaboration in 20th Century Chemistry. *Journal of the American Society for Information Science & Technology*, 55(2), pp.160–168.

Cummings, J.N., 2004. Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization. *Management Science*, 50(3), pp.352–364.

Cummings, J.N. & Kiesler, S., 2007. Coordination costs and project outcomes in multi-university collaborations. *Research Policy*, 36(10), pp.1620–1634.

Dasgupta, P. & David, P.A. 1984. Towards a new economics of science. *Research Policy*, 23, pp.487-521.

Diehl, M. & Stroebe, W., 1987. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of personality and social psychology* 53.3 (1987): 497-509.

Ding, W.W. et al., 2010. The Impact of Information Technology on Academic Scientists' Productivity and Collaboration Patterns. *Management Science*, 56(9), pp.1439 – 1461.

Dumaine, B., 1994. The Trouble with Teams. *Fortune*.

Dumaine, B. & Gustke, C., 1990. Who Needs a Boss? *Fortune*.

Edmondson, A., 1999. Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly*, 44(2), pp.350–383.

Engers, M., Gans, J.S., Grant, S. & King, S.P., 1999. First Author Conditions. *Journal of Political Economy*, 107 (4), pp.859-883.

Fleming, L., 2007. Breakthroughs and the "Long Tail" of innovation. *MIT Sloan Management Review*, 49(1), p.69.

Fleming, L., Mingo, S. & Chen, D., 2007. Collaborative Brokerage, Generative Creativity, and Creative Success. *Administrative Science Quarterly*, 52(3), pp.443–475.

Furman, J. & Stern, S., 2011. Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *American Economic Review*, 101(5), pp.1933–1963.

Gans, J.S. & Murray, F. 2013. Markets for Scientific Attribution. *mimeo*., MIT.

Gilfillan, S.C., 1935. *Inventing the ship*, Follett Chicago, IL.

Girotra, K., Terwiesch, C. & Ulrich, K.T., 2010. Idea Generation and the Quality of the Best Idea. *Management Science*, 56(4), pp.591–605.

Green, J. & S. Scotchmer (1995), RAND Journal of Economics.

Hara, N. et al., 2003. An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, 54(10).

Hargadon, A., 2008. Creativity That Works. In *Handbook of Organizational Creativity*. Psychology Press.

Hargadon, A., 2003. *How breakthroughs happen: The surprising truth about how companies innovate*, Harvard Business School Press.

Hargadon, A. & Bechky, B., 2006. When Collections of Creatives Become Creative Collectives: A Field Study of Problem Solving at Work. *Organization Science*, 17(4), pp.484–500.

Hargadon, A. & Sutton, R., 1997. Technology Brokering and Innovation in a Product Development Firm. *Administrative Science Quarterly*, 42(4), pp.716–749.

Häussler, C. & Sauermann, H., 2013. Credit Where Credit is Due? The Impact of Project Contribution and Social Factors on Authorship and Inventorship. *Research Policy*, 42(3), pp.688-703.

Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), pp.16569–16572.

Hoerr, J., 1989. The Payoff from Teamwork: The Gains in Quality are Substantial-So Why Isn't It Spreading Faster. *Business Week (European Edition)*, pp.36–42.

Holmstrom, B. 1982. Moral Hazard in Teams. *Bell Journal of Economics*, 13, pp.324-340.

Jiang, L., Thursby, J. and Thursby, M. 2012. Scientific Disclosure and the Faces of Competition. REER Conference Presentation, Georgia Tech.

Johansson, F., 2004. *The Medici effect: breakthrough insights at the intersection of ideas, concepts, and cultures*, Harvard Business Press.

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

Jones, B., Wuchty, S. & Uzzi, B., 2008. Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. *Science*, 322(5905), p.1259.

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

Katzenbach, J.R., Smith, D.K. & Bookspan, M., 1993. *The wisdom of teams*, Harvard Business School Press Boston.

Knorr-Cetina, K., 1999. *Epistemic Cultures: How the Sciences Make Knowledge*, Harvard University Press Cambridge, MA.

Koplowitz, R. et al., 2009. Benchmarking Your Collaboration Strategy. Available at: http://www.forrester.com/rb/Research/benchmarking_collaboration_strategy/q/id/48336/t/2.

Landry, R. & Amara, N., 1998. The impact of transaction costs on the institutional structuration of collaborative academic research. *Research Policy*, 27(9), pp.901–913.

Latour, B. & Woolgar, S., 1986. *Laboratory life: The construction of scientific facts*, Princeton University Press.

Leahey, E., 2007. Not by productivity alone: How visibility and specialization contribute to academic earnings. *American sociological review*, 72(4), pp.533–561.

Lee, S. & Bozeman, B., 2005. The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*, 35(5), pp.673–702.

Maciejovsky, B., Budescu, D.V. & Ariely, D., 2009. Research Note--The Researcher as a Consumer of Scientific Publications: How Do Name-Ordering Conventions Affect Inferences About Contribution Credits? *Marketing Science*, 28(3), pp.589–598.

McAfee, R.P. & McMillan, J., 1991. Optimal Contracts for Teams. *International Economic Review*, 32, pp.561-577.

Melin, G., 2000. Pragmatism and self-organization: Research collaboration on the individual level. *Research Policy*, 29(1), pp.31–40.

Merton, R.K., 1968. The Matthew Effect in Science The reward and communication systems of science are considered. *Science*, 159(3810), p.56.

Merton, R.K., 1988. The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. *Isis*, 79(4), pp.606–623.

Newman, M.E.J., 2001. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences,* 98(2) pp.404-409.

Obstfeld, D., 2005. Social networks, the tertius iungens orientation, and involvement in innovation. *Administrative Science Quarterly*, 50(1), pp.100–130.

Orsburn, J.D. & Moran, L., 2000. *The new self-directed work teams: Mastering the challenge*, McGraw-Hill.

Paulus, P.B., 2007. Fostering creativity in groups and teams. *The handbook of organizational creativity*, pp.159–182.

Paulus, P.B. & Nijstad, B.A., 2003. *Group creativity: Innovation through collaboration*, Oxford University Press.

Porac, J.F. et al., 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), pp.661–678.

Price, D.J. & Beaver, D.B., 1966. Collaboration in an invisible college. *American Psychologist*, 1966, pp.1011–8.

Reagans, R., Argote, L. & Brooks, D., 2005. Individual Experience and Experience Working Together: Predicting Learning Rates from Knowing Who Knows What and Knowing How to Work Together. *Management Science*, 51(6), pp.869–881.

Reagans, R. & Zuckerman, E.W., 2001. Networks, Diversity, and Productivity: The Social Capital of Corporate R&D Teams. *Organization Science*, 12(4), pp.502–517.

Salganik, M.J., Dodds, P.S. & Watts, D.J., 2006. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, 311(5762), pp.854–856.

Simcoe, T.S. & Waguespack, D.M., 2011. Status, Quality, and Attention: What's in a (Missing) Name? *Management Science*, 57(2), pp.274 –290.

Singh, J. & Fleming, L., 2010. Lone Inventors as Sources of Breakthroughs: Myth or Reality? *Management Science*, 56(1), pp.41–56.

Stein, M.I., 1953. Creativity and culture. *Journal of Psychology*, 36(2), pp.311–322.

Stokols, D. et al., 2005. In vivo studies of transdisciplinary scientific collaboration Lessons learned and implications for active living Research. *American journal of preventive medicine*, 28(2S2), pp.202–213.

Subramanyam, K., 1983. Bibliometric studies of research collaboration: A review. *Journal of information Science*, 6(1), p.33.

Taylor, A. & Greve, H.R., 2006. Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal*, 49(4), p.723.

Tushman, M.L. & Katz, R., 1980. External Communication and Project Performance: An Investigation into the Role of Gatekeepers. *Management Science*, 26(11), pp.1071–1085.

Uzzi, B. & Spiro, J., 2005. Collaboration and Creativity: The Small World Problem. *American Journal of Sociology*, 111(2), pp.447–504.

Valderas, J.M., 2007. Why do team-authored papers get cited more? *Science (New York, NY)*, 317(5844), p.1496.

Von Hippel, E.A., 2003. *Democratizing Innovation*. MIT Press Book

Woodman, R.W., Sawyer, J.E. & Griffin, R.W., 1993. Toward a Theory of Organizational Creativity. *The Academy of Management Review*, 18(2), pp.293–321.

Wray, B.K., 2002. The Epistemic Significance of Collaborative Research. *Philosophy of Science*, 69(1), pp.150–168.

Wuchty, S., Jones, B. & Uzzi, B., 2007. The Increasing Dominance of Teams in Production of Knowledge. *Science*, 316(5827), p.1036.

Zare, R.N., 1997. Editorial: Knowledge and Distributed Intelligence. *Science*, 275(5303), p.1047.

Zuckerman, H., 1972. Age, Aging, and Age Structure in Science, reprinted in: Robert K. Merton, 1973. *The Sociology of Science*.