Productivity and Helpfulness: Implications of a New Taxonomy for Star Scientists*

Alexander Oettl†
College of Management
Georgia Institute of Technology‡

June 27, 2008

Abstract

It is surprising that the prevailing performance taxonomy for scientists (Star versus Non-Star) focuses only on individual output and ignores social behavior since, innovation is often characterized as a communal process. To address this deficiency I expand the traditional taxonomy that focuses solely on productivity and add a second, social dimension to the taxonomy of scientists: helpfulness to others. Using academic paper citations to capture scientist productivity and the receipt of academic paper acknowledgements to measure helpfulness, I classify a group of 415 immunologists into four distinct categories of human capital quality: *All-Stars* who have both high productivity and helpfulness; *Lone Wolves* who have high productivity but average helpfulness; *Mavens* who have average productivity but high helpfulness; and *Non-Stars* who have both average productivity and helpfulness. Looking at the change in quality-adjusted publishing output of an immunologist's coauthors after the immunologist's death, I find that the productivity of coauthors of *All-Stars* decreases on average by 35%, coauthors of *Mavens* by 30% on average, and the coauthors of *Lone Wolves* by 19%, all relative to the decrease in productivity of coauthors of *Non-Stars*. These findings suggest that our current conceptualization of star scientists, which solely focuses on individual productivity, is both incomplete and potentially misleading as *Lone Wolves* may be systematically overvalued and *Mavens* undervalued.

JEL Classifications: J24, 049, Z13

Keywords: star scientists, human capital, knowledge spillovers, taxonomy, immunologists

*Special thanks to my dissertation committee of Olav Sorenson, Joanne Oxley, and especially my dissertation chair Ajay Agrawal for immeasurable advice and guidance. I would also like to thank Pierre Azoulay, Ken Corts, Alberto Galasso, Anita McGahan, Samila Sampsa, Scott Stern, Colleen Stuart and seminar participants at the 15th CCC Conference for helpful comments and suggestions. Dr. Jennifer Gommerman of the immunology department at the University of Toronto provided valuable insight. This research was graciously funded by the Social Sciences and Humanities Research Council of Canada (Grant No. 410-2004-1770) and the Martin Prosperity Institute Program on Innovation and Creative Industries. All errors and omissions are my own.

†alex.oettl@mgt.gatech.edu

‡800 West Peachtree St., NW, Atlanta, GA, 30308

1 Introduction

The need to hire the best and the brightest - "the war for talent" - has long been one of the most pressing strategic concerns facing managers (Kapur and McHale 2005, Guthridge, Komm, and Lawson 2008). This concern is largely driven by the observation that high performers, or stars, account for the generation of a disproportionately large level of output. Google's vice-president of engineering, Alan Eustace, noted to the Wall Street Journal in 2005 that "one top-notch engineer is worth 300 times or more than the average" and that he "would rather lose an entire incoming class of engineering graduates than one exceptional technologist" (Tam and Delaney 2005). Why is this? How do stars so greatly influence the performance of organizations?

The existing performance taxonomy for scientists focuses exclusively on individual output, classifying a scientist as either a Star or a Non-Star. The seminal work of Zucker, Darby, and Brewer (1998), for example, defines stars as the top 0.75% of contributors to the genetic sequence database GenBank, a group that accounts for almost 17% of contributions. Recent work by Groysberg, Lee, and Nanda (2008) examines the skill portability of the top 3% of security analysts when they move firms using a ranking of the perceived effectiveness of security analysts, while Azoulay, Graff Zivin, and Wang (2008) look at the impact of eminent scientists using a variety of measures such as research funding, citations, and patenting. In all of these articles, the definition of a star is based solely on individual productivity; in other words, we define stars by what they physically produce.[1]

This uni-dimensional classification of star scientists is surprising as innovation is most often characterized as a communal process. Communal interactions matter for two reasons. First, innovation is more often a result of the recombination of existing knowledge and ideas rather than the discovery of something fundamentally novel (Gilfillan 1935, Nelson and Winter 1982). As knowledge frontiers continue to expand, combinations of increasingly specialized levels of human capital are required to reach the forefront of knowledge (Wuchty, Jones, and Uzzi 2007, Jones 2009). It is this recombination of specialized ideas, either through formal collaborations (coauthorships, joint ventures, etc.) or informal means (discussions and comments from helpful individuals), that leads

[1] While I am unable to directly measure individual productivity, I assume that an individual's inputs are uniform and constant and thus use the term productivity as a measure of an individual's output.

to innovation. Second, the exchange of knowledge is to a large extent governed through social channels. Individuals possess only finite levels of knowledge, and knowledge search is costly. Social forces can reduce barriers to knowledge flow through geographic proximity (Jaffe, Trajtenberg, and Henderson 1993), labor mobility (Almeida and Kogut 1999, Oettl and Agrawal 2008), interpersonal networks (Singh 2005), and membership in ethnic communities (Agrawal, Kapur, and McHale 2008).

While innovation is a communal process, the inability for parties to perfectly contract on knowledge exchange leads to failures in the market for knowledge and a decrease in knowledge transfer (Arrow 1962). As such, conditions that facilitate knowledge sharing or spillovers in the absence of formal contractual environs are of great value to firms. Ultimately, if our concern is to understand the mechanisms by which an individual maximizes his performance, simply understanding the productivity inputs of an individual would suffice. However, the strategy and economics literatures focus on performance measures at the organization and regional levels, and as such, mechanisms in which individuals influence the productivity of others become important as these mechanisms directly influence the performance of organizations and regions. Hence, mechanisms by which individuals improve the performance of others are of paramount concern to scholars of strategy and economics.

The importance of social factors on innovation illuminates the deficiency of our current productivity-focused conceptualization of star scientists (Stars versus Non-Stars). To expand our current conceptualization of star scientists, I develop a new taxonomy of star scientists by incorporating a social dimension: helpfulness to others. This new taxonomy allows an individual to not only vary along a productivity dimension but also a helpfulness dimension.

The objective of this chapter is threefold. First, I expand upon the current dichotomous conceptualization of stars by developing a taxonomy that not only incorporates a star's individual productivity but also his helpfulness. In doing so, I move beyond the current uni-dimensional classification and redefine what it means to be a star. Second, I propose a measure to classify individuals into this new taxonomy. Third, I use this taxonomy to assess the extent to which different star types influence the productivity of others.

Following prior studies (Allison and Long 1990, Azoulay, Graff Zivin, and Wang 2008), I measure individual productivity using Impact Factor-weighted publication counts.[2] On the other hand, I measure helpfulness by academic journal acknowledgements, since such acknowledgements are generally made to those who have helped in the development of the work. Using these measures of productivity and helpfulness, I classify a sample of 415 immunologists and examine their influence on the productivity of their coauthors. I use coauthorship to pinpoint the timing of the formation of an interpersonal tie between an immunologist and a potential recipient of performance benefits. It is this co-location in social space that allows stars to impact the performance of their peers.

By placing a star in both productivity and helpfulness space while keeping the classifications discrete, I am able to classify an individual as one of four types: All-Star, Lone Wolf, Maven, or Non-Star.

I define an All-Star as an individual with both high productivity and high helpfulness. A Lone Wolf is someone who has high productivity but average helpfulness. A Maven is an individual with average productivity but high helpfulness, and a Non-Star has both average productivity and average helpfulness. Restrictively, the current dichotomous conceptualization of stars groups both All-Stars and Lone Wolves together, while completely overlooking Mavens. By expanding on the current classification, I am able to examine the influence of individuals who vary in both their productivity and helpfulness.

Examining the changes in productivity from coauthoring with various star types would be an appropriate empirical exercise if coauthoring relationships were chosen at random, but clearly they are not. The problem with endogenous coauthor selection is that the coauthors selected by an immunologist may be chosen due to their own productivity, thus producing spurious correlations between an individual's productivity and their coauthorship network. For this paper, I examine the *decrease* in productivity of coauthors when an immunologist dies.

Across a number of specifications, the performance of the coauthors of All-Stars who die decreases on average by 35% relative to the decrease in performance when a Non-Star dies. More

[2] The Impact Factor is a time-varying journal-level measure of quality, which captures the rate at which each article in the journal is cited. Thus, journals with articles that are cited more often will have higher Impact Factors. The Institute for Scientific Information (ISI), a subsidiary of Thomson Scientific, constructs this measure annually.

interestingly, coauthors of Mavens who die experience a 30% decrease in performance, while the coauthors of Lone Wolves who die experience decreases in performance of only 19% on average.

By expanding the current conceptualization of star scientists and focusing on both the productivity and helpfulness dimensions of scientists, I find that spillovers are most likely generated from individuals with high helpfulness. As a result, the literature has largely overemphasized the importance of Lone Wolves yet has overlooked and consequently underemphasized Mavens.

2 Star Scientists and Spillovers

While the quality of human capital may be uniformly distributed, the returns to human capital are wildly skewed (Ernst, Leptien, and Vitt 2000). Individuals in the right tail of the distribution, so-called stars, generate a disproportionately large share of output (Rosen 1981). As Lotka (1926) observed, the top 6% of physicists produce more than 50% of all papers. This skewed distribution - termed the Pareto Principle - is ubiquitous across industries and is a strong determinant of inventive productivity (Narin and Breitzman 1995). However, as I argue, even though stars contribute disproportionately to output production, they cannot alone act as a source of sustained competitive advantage.

In the resource-based view (RBV) of the firm, firms generate sustainable competitive advantage through their use of strategic resources (Wernerfelt 1984). Resources are only sources of sustainable competitive advantage, however, if they are valuable, rare, inimitable, and difficult to substitute (Barney 1991). In efficient factor markets, a star will be perfectly compensated for his productivity (Hirshleifer, Glazer, and Hirshleifer 1998), suggesting that they cannot be a source of rents for the firm. The resource-picking literature, nonetheless, theorizes that firms can capture economic rents by employing superior information or analysis to pick undervalued resources in the factor market, much in the same way that a fund manager attempts to outsmart the financial markets by picking stocks (Barney 1986, Makadok 2001). The central assumption of this resource-picking mechanism is that the factor markets must be characterized by imperfect information, thus providing a forum for a firm with superior information to pick undervalued resources. While this situation may well exist in a number of factor markets, it seems unlikely to exist in the market for stars, as the output of stars is both easy to measure and highly visible. Consequently, the likelihood of a star being mispriced is low, rendering the resource-picking mechanism ineffective at generating sustainable competitive advantage. Furthermore, because of the high visibility of the output of stars, stars themselves are more mobile, further attenuating the value of stars as resources (Lazear 1986).

Human capital, however, can be important for firm strategy in ways other than generating direct output: human capital can generate spillovers. Since the early work of Lucas (1988) human capital spillovers have been at the center of economic growth models. Lucas classifies human capital into two types: internal and external. Internal effects of human capital capture the extent to which human capital affects the individual's own productivity, while the external effects of human capital capture the influence individuals have on the performance of others. If these external effects generate an unpriced spillover onto the productivity of others, then the spillover constitutes an externality (Acemoglu 1996). The notion that these human capital externalities and their effect on the increase in knowledge stocks can lead to increasing economic returns is captured by the endogenous growth theory of Romer (1990). As knowledge flows and spillovers lie at the center of many of our models of innovation (Audretsch and Feldman 1996) and as human capital externalities are a key input in the generation of knowledge flows, understanding the parameters within which human capital spillovers are generated is of utmost importance. In addition, from a firm strategy standpoint, human capital spillovers may be more difficult to observe, thus allowing for the possibility of resource picking and conversely serving as a source of sustainable competitive advantage.

Despite the importance of human capital spillovers, the strategy and economics literature has mostly focused on the skewed nature of the productivity distribution when examining the relationship between stars and performance. The seminal work of Zucker, Darby, and Brewer (1998) reports strong correlations between the location of star scientists and the formation of biotechnology ventures. In more recent work, Groysberg, Lee, and Nanda (2008) examine the firm specificity of the human capital of stars. They find that much of the performance premiums accruing to stars is firm specific and that when stars move, their productivity decreases. In a related work, Brown (2008) explores the relationship between effort exertion and differences in relative ability. Using a sample of professional golfers, Brown finds that golfers exert less effort in the presence of a star (Tiger Woods), thus indicating that the presence of stars can have a negative effect on organizational outcomes. None of these studies, however, explicitly examines these stars' human capital spillovers.

One notable and important exception is the work of Azoulay, Graff Zivin, and Wang (2008), henceforth referred to as AGW. AGW examine the effect that the death of an eminent life scientist has on the performance of his coauthors.[3] They find that, following the death of a star, coauthors' performance decreases by 5% to 10%. Since coauthors were benefiting from their tie with a star, the cessation of the coauthoring relationship (because of the death) ended these benefits, resulting in a decline in coauthor performance. Their study, however, is unable to rule out two possibly conflating effects. First, because their sample consists only of the top 5% of life scientists who died and their analysis solely examines the influence of those deaths on the performance of their coauthors, they are unable to examine the impact non-stars have on the performance of others, as non-stars are not included in the sample. As such we are unable to determine how much greater the performance benefits from stars are than non-stars. They do show evidence that coauthor performance gains are increasing in a scientist's citations, but this is still conditioned on being in the top 5%. Second and in a related vein, without the inclusion of different star types in their sample, they are unable to disentangle the decrease in performance of coauthors that is due to the loss of an intellectual link versus the decrease in performance arising from the emotional toll of the death of a former colleague. Lastly, they do not have access to pricing data and as such cannot properly determine whether or not the changes in the performance of their sample's coauthors are spillovers.

This paper builds upon and extends the pioneering work of AGW by developing a new taxonomy of star scientists that allows for more precise identification of star scientists of different productivity types and examines the conditions under which stars are more likely to impact the performance of their coauthors.

[3]AGW classify a scientist as eminent if he matches a number of performance-related criteria. In general, one can view these life scientists as being in the top 5% of the life scientist productivity distribution.

3 A New Taxonomy of Star Scientists

The difficulty in finding empirical examples of human capital externalities is due to the uneven distribution of the externality generating process. In others words, not all individuals have an equal impact on the performance of others. The current strategy and economics literature classify stars along a single dimension - productivity. That is, an individual is classified a star if he falls in the right tail of some productivity distribution, normally output. For example, Zucker, Darby, and Brewer (1998) classify the top 0.75% of GenBank contributors as stars. Yet if social behaviors influence the impact that stars have on the performance of others, then including a dimension of social behavior in our conceptualization of star scientists is surely needed. I extend our current conceptualization of star scientists that solely focuses on productivity and add a second, social dimension to the taxonomy of scientists: helpfulness. Where productivity encapsulates an individual's output that is beneficial to himself, helpfulness encapsulates an individual's output that is beneficial to others.

Examining the helpfulness of individuals in organizations is a well-trodden research stream. A large amount of literature on organizational psychology examines what is known as Organizational Citizenship Behavior (OCB) (Smith, Organ, and Near 1983). The literature finds that a combination of altruism and courtesy greatly influences the level of helpfulness individuals extend to one another within organizations. A large literature in social psychology exists on the personality characteristics associated with helpful behavior. Among the many factors that influence an individual's helpfulness, three are most applicable to the setting of academic scientists: situational, social, and person factors. Situational factors deal with the costs associated with helping, social factors involve the influence of social norms on helpful behavior, and person factors capture the prosocial traits of an individual. I believe that holding situational and social factors constant within my setting of academic scientists is acceptable, and as such, person factors (which are innate) should be the only traits that influence helpfulness within my study (Fletcher and Clark 2003). Consequently, I take helpfulness to be an innate, continuous measure that captures an individual's output that is beneficial to others.

To provide a more concrete description of what helpfulness in the sciences entails, below are some quotes from the obituaries of Maurice Landy and D. Bernard Amos, two scientists who are members of my empirical sample and both of whom are in the top 5% of the helpfulness distribution.

On Maurice Landy:

> "During his last year, Maurice Landy was a member of my laboratory, where he gave generously of his wisdom and experience. He was particularly attentive to younger scientists, teaching them to present their work in its optimal light and to respect and critically enjoy the work of others. Friends are valued for much more than just their contributions to knowledge. Thusly, Maurice, as a friend, turned our failures into successes, our parochialisms into worldliness, and our desperations into hopes. He demanded the best out of us and enjoyed our accomplishments as he might his own. And, in the end, he left a memory that we all cherish." (Lawrence and Cohn 1993)

On D. Bernard Amos:

> "Bernard has had a profound impact on many individuals during his life. He has been instrumental in the training, education and development of generations of clinical and basic scientists." (Tedder and Dawson 2003)

and

> "Only two months before his death I asked him about a research problem I had encountered testing individuals exposed to tuberculosis that have a negative skin test. He remembered his NLT work of 40 years earlier and suggested that I look for antibodies to Class II and antibodies to tuberculin that could prevent delayed hypersensitivity reactions tested by intradermal inoculation of tuberculin." (Yunis 2004)

In comparison, the obituaries of Zanvil Cohn and Philip Gell, two immunologists in the top 5% of the productivity distribution but not the top 5% of the helpfulness distribution, focus more on the scholastic accomplishments of the two scholars. On Zanvil Cohn:

> "His incisive manner, his admiration of clever new experiments, his sense of fairness and respect, and his wit all will be sorely missed." (Steinman and Moberg 1994)

On Philip Gell:

> "He was one of the founders of the British Immunological Society and started the first postgraduate M.Sc. course in immunology in 1963, which is still running today. In this course, and with a stream of postdoctoral fellows in his own laboratory, Gell helped to influence many of the basic and clinical scientists, both British and international, who lead the field today. He was elected Fellow of the Royal Society in 1969. (Silverstein and Benacerrafe 2001)

Table 1 presents a new taxonomy for star scientists that incorporates productivity and helpfulness. Not only does a scientist vary along a dimension of productivity, he also varies along a measure of helpfulness. In doing so, I define three new star types. An All-Star is an individual with both high productivity and helpfulness. A Lone Wolf is an individual with high productivity but average helpfulness. A Maven is an individual with average productivity but high helpfulness. A Non-Star, has both average productivity and helpfulness.

Why does this taxonomy matter? Conventionally, both All-Stars and Lone Wolves are classified as stars as they both have high productivity. This aggregation has large strategy implications if the effects of All-Stars and Lone Wolves on organizations vary. Mavens on the other hand are currently classified as individuals with average productivity. But Mavens may have the largest impact on the performance of others due to their level of helpfulness. As such, we may be overvaluing Lone Wolves while undervaluing Mavens. Given that human capital spillovers are at the core of our innovation and economic growth models, it is paramount to identify which inputs into the economic production function have the potential to generate spillovers.[4]

4 Data

An ideal empirical setting for this study satisfies three criteria. First, it should take place in an organizational setting where collaboration exists. As the goal of this paper is to identify which

[4] To be clear, this study cannot make the claim that the positive performance impact of a star on his coauthors is a spillover as pricing data are not available for the sample studied. Chapter 3, however, includes preliminary evidence that these performance gains may indeed be uncompensated.

types of individuals have the largest impact on the performance of their coauthors, a setting in which collaboration takes place is clearly necessary. Second, from a measurement standpoint, the ability to separate individual from group or organizational level performance is necessary as the focus of interest is on star individuals and not star teams or firms. Third, a field or discipline that engages in the practice of manuscript acknowledgements is necessary to identify individual helpfulness, which I will discuss in more detail later in this dissertation. A discipline that satisfies all three of these conditions is the field of immunology.

From a research standpoint, immunology is an incredibly important discipline. The National Institute of Allergy and Infectious Diseases (NIAID), which oversees the distribution of immunology-related research grants, allocated $940 million to immunology research in 2005, up from $646 million in 2003 (Hackett, Rotrosen, Auchincloss, and Fauci 2007). More importantly, however, the structure of immunology research is organized in a very similar fashion to other medical sciences, such as biochemistry, microbiology, and pharmacology.

4.1 Measures

One major hurdle to extending the dichotomous conceptualization of stars has been the lack of data. I propose to use the receipt of acknowledgements as a measure of an individual's helpfulness. Academic acknowledgements are a central and convenient way of recognizing a non-author's contributions to the development of a manuscript without extending ownership rights in the form of coauthorship (Merton 1973).[5]

The goal of this study is to classify an immunologist along the dimensions of productivity and helpfulness and then to examine the change in output of his coauthors when he dies. I measure productivity as the total number of citations received for papers written by the focal immunologist prior to 1966. Citation data come from the Institute for Scientific Information (ISI) Web of Science. I measure helpfulness as the total number of acknowledgements received by the focal immunologist

[5] Of course, acknowledgements can come in two forms. They may represent an acknowledgement of another author's useful comments (that is, the author is selected on quality) or they may accrue as a result of the author's influence on the publishing process, the field, etc. (the author is selected on status). While I am unable to empirically separate out these two types of acknowledgements, "status" acknowledgements should add noise to the empirical analysis and thus, due to attenuation bias, result in conservative estimates. I discuss this further in Section 8.

between the years 1960 and 1965 (inclusive) in *The Journal of Immunology.* I choose *The Journal of Immunology* because during this time period it was the pre-eminent academic journal for the discipline of immunology.[6] Acknowledgements operate very similarly in immunology as they do in the social sciences, albeit with fewer acknowledgements per paper. Of the 1,324 articles published in *The Journal of Immunology* between 1960 and 1965, 50% had at least one acknowledgement. Of the articles that did have at least one acknowledgement, 40% of them had at least one acknowledgement for criticism and encouragement, the measure used for this study.[7] As an example, the following was in the acknowledgement section of Bennett (1965): "The author wishes to thank Drs. L. J. Old and E. A. Boyse of the Sloan-Kettering Institute, New York, for their suggestions and encouragement, and Mrs. Patricia Hubertus for technical assistance."

I measure a coauthor's productivity by their Impact Factor-weighted publications. I obtain Impact Factor weights from the Journal Citation Reports from the ISI, which published Impact Factors for all immunology journals on a yearly basis between 2000 and 2007. I use the average Impact Factor across these eight years to create a time-invariant quality measure of the 136 immunology journals indexed by ISI.

I collect data on deaths in a hybrid form by way of extracting obituaries and memoriams from the titles of over 400,000 immunology articles from the Web of Science as well as through manual Internet searches. While ideally I would like to identify unexpected deaths so that the "treatment" of losing a coauthor is fully exogenous, none of the deaths in my sample are of this nature and as such may not be fully exogenous. Allowing for the possibility that the deaths were anticipated by coauthors should generate conservative estimates of the productivity effect, as presumably the coauthors had time to make alternate arrangements to minimize the anticipated decrease in productivity. As such, the regression estimates should be viewed as the changes in productivity net of an anticipation of death.

[6] *The Journal of Immunology* in 2007 had an Impact Factor of 6.068, ranking it 13th among all immunology journals. It is, however, by far the most widely cited journal in immunology and has been in print since 1916, making it one of the oldest immunology journals in the world. Furthermore, the *Journal of Immunology* was chosen at random, and I have no reason to believe that immunologists providing feedback or criticism alter their behavior on publications intended for the *Journal of Immunology.*

[7] Acknowledgements thanking lab technicians and assistants are removed, although my sampling process requiring at least three lifetime publications most likely would have removed these individuals anyway.

4.2 Sample

The sample for this study draws on all immunologists who published at least one article in an immunology journal between 1960 and 1965, inclusive.[8] There are 5,323 of these scientists. I apply the requirement constraint that every immunologist must have at least three lifetime papers in an attempt to remove post-doctoral students, graduate students, and any other scientists who did not become academic scientists. This reduces my sample to 1,543 scientists. This remaining set of scientists must meet one final criterion: they must have at least one coauthoring relationship, formed after 1965, with a scientist who also has at least three lifetime publications. After applying this final condition, I am left with a final sample of 415 immunologists.

I divide my sample of 415 immunologists into four discrete categories based on their location in productivity and helpfulness space. Figure 1 graphically shows the sample's placement. While the constructs of productivity and helpfulness are continuous measures, the new taxonomy for star scientists requires discrete allocations and as such a cut-off point must be established to discern between high and average levels of productivity and helpfulness. The goal of this taxonomy is not to quibble about cut-off points but rather to demonstrate that an individual at the extreme end of one distribution is of a different type than the average individual in the rest of the distribution. Therefore, I define high productivity and helpfulness as being in the top 5%, which is a similar cut off to other studies looking at stars (Azoulay, Graff Zivin, and Wang 2008). An immunologist has high productivity (in the top 5%) if he receives more than 2,028 citations prior to 1965, and an immunologist has high helpfulness (in the top 5%) if he receives three or more acknowledgements in *The Journal of Immunology* in the six years between 1960 and 1965, inclusive.

All-Stars are immunologists who have both high productivity and helpfulness (upper right quadrant of the graphic). Mavens are immunologists with average productivity but high helpfulness (upper left quadrant). Lone Wolves have high productivity but average helpfulness (bottom right quadrant). Non-Stars have both average productivity and helpfulness (bottom left quadrant). I show the classification of the sample in tabular format in Table 2.

[8] I draw the list of immunology journals from the Thomson Corporation's ISI Web of Science database. While I include the major immunology journals, such as *The Journal of Immunology* and *The Journal of Experimental Medicine*, in this list, general field journals such as *Nature* and *Science* are not. Consequently, I do not use "general" journals when constructing measures and defining the sample.

Assigning the sample of 415 immunologists results in the following classification: four scientists are All-Stars, five are Mavens, 16 are Lone Wolves, and 390 are Non-Stars.[9] Of these 415 immunologists, 28 of them have died: two All-Stars, three Mavens, five Lone Wolves, and 18 Non-Stars.

4.3 Unit of Analysis

To what extent do different star types influence the productivity of others? To answer this question, I look at the change in performance of coauthors of stars who die. As such, my unit of analysis is an immunologist-coauthor-year triad. The cross-sectional unit, however, is the immunologist-coauthor dyad, where the immunologist is one of four star types.[10] To identify coauthors, I identify all coauthorships formed after 1966 with scientists who have at least three lifetime publications. The immunologists in the sample have 58.3 coauthors on average, resulting in 24,175 immunologist-coauthor dyads.[11] The average publishing lifespan for immunology coauthors in my sample is 23.8 years, resulting in a final sample size of 575,483 observations. For the 28 immunologists who die, I reduce the sample to 816 dyads generating a subsample of 25,968 observations.

Using 1965 as a cutoff for both the productivity and helpfulness measures allows me to hold constant each scientist's "type" for 1965.

While the new taxonomy for star scientists is meant to classify an individual at a given point in time, for this paper I assume that distributions of skill along both the productivity and helpfulness dimensions are innate and thus do not vary across my sample. By looking at all coauthorships formed after 1965, I hold the window of evaluation constant for all star types and only focus on a star's influence on the performance of all new coauthors, thus reducing the likelihood that an im-

[9]The observant reader will notice that while the high productivity stars (All-Stars and Lone Wolves) account for 5% of the total sample ($\frac{4+16}{415}$), the high helpfulness scientists do not ($\frac{4+5}{415}$). The reason for this is that because of the discrete nature of acknowledgements and the requirement that membership in the top 5% be *greater than* the 95th percentile cut-off values, I classify fewer than 5% of all scientists as highly helpful. When I change the cut-off designation to define scientists in the top 5% as scientists with *greater than or equal to* the 95th percentile cut-off values, the results are statistically and quantitatively largely unchanged.

[10]Non-Stars, the fourth type of star, are of course not truly stars, but for ease of classification I will consider them as one of the four star types.

[11]What happens if a scientist is the coauthor of multiple focal immunologists? While 78% of the coauthors in the sample only coauthor with one of the focal 415 immunologists, 15% coauthor with two immunologists in the sample, 4% coauthor with three immunologists in the sample, and 3% coauthor with four or more immunologists in the sample. I can adjust standard errors to account for this serial correlation. In practice, however, standard errors that have been adjusted for both immunologist and coauthor serial correlation differ marginally from standard errors adjusted solely for immunologists.

munologist's productivity and helpfulness is a function of the coauthor's performance. In addition, by only looking at newly formed coauthorships after immunologist types have been established in 1965, I remove the conflation of star definition with the performance of coauthors in that I only examine the performance impact of scientists who are not coauthors of the star at the time of classification. Furthermore, to reduce the conflation of star type classification with the productivity lifecycle of scientists, I include various age group dummies for both immunologists and coauthors in all regression models. I further discuss estimation and control variables in Section 5.

4.4 Descriptive Statistics

Tables 3 and 4 present descriptive statistics and correlation matrices for both the full sample of 24,175 dyads and the subsample of 816 dyads where the focal immunologist dies. For the full sample (Table 3), the average coauthor publishes 9.737 Impact Factor weighted publications a year, while the average coauthor of an immunologist who dies publishes on average 11.029 Impact Factor weighted publications a year. Table 5 presents means of four performance measures split both by the type of star the focal immunologist in the dyad is and whether or not the immunologist died. The coauthors of All-Stars and Lone Wolves have higher average Impact Factor weighted papers than Mavens and Non-Stars, but the coauthors of Mavens receive, on average, the most citations for their papers. Across almost all star types and measures, the coauthors of immunologists who die have higher output, on average, than the coauthors of immunologists who do not die.

5 Econometric Estimation

The empirical objective of this study is to examine the extent to which different star types influence the performance of others. As discussed earlier, a star has the ability to influence the performance of individuals across multiple levels: coauthors, peers in the same department, peers within the same institution, etc. For this study, I solely focus on a star's influence on the performance of his coauthors.[12] The most straightforward empirical approach would be to examine the change in a

[12]In a working paper, Waldinger (2008) explores changes in the productivity of scientists (peer effects) from the exogenous dismissal of colleagues in Nazi Germany. He explores peer effects at three levels: the department level, the same specialization within the same department, and the coauthor level. He only finds evidence of peer effects

coauthor's performance after the formation of the coauthoring relationship (i.e., after the first time the two scientists collectively author a paper). Unfortunately, both the decision to coauthor at all and the decision of whom to coauthor with are clearly not random decisions. This endogeneity would bias my regression coefficients as the choice of coauthors may very well be related to their future productivity, resulting in a spurious relationship between a coauthor's productivity and his coauthorship network. As such, the empirical challenge becomes finding an exogenous change in the coauthoring relationship. An alternative to examining the formation of coauthoring ties is to examine the cessation of coauthoring ties but one that is exogenous. For this paper, I examine the change in productivity of a coauthor when an immunologist dies.

The empirical model to be estimated is

$$Y_{-ijt} = exp[\beta_1 Death_{it} + \beta_2 Death_{it} \times AllStar_i + \beta_3 Death_{it} \times LoneWolf_i + \\ \beta_4 Death_{it} \times Maven_i + \gamma_{it} + \mu_{jt} + \delta_t + \phi_{ij} + \varepsilon_{ijt}] \quad (1)$$

Since my objective is to capture the change in performance of a coauthor after an immunologist's death, the dependent variable, $Y_{-i,j,t}$, measures the number of Impact Factor weighted publications coauthor j wrote in year t where star i is *not* a coauthor. I use quality adjusted publication counts instead of raw publication counts to ensure that I am observing changes in the quality of publishing rather than changes in the frequency of publishing. $Death_{ijt}$ is an indicator variable that switches to 1 the year immunologist i dies. β_1 captures the net change in productivity of coauthor j after star i dies, irrespective of his star type. β_2, β_3, and β_4 captures the change in productivity of coauthor j if immunologist i is an All-Star, Lone Wolf, or Maven, respectively. I omit the Non-Star category, and so the coefficients of β_2, β_3, and β_4 should be interpreted as the change in productivity relative to the productivity change when a Non-Star dies. Because the star types of i are time invariant, they can only be identified through the interaction with *Death*. γ_{it}, and μ_{jt} are sets of age cohort dummies that capture the changes in research productivity across the academic lifecycle (Levin

at the coauthor level.

and Stephan 1991).[13] I capture time effects with δ_t. ϕ_{ij} is a series of dyad fixed effects, which in practice are conditioned out during estimation and as such are not directly estimated. ε_{ijt} is an identically distributed error term but not independent. Errors are correlated due to star i's death affecting all of his coauthors at the same time. Clustering of the standard errors by the star will correct for this non-independence during estimation. If the coefficients on β_1 through β_4 are less than zero, then the death of star i has a negative influence on the performance of coauthor j, which provides some evidence that star i is indeed influencing the performance of coauthor j.

The identification of $Death_{ijt}$ comes from the variation in the deaths of immunologist i. By employing dyad fixed effects, I capture all time invariant attributes common to the dyad by these fixed effects, forcing the parameters to be solely identified from within dyad variation. Because of the count nature of the dependent variable and the high percentage of zero values (33%) across the sample, a count model is most appropriate. Specifically, I employ the Fixed Effects Poisson (FEP) estimator developed by Hausman, Hall, and Griliches (1984). Apart from being computationally straightforward, the Fixed Effects Poisson estimator estimated via quasi maximum likelihood (QML) has strong robustness features, even allowing for consistent parameter estimates of non-count dependent variables (Wooldridge 2002). In addition, standard errors can be made robust to deviations from the poisson distribution, in particular the equality requirement of the first and second moments (Wooldridge 1999). I report these robust standard errors for all QML specifications.

In addition to the QML Fixed Effects Poisson specification used, I implement a linear model of the following form for robustness checks:

$$\begin{aligned}\log \tilde{Y}_{-ijt} = d_{ijt} + \beta_1 Death_{ijt} + \beta_2 Death_{ijt} \times AllStar_i + \beta_3 Death_{ijt} \times LoneWolf_i + \\ \beta_4 Death_{ijt} \times Maven_i + \gamma_{it} + \mu_{jt} + \delta_t + \phi_{ij} + \varepsilon_{ijt} \quad (2)\end{aligned}$$

where the new dependent variable $\tilde{Y}_{-ijt} = Y_{-ijt}$ if $Y_{-ijt} \geq 1$, and $\tilde{Y}_{-ijt} = 1$ if $Y_{-ijt} = 0$. To distinguish between the values of $\tilde{Y}_{-ijt}$ where $Y_{-ijt} = 1$ and where $Y_{-ijt} = 0$, I create a variable

[13] In practice, I generate these age cohort dummies in four-year intervals, whereby the first dummy will capture a scientist in his first four years, the second dummy will capture a scientist in his fifth through eighth year, etc.

d_{ijt} where $d_{ijt} = 1$ if $Y_{-ijt} = 0$ and add it as an independent variable. This technique allows for straightforward interpretation of the coefficients as well as allowing the variables of interest to be interacted with continuous variables (Pakes and Griliches 1980, Acemoglu and Linn 2004).[14] In addition, by adopting a linearization of the functional form of the regression, I can carry out instrumental variable analysis through two-stage least squares estimation (2SLS), which is much more cumbersome in non-linear models such as the poisson outlined in Equation 1.

I present both specifications (Equations 1 and 2) across two main samples. The first sample only includes dyads in which a star dies. Twenty-eight of the 415 immunologists die, and they each have an average of 29.1 coauthors, resulting in a sample consisting of 816 dyads. The second sample includes all immunologist-coauthor dyads, regardless of whether the immunologist dies or not. The full set of 415 immunologists has, on average, 58.3 coauthors, resulting in a sample of 24,175 dyads. Across both specifications, the regression analysis follows a differences-in-differences style estimation. For the sample consisting solely of dyads where the immunologist dies, I use variation in the death of immunologist i to estimate the relationship between death and the change in performance of coauthor j. For the sample consisting of all dyads, I introduce a second dimension of variation, comparing the possible disparity between scientists whose coauthors die and those whose do not. Since only 7% of the 415 immunologists in the sample die (the "treated" group), the remaining 93% in the full sample serve as a de facto control group.[15]

6 Results

6.1 Main Results

This study asks two main research questions. First, using the old dichotomous definition of star scientists, do productivity stars have a larger impact on the performance of their coauthors than Non-Stars? Second, using my new taxonomy of star scientists, to what extent do different star types influence the performance of others? The first question was already asked and partially answered in

[14]The magnitude of interaction effects in nonlinear models do not equal their marginal effects and thus makes direct interpretation of results difficult (Ai and Norton 2003).

[15]In addition, the control group allows me to estimate the cohort effects more precisely.

the very important work by Azoulay, Graff Zivin, and Wang (2008) (AGW). As mentioned earlier, however, the empirical results that AGW present, whereby a scientist's performance decreases after the death of a coauthor, are unable to rule out two possibly conflating effects. First, because AGW condition their sample on the top 5% of life scientists and solely examine the influence of their deaths on the performance of their coauthors, they are unable to examine the impact of non-stars on the performance of their coauthors, as non-stars are omitted from their sample. They do show evidence that as a scientist's citations increase so does his impact on the performance of his coauthors, but this still is conditioned on being in the top 5%. Second and similarly, without the inclusion of different star types in their sample, they are unable to disentangle the decrease in performance of coauthors due to the loss of an intellectual link rather than the decrease in performance due to the disruption caused by the death of a former colleague.

Before I explore the extent to which different star types influence the performance of others, I feel it prudent to address the question of the extent to which stars have a larger impact on the performance of their coauthors than non-stars, first by replicating AGW's results and then by including controls for non-stars to appropriately deal with the alternate explanations outlined above.

Table 6 presents the main results of Equation 1 from section 5. To allow for readily comparable coefficients to those presented by AGW, the sample used in Table 6 only includes dyads where the immunologist dies. Specification 1 further restricts the sample by only including immunologists who have high productivity, that is, All-Stars and Lone Wolves. This definition and sample are identical to those used in AGW. The coefficient on death is -0.173, which translates into a 15.9%[16] decrease in performance, somewhat larger than AGW's range of estimates between -5 and -10%. Specification 2 returns the sample to all dyads with a death. The death variable is interacted both with immunologists who have high (All-Stars and Lone Wolves) and average (Mavens and Non-Stars) productivity. The coefficient on the high productivity interaction, significant at the 5% level, indicates that the performance of coauthors of high productivity stars decreases by 17.4% (-0.191), again somewhat larger than the findings of AGW. The coefficient on average productivity

[16] $exp(-0.173) - 1 = -0.159$

immunologists is not statistically distinct from 0, indicating that average productivity immunology stars in the aggregate have little influence on the performance of their coauthors. Specification 3 is identical to Specification 2, but instead of interpreting the coefficients on the influence of high productivity on coauthor performance relative to when the immunologist was alive, the omitted category in Specification 3 is the death of an average productivity immunologist. Interestingly, the null hypothesis that the performance effects of the death of a high productivity star is different from the death of an average productivity star cannot be rejected. Specification 4 looks at the average effect of death on a coauthor's performance for all star types. That is, the average effect of the death of an immunologist in the sample decreases the performance of their coauthors by 4.4%. This value, however, is highly insignificant.

Specification 5 introduces the second dimension of the star taxonomy: helpfulness. The omitted category for Specification 5 is the death of a Non-Star, and so all coefficients should be interpreted as relative to the performance effects of the death of a Non-Star. Most strikingly, the death of an All-Star decreases the performance of his coauthors by 38.6% (-0.488), the death of a Lone Wolf decreases the performance of his coauthors by 23.6% (-0.269), and the death of a Maven decreases the performance of his coauthors by 38.7% (-0.489). All coefficients are significantly different from 0 at the 10% level, while the Maven coefficient is significant at the 1% level.

Table 7 continues to present the main results of Equation 1 but differs from the results presented in Table 6 in that I estimate the full sample (dyads where the focal immunologist dies and dyads where the focal immunologist does not die). The models run in Table 7 are identical to those run in Table 6 apart from the different sample. As in the previous table, the sample in Specification 1 only includes high productivity immunologists (All-Stars and Lone Wolves). The average effect of the death of a high productivity star is not only quantitatively smaller than in the sample with only dyads where a death occurs but also statistically insignificant. Specification 2 returns to the full sample and examines the effect of death on coauthor performance by whether or not the immunologist has high or average productivity. The death of a high productivity star results in a 31.7% (-0.381) decrease of his coauthor's performance and is highly significant. Interestingly, the death of an average productivity immunologist also has a negative impact on his coauthor's performance, decreasing the subsequent quality adjusted output by 17.1% (-0.188). This indicates that when thinking about stars on a productivity continuum, the death of a coauthor, irrespective of his productivity, negatively influences the performance of his coauthors, albeit with different magnitudes. Specification 3 changes the omitted variable to the death of an average productivity immunologist, allowing for the direct test of the null hypothesis that high productivity immunologists have an equal impact on the performance of their coauthors than average productivity immunologists. While the coefficient indicates that the performance of coauthors of high productivity immunologists who die is 17.6% lower than average productivity immunologists who die, this value is not statistically distinct from 0. Specification 4 makes no distinction between immunologist star types and shows that the average decrease in performance of coauthors of immunologists who have died is 23.2% (-0.264).

Specification 5 introduces the new taxonomy star types. In this specification, death has an average 11% negative effect on the performance of coauthors but is statistically insignificant. The change in a coauthor's performance after a death varies by the type of star that dies. Coauthors of All-Stars produce 40.0% (-0.510) less quality-adjusted output, and coauthors of Mavens produce 39.0% (-0.494) less quality-adjusted output. The coauthors of Lone Wolves produce 21.4% (-0.241) less quality-adjusted output, but this value is not statistically distinct from 0. All of these coefficients are relative to the decrease in performance of coauthors of a Non-Star who dies. The variables that appear in Specification 5 are the main variables of interest for this study, and their relationships will be shown in various functional forms throughout this paper. The Lone Wolf and Maven coefficients are statistically distinct from one another at the 5% level.

To compare these results with those of AGW, recall that their range in performance decreases due to a coauthor's death is between -5 and -10%. As can be seen from Tables 6 and 7, the estimates I present are at times outside of this bound. One of the reasons for this difference may lie in the change of setting. Where AGW look at a range of life science disciplines, this study only examines immunologists. Second, AGW construct a control group using a form of propensity score matching to find two "nearest neighbor" control stars for every star who dies. My study, instead, uses all immunologists who do not die in my sample of 415 immunologists to form a control group. If either the immunologists used in the control group or their coauthors have lower performance (for a myriad of factors) than the immunologists who die or their coauthors, then my estimates will appear larger in magnitude. As can be seen from Table 5, the coauthors of immunologists who die do have higher average publication, citation, and Impact Factor-weighted publication counts than the coauthors of immunologists who do not die. This, however, does not directly imply that there should be any difference in performance impact from coauthor death across these two groups. Regardless, the range of parameter estimates do appear similar enough to assuage a reader's concern that AGW and this study are examining different phenomenon.

Overall, the significance of the star types from the new taxonomy appear both significant and stable. It does appear that Mavens are different from Lone Wolves. Furthermore, high productivity immunologists – All-Stars and Lone Wolves – each have different effects on the performance of their coauthors when they die. The next section examines the robustness of these relationships.

6.2 Robustness Checks

Table 9 presents OLS estimates of Equation 2 from Section 5. Specifications 1 and 2 replicate the results from Specification 5 from Tables 6 and 7, where a coauthor's change in performance is a function of the type of star who died. The coefficients from this linear model are qualitatively similar to those presented in the poisson fixed effects models estimated by quasi maximum likelihood (QML), and as such, I feel confident using the OLS model to estimate the relationship despite the count nature of the dependent variable. Specification 2, which makes use of the entire sample, shows that the death of an All-Star decreases the performance of his coauthors by 24% (-0.274), while the death of a Maven decreases the performance of his coauthors by 21.6% (-0.243). Both of these estimates are significant at the 1% level. Again, as in previous specifications, the effect of the death of a Lone Wolf is statistically indistinguishable from the effect the death of a Non-Star has on his coauthors.

Of the 415 immunologists in the sample, 28 die. While my identification of the impact of their deaths comes from the variation in performance of their coauthors, concerns of outliers driving the results may still exist. In Specifications 3 and 4, instead of looking at the death of the sample of immunologists, I examine their "exit" decisions. An exit is defined as the year in which an immunologist stops publishing for at least a four-year period. According to academic immunologists, if an immunologist has failed to publish a single manuscript in four years, then it is fairly reasonable to assume that this person has exited the risk set of publishing. An exit can occur for a number of reasons: retirement, decrease in productivity, move to industry, and of course death. Using the full sample in Specification 4, the results differ somewhat from the other tables. The exit of an All-Star is both economically and statistically insignificant, and the decrease in performance of coauthors of Lone Wolves and Mavens are not statistically distinct from one another. Three major explanations for these results exist. First, All-Stars may announce their exits well in advance, so that coauthors have ample time to adjust to the anticipated loss of the All-Star. Second, if an All-Star exits, he may still be active in an advising capacity, thus mitigating any losses that may have befallen his coauthors. Third, coauthors of exited All-Stars aren't very productive before or after the exit, and thus no decline is observed. These alternate reasons are unknown to the econometrician, and so I must search for some form of exogenous variation.

To remedy the endogeneity of the exit decision, Specifications 5 and 6 report results from an instrumental variable (IV) two stage least squares (2SLS) estimation where the death measure acts as an instrument for the exit decision. While results presented indicate a relationship between the death of an immunologist and the productivity of his coauthors, the relationship is driven entirely through the exit of the immunologist from the coauthor's coauthorship network. As such, the death of an immunologist should not affect his coauthor's productivity in any way other than through his exit, making death an appropriate candidate for an instrument of exit. I present linear probability models of the first stage estimates whereby the likelihood of an exit is a function of death in Table 8. As can be seen from the first stage estimates in Table 8, the death of an immunologist has a highly significant impact on the likelihood of the immunologist exiting the sample, almost by definition. Having an appropriate instrument, however, increases the level of variation and attenuates the likelihood that outliers are driving the results. The IV estimates presented in Specification 6 from Table 9 reveal that the exit of an All-Star decreases the performance of his coauthors by 42.1% (-0.546), while the exit of a Maven decreases the performance of his coauthors by 35.9% (-0.445).

Both of these coefficients are statistically significant at the 5% level, while the effect of a Lone Wolf's exit on his coauthor's performance is statistically insignificant.

Table 10 replicates Table 9 but changes the dependent variable from the count of Impact Factor-weighted publications to the count of citation weighted publications. As in Table 9, I provide OLS estimates for both the effect of a death and exit on a coauthor's performance in addition to instrumenting exit with death. With the change of the dependent variable, the results are somewhat different. The death of a Maven has a larger effect on the decrease in performance of his coauthors than the death of an All-Star, although these coefficients are not statistically different from one another. More interesting, however, is that for the first time the exit of a Lone Wolf has a larger negative effect on the performance of his coauthors than the exit of an All-Star, even after instrumenting for exit with death. The effect of a Maven's exit still has the largest negative effect on his coauthors, but the coefficient is not statistically distinct from the Lone Wolf coefficient. While all three major star types – All-Stars, Lone Wolves, and Mavens – negatively affect the performance of their coauthors with respect to citation-weighted publications, no one star type has a larger effect that is statistically significant. One of the main reasons for this finding may be the idiosyncratic way in which citations are awarded and the effect an individual has on generating citations. While a lone individual can greatly influence the quality of a manuscript and thus the quality of the journal it is accepted by, the process by which an article collects citations is less understood.

Table 11 presents results from estimating my main regression specification of interest by means of a fixed effects negative binomial. While the fixed effects poisson estimator with robust standard errors is much preferred over the negative binomial estimator (as the poisson estimator presents more conservative [larger] standard errors), the negative binomial estimator is still quite heavily used in strategy, sociology, and economics in the application of count models and as such is included to ensure that the results still appear reasonable. Specification 2 shows results from the full sample and finds results consistent with other estimators, whereby the death of an All-Star or Maven decreases the performance of his coauthors in excess of 20%.

Table 12 moves away from the discrete definition of productivity and helpfulness and instead looks at the continuous relationship between the effect of an exit or death interacted with continuous measures of productivity and helpfulness. I convert the measures of productivity and helpfulness to logs so that the coefficients may be interpreted as elasticities, which allows for a unit-free interpretation of the parameters. Specification 1 shows the effect of an immunologist's exit on the performance of his coauthors. A doubling in the productivity of the immunologist, where productivity is measured by citations received, results in a 5.2% decrease in the performance of his coauthors. Helpfulness has no statistical effect on the performance of coauthors when an immunologist exits. Specification 2 looks at deaths instead of exits. A doubling in the helpfulness of an immunologist who dies reduces the performance of his coauthor by 13.4%, a much higher level than in the case of exits, again raising the need for an IV estimation to address the endogeneity of exits. Specification 3 instruments exits with death and finds that a doubling of the helpfulness of an immunologist is associated with an 18.7% decrease in performance of his coauthors, while changes in productivity have no effect on the coauthor's performance. I change the dependent variable to citation-weighted publication counts in Specifications 4 through 6. Immunologist productivity has a statistically significant effect on the performance of coauthors when an immunologist exits, yet looking at deaths and exits instrumented by death, the effect of helpfulness strengthens but retains only marginal statistical significance.

7 Alternate Explanations

At least three alternate explanations may account for the relationship observed between the death of various star types and the subsequent decrease in performance of their coauthors. First, the decrease in a coauthor's performance may come from a decrease in funding that had been provided by the now deceased immunologist. Second, the status of the immunologist is artificially increasing the perceived performance of his coauthors and thus, after his death, the coauthors return to their natural steady-state. This casts doubt on the claim that the decrease in performance of a coauthor when an immunologist dies is due to the elimination of performance enhancing benefits such as providing helpful critique. Third, the effects observed are due to the influence of institution- and university-specific factors that influence the performance of coauthors.

The first alternate explanation of the reported results comes from the concern that an immunologist's funding largely drives coauthor performance. The concern is that because funding is almost always linked to a primary investigator, funding would stop with an immunologist's death, and so the relationship we witness whereby a coauthor's performance decreases with the death of an immunologist is simply being driven by the omitted funding variable. For this to be a viable alternate explanation, however, I must assume that a coauthor is able to benefit or make use of the immunologist's funding yet not include the immunologist as a coauthor. The probability of such an occurrence is very low as any use of funds will surely be tied to a coauthoring arrangement. Furthermore, recall that the dependent variable used throughout is the quality adjusted count of papers written by the coauthor without the focal immunologist. If the coauthor is benefiting from the immunologist's funding and consequently frequently coauthoring with the immunologist, then the coauthor's pre-death publications with the immunologist will not be counted. The empirical exercise of this paper is to compare pre- and post-death publishing rates of coauthors. If the pre-death rates are lower due to frequent coauthorship with the focal immunologist (and thus netted out), then the observed change in productivity after the immunologist dies should be negligible, thus biasing my results in an opposite direction from what is observed.

The second alternate explanation for the reported results is the conflation of performance with status. The concern is that a coauthor experiences positive performance due to an association with a high status immunologist (Merton 1973). For this to be true, status cannot act as an information signal, wherein due to information asymmetry association with an immunologist conveys quality onto the coauthor, consequently increasing his performance. If in this context status acts as a quality signal, then the signal should not be weakened once the immunologist dies, and consequently a decrease in a coauthor's performance after the death of an immunologist is unlikely to be associated with status effects. Furthermore, if status is driving the results reported, then I would not necessarily predict the strong effect of the death of a Maven. Moreover, if status is driving my results, then one would expect the effect of a Lone Wolf's death to have even larger negative effects on their coauthor's productivity, as All-Stars and Lone Wolves are considered high status individuals.

Third, heterogeneity in resources available to institutions or universities may increase the performance of immunologists. While this is certainly true, the only way for institutional effects to influence the performance of coauthors is if at the time of an immunologist's death the coauthor changes institutions. I employ dyad fixed effects across all specifications, and so any time-invariant characteristics that do not alter over the course of the panel, such as institutional setting, will be captured by the fixed effects. Lastly, it is very unlikely that a coauthor changes institutions in response to a former colleague's death, and thus institutional effects should have no influence on the results observed.

8 Discussion and Conslusion

This paper develops a new taxonomy of star scientists for the purpose of identifying scientists most likely to impact the performance of others. I expand the current conceptualization of star scientists, which presently only examines an individual's productivity, by adding a social behavior dimension: helpfulness. Helpfulness is the extent to which an individual is beneficial to others. By dividing immunologists into four classifications along the dimensions of productivity and helpfulness, I examine the magnitude of the decrease in performance of coauthors of immunologists who have died. I define All-Stars as scientists with both high productivity and helpfulness, where high is defined as being in the top 5% of the distribution. Lone Wolves are scientists with high productivity but average helpfulness. Mavens are scientists with average productivity but high helpfulness. Non-Stars make up the fourth classification and are average in productivity and helpfulness. Mavens and All-Stars have the largest negative impact on the publishing rates of their coauthors when they die, indicating the loss of a source of performance. The death of an All-Star on average decreases the performance of his coauthors by 35%, while the death of a Maven decreases the performance of his coauthors by 30%. These findings are robust to a series of controls and specifications.

These findings have several important implications for both theory and practice. First, the resource-picking mechanism in the resource-based view of the firm requires somewhat imperfect information to outsmart the resource market. This is difficult to do with public and highly verifiable resources such as a high productivity star. High helpfulness stars, on the other hand, are more difficult to measure, resulting in possible information asymmetries and a possible source of sustainable competitive advantage. Second, the learning by hiring literature (Song, Almeida, and Wu 2003) has mostly focused on the acquisition of knowledge through the hiring of individuals. Yet, as has been argued throughout this paper, if factor markets are efficient, then a recruited engineer, for example, should be able to capture his full output, which includes his embodied knowledge, through his salary. A logical extension to this literature is to think about the extent to which different star types affect a firm's ability to learn through the hiring process. Third, alongside AGW, this paper is one of the first studies to both measure and find conclusive support of the impact of star scientists on the performance of others. I find large economic and statistically significant results by forming a new taxonomy that focuses on scientists most likely to impact the performance of others.

Strategic hiring decisions are no longer the sole domain of human resources but rather require top corporate-level decision making as human capital has a large impact on firm performance. For many knowledge-based companies, the largest expense is the cost of human capital in the form of salaries. Developing an effective organizational design to properly manage the main inputs of a firm's innovation production function is of great importance.

From a public policy standpoint, a more nuanced understanding of human capital will surely inform the debate on regional clusters. One of the main benefits of clusters is both the generation of and the ability to absorb non-rival knowledge spillovers. Policy makers go to great lengths in developing appropriate incentives to optimize cluster structure with the goal of maximizing welfare. Understanding what types of human capital are most likely to impact the performance of others is of critical concern.

While the development of this new taxonomy provides a tool with which to examine human capital along a second dimension, its construction raises additional questions. How are Mavens priced relative to Lone Wolves? Does the market appropriately price the observed positive performance effects of different star types? The mechanism presented in this study is that of a strong tie formation. How then do All-Stars and Mavens diffuse knowledge and spillovers across weak ties (Granovetter 1973)? These questions are left to future research.

A number of limitations with this study still exist. First is the endogenous nature of acknowledgements. No clear or externally enforced rule exists over the administration of acknowledgements. Consequently, acknowledgements may be strategically applied for personal gains, such as winning favor with a journal referee or editor. However, if acknowledgements are bestowed upon individuals who are not helpful but instead in a position of authority, then we would expect additional noise to enter the helpfulness measure and consequently bias results towards 0 and in the opposite direction from the results presented here.

Second, the main conduit in this study by which stars impact the performance of others comes from the establishment of a social tie through the formation of a coauthoring relationship. Clearly, this is not the only forum by which stars impact the performance of others nor do I purport this to be so. This conduit should be viewed within the context of a larger research agenda that explores the ways in which different star types impact the performance of others.

Lastly, the external validity of this setting may be limited. While the objective functions of firms and academic departments are not entirely orthogonal, they are certainly different. While a firm setting would be ideal, the difficulty in obtaining both individual level productivity and helpfulness measures is greatly encumbering. Nonetheless three factors contribute to the attractiveness of studying immunologists. First, because acknowledgement patterns vary across disciplines, it is important to isolate this heterogeneity by focusing on a single distinct discipline. Second, because of this heterogeneity in acknowledgement norms found across academia, it is important to look at a discipline where acknowledgements are both present and applicable to my measure of helpfulness. And third, measures on helpfulness are incredibly hard to identify, let alone access for a cross section of firms, and as such are not the appropriate focus of a study such as this. A survey instrument, however, may be able to capture individual levels of helpfulness in a firm setting. This avenue is left open for future research.

This study presents preliminary evidence of the performance gains associated with coauthoring with helpful scientists. In doing so, it makes three important contributions. First, it extends the current dichotomous conceptualization of star scientists by explicitly defining star classification not only along the dimension of productivity but also the spectrum of helpfulness, thus developing a new taxonomy of star scientists. Second, it provides a measure by which helpfulness can be empirically tested: acknowledgements. Third, it attempts to establish a causal link between coauthoring with All-Stars and Mavens and an increase in performance.

The traditional method of bundling together All-Stars and Lone Wolves is quite problematic. All-Stars and Lone Wolves are quite different in their impact on the performance of others. Furthermore, Mavens, who under the current dichotomous conceptualization of star scientists are classified as Non-Stars, are actually quite important in affecting the performance of others. As such, our current classification of star scientists has possibly been systematically overvaluing Lone Wolves while undervaluing Mavens.

References

Acemoglu, D. (1996): "A Microfoundation for Social Increasing Returns in Human Capital Accumulation," *Quarterly Journal of Economics*, 111(3), 779–804.

Acemoglu, D., and J. Linn (2004): "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *Quarterly Journal of Economics*, 119(3), 1049–1090.

Agrawal, A., D. Kapur, and J. McHale (2008): "How do spatial and social proximity influence knowledge flows? Evidence from patent data," *Journal of Urban Economics*, 64(2), 258–269.

Ai, C., and E. C. Norton (2003): "Interaction Terms in Logit and Probit Models," *Economics Letters*, 80, 123–129.

Allison, P. D., and J. S. Long (1990): "Departmental Effects on Scientific Productivity," *American Sociological Review*, 55(4), 469–478.

Almeida, P., and B. Kogut (1999): "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science*, 45(7), 905–917.

Arrow, K. J. (1962): "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, ed. by R. R. Nelson, National Bureau of Economic Research, Conference Series. Princeton University Press, Princeton.

Audretsch, D. B., and M. P. Feldman (1996): "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86(3), 630–640.

Azoulay, P., J. Graff Zivin, and J. Wang (2008): "Superstar Extinction," Working Paper.

Barney, J. (1986): "Strategic Factor Markets," *Management Science*, 32, 1231–1241.

——— (1991): "Firm Resources and Sustained Competitive Advantage," *Journal of Management*, 17(1), 99–120.

Bennett, B. (1965): "Specific Suppression of Tumor Growth By Isolated Peritioneal Macrophages from Immunized Mice," *Journal of Immunology*, 95(4), 656–664.

Brown, J. (2008): "Quitter Never Win: The (Adverse) Incentive Effects of Competing with Superstars," Working Paper.

Ernst, H., C. Leptien, and J. Vitt (2000): "Inventors are not alike: The distribution of patenting output among industrial R&D personnel," *IEEE Transactions on Engineering Management*, 47(2), 184–199.

Fletcher, G. J. O., and M. S. Clark (2003): *Blackwell Handbook of Social Psychology: Interpersonal Processes*. Blackwell Publishing, Malden, MA.

Gilfillan, S. C. (1935): *Inventing the Ship*. Follett Publishing Company, Chicago.

Granovetter, M. S. (1973): "The Strength of Weak Ties," *The American Journal of Sociology*, 78(6), 1360–1380.

Groysberg, B., L.-E. Lee, and A. Nanda (2008): "Can They Take It With Them? The Portability of Star Knowledge Workers' Performance," *Management Science*, 54(7), 1213–1230.

Guthridge, M., A. B. Komm, and E. Lawson (2008): "Making Talent a Strategic Priority," *McKinsey Quarterly*, 1, 48–59.

Hackett, C. J., D. Rotrosen, H. Auchincloss, and A. S. Fauci (2007): "Immunology research: challenges and opportunities in a time of budgetary constraint," *Nature Immunology*, 8(2), 114–117.

Hausman, J., B. H. Hall, and Z. Griliches (1984): "Econometric Models for Count Data with an Application to the Patents-R & D Relationship," *Econometrica*, 52(4), 909–938.

Hirshleifer, J., A. Glazer, and D. Hirshleifer (1998): *Price Theory and Applications*. Prentice Hall, Upper Saddle River, NJ.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108(3), 577–598.

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

Kapur, D., and J. McHale (2005): *Give Us Your Best and Brightest*. Center for Global Development, Washington DC.

Lawrence, H. S., and M. Cohn (1993): "In Memoriam: Dr. Maurice Landy," *Cellular Immunology*, 152, 1–6.

Lazear, E. P. (1986): "Raids and Offer Matching," *Research in Labor Economics*, 8, 141–165.

Levin, S. G., and P. E. Stephan (1991): "Research Productivity Over the Life Cycle: Evidence for Academic Scientists," *The American Economic Review*, 81(1), 114–132.

Lotka, A. J. (1926): "The Frequency Distribution of Scientific Productivity," *Journal of the Washington Academy of Science*, 16, 317–325.

Lucas, R. E. (1988): "On the Mechanisms of Economic Development," *Journal of Monetary Economics*, 22(1), 3–42.

Makadok, R. (2001): "Toward a Synthesis of the Resource-Based and Dynamic-Capability Views of Rent Creation," *Strategic Management Journal*, 22, 387–401.

Merton, R. K. (1973): *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.

Narin, F., and A. Breitzman (1995): "Inventive Productivity," *Research Policy*, 24(4), 507–519.

Nelson, R. R., and S. G. Winter (1982): *An Evolutionary Theory of Economic Change*. Belknap Press of Harvard University Press, Cambridge, MA.

Oettl, A., and A. Agrawal (2008): "International labor mobility and knowledge flow externalities," *Journal of International Business Studies*, 39(8), 1242–1260.

Pakes, A., and Z. Griliches (1980): "Patents and R&D at the firm level: A first report," *Economics Letters*, 5(4), 377–381.

Romer, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), S71–S102.

Rosen, S. (1981): "The Economics of Superstars," *The American Economic Review*, 71(5), 845–858.

Silverstein, A. M., and B. Benacerrafe (2001): "In Memoriam: Philip Gell," *Cellular Immunology*, 213, 1–3.

Singh, J. (2005): "Collaborative Networks as Determinants of Knowledge Diffusion Patterns," *Management Science*, 51, 756–770.

Smith, C. A., D. W. Organ, and J. P. Near (1983): "Organizational Citizenship Behavior: Its Nature and Antecedents," *Journal of Applied Psychology*, 68(4), 653–663.

Song, J., P. Almeida, and G. Wu (2003): "Learning-by-Hiring: When Is Mobility More Likely to Facilitate Interfirm Knowledge Transfer?," *Management Science*, 49(4), 351–365.

Steinman, R. M., and C. L. Moberg (1994): "Zanvil Alexander Cohn 1926-1993," *Journal of Experimental Medicine*, 179(1), 1–30.

Tam, P.-W., and K. J. Delaney (2005): "Talent Search: Google's Growth Helps Ignite Silicon Valley Hiring Frenzy," *The Wall Street Journal*, 23 November 2005, A1.

Tedder, T. F., and J. R. Dawson (2003): "In Memoriam: D. Bernard Amos," *Journal of Immunology*, pp. 6316–6317.

Waldinger, F. (2008): "Peer Effects in Science – Evidence from the Dismissal of Scientists in Nazi Germany," Mimeo, London School of Economics.

Wernerfelt, B. (1984): "A Resource-Based View of the Firm," *Strategic Management Journal*, 5(2), 171–180.

Wooldridge, J. M. (1999): "Distribution-free estimation of some nonlinear panel data models," *Journal of Econometrics*, 90(1), 77–97.

——— (2002): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, MA.

Wuchty, S., B. F. Jones, and B. Uzzi (2007): "The Increasing Dominance of Teams in Production of Knowledge," *Science*, 316, 1036–1039.

Yunis, E. J. (2004): "D. Bernard Amos," in *Biographical Memoirs*, ed. by N. A. o. S. Office of the Home Secretary, vol. 85, pp. 1–19. The National Academies Press.

Zucker, L. G., M. R. Darby, and M. B. Brewer (1998): "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises," *The American Economic Review*, 88(1), 290–306.

Figure 1: Productivity and Helpfulness Distribution: N=415

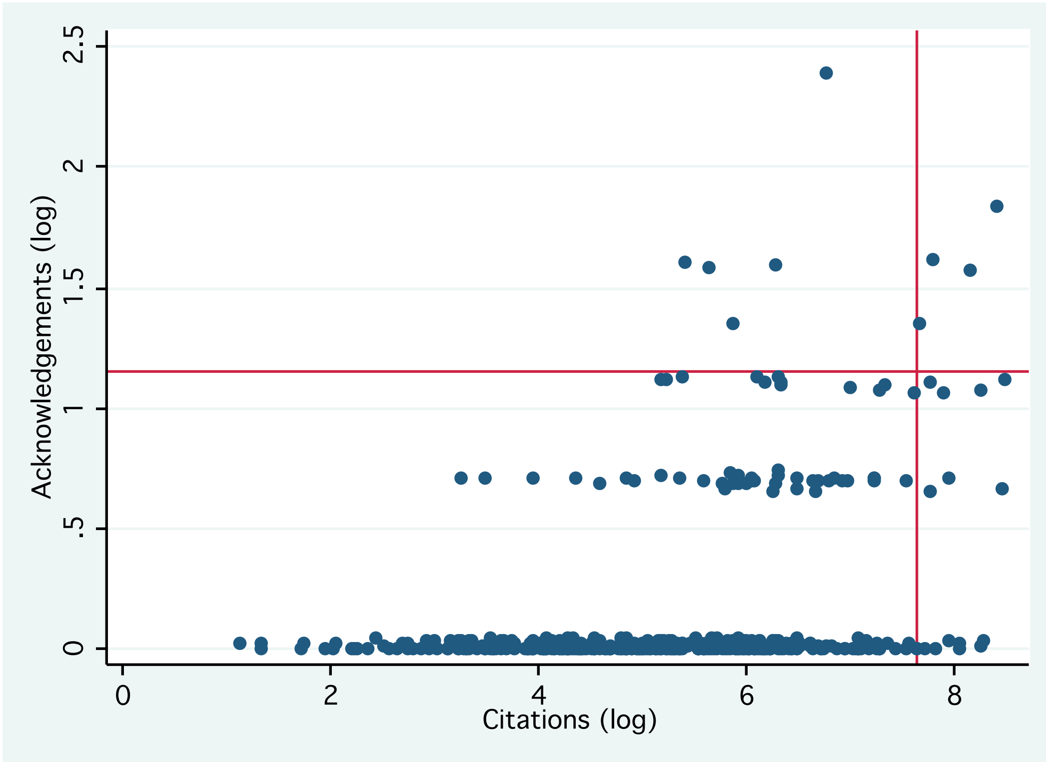

Table 1: A New Taxonomy for Star Scientists

	Average Productivity	High Productivity
High Helpfulness	Maven	All-Star
Average Helpfulness	Non-Star	Lone Wolf

Table 2: Star Scientist Classifications - Top 5%

	Average Productivity	High Productivity
High Helpfulness	**Maven** Total N = 5 Dyads = 189 Average coauthors = 37.8 Died N = 3 Died Dyads = 68 Average coauthors = 22.66	**All-Star** N = 4 Dyads = 188 Average coauthors = 47 Died = 2 Died Dyads = 22 Average coauthors = 11
Average Helpfulness	**Non-Star** Total N = 390 Dyads = 23,125 Average coauthors = 59.29 Died N = 18 Died Dyads = 490 Average coauthors = 27.22	**Lone Wolf** N = 16 Dyads = 673 Average coauthors = 42.06 Died = 5 Died Dyads = 236 Average coauthors = 47.2

Table 3: Variable Descriptive Statistics. Full Sample: N = 575,483

	Variable	Mean	Std. Dev.	1	2	3	4	5	6	7	8	9	10	11	12
1	Impact Factor Pubs	9.737	18.43												
2	Ln Impact Factor Pubs	1.376	1.34	0.74											
3	Citation-Weighted Pubs	22.239	56.10	0.20	0.33										
4	Death	0.012	0.11	0.02	0.04	−0.01									
5	All-Star Death	2.E-4	0.01	0.01	0.01	0.00	0.13								
6	Lone Wolf Death	0.003	0.06	0.02	0.03	0.00	0.55	0.00							
7	Maven Death	0.001	0.02	0.01	0.01	0.00	0.22	0.00	0.00						
8	Non-Star Death	0.007	0.08	0.01	0.02	−0.01	0.79	0.00	−0.01	0.00					
9	Exit	0.085	0.28	0.07	0.11	−0.03	0.23	0.05	0.11	0.08	0.18				
10	All-Star Exit	0.001	0.04	0.02	0.02	0.00	0.05	0.40	0.00	0.00	0.00	0.12			
11	Lone Wolf Exit	0.007	0.08	0.02	0.04	0.00	0.23	0.00	0.42	0.00	−0.01	0.28	0.00		
12	Maven Exit	0.002	0.05	0.01	0.02	0.00	0.10	0.00	0.00	0.47	0.00	0.16	0.00	0.00	
13	Non-Star Exit	0.074	0.26	0.06	0.10	−0.03	0.15	0.00	−0.02	−0.01	0.20	0.93	−0.01	−0.02	−0.01

Table 4: Variable Descriptive Statistics. Died Dyads Sample: N = 25,968

	Variable	Mean	Std. Dev.	1	2	3	4	5	6	7	8	9	10	11	12
1	Impact Factor Pubs	11.029	17.392												
2	Ln Impact Factor Pubs	1.569	1.348	0.78											
3	Citation-Weighted Pubs	25.142	55.955	0.23	0.33										
4	Death	0.255	0.436	0.07	0.11	−0.09									
5	All-Star Death	0.005	0.068	0.03	0.03	−0.01	0.12								
6	Lone Wolf Death	0.077	0.267	0.07	0.10	−0.02	0.49	−0.02							
7	Maven Death	0.012	0.110	0.02	0.04	−0.03	0.19	−0.01	−0.03						
8	Non-Star Death	0.161	0.367	0.02	0.03	−0.08	0.75	−0.03	−0.13	−0.05					
9	Exit	0.255	0.436	0.13	0.18	−0.06	0.57	0.12	0.24	0.19	0.43				
10	All-Star Exit	0.014	0.117	0.07	0.06	0.00	0.02	0.58	−0.03	−0.01	−0.05	0.20			
11	Lone Wolf Exit	0.056	0.229	0.07	0.11	−0.01	0.33	−0.02	0.70	−0.03	−0.11	0.41	−0.03		
12	Maven Exit	0.042	0.201	0.05	0.07	−0.02	0.02	−0.01	−0.06	0.53	−0.09	0.36	−0.02	−0.05	
13	Non-Star Exit	0.144	0.351	0.07	0.09	−0.06	0.48	−0.03	−0.12	−0.05	0.67	0.70	−0.05	−0.10	−0.09

Table 5: Coauthor Means (Std. Dev.) by Dyad Type

	Papers	Citations	Impact Factor Papers	Citation Weighted Papers
Full Sample				
All-Star	2.43	94.65	11.37	28.67
N = 4,659	(4.07)	(206.4)	(20.27)	(65.1)
Lone Wolf	2.39	96.04	11.31	30.27
N = 18,908	(4.09)	(209.7)	(18.92)	(65.53)
Maven	2.61	105.70	11.18	30.43
N = 5,285	(3.48)	(209.07)	(15.54)	(63.8)
Non-Star	2.33	71.19	9.65	21.83
N = 546,631	(4.2)	(172.93)	(18.42)	(55.55)
Died = 1				
All-Star	3.16	139.47	15.86	35.80
N = 792	(4.13)	(281.01)	(23.48)	(107.65)
Lone Wolf	2.83	103.65	12.24	28.82
N = 7,879	(3.94)	(215.39)	(18.61)	(62.2)
Maven	3.44	124.25	12.91	27.62
N = 2,551	(4.28)	(235.42)	(17.31)	(48.59)
Non-Star	2.44	74.58	9.80	22.18
N = 14,746	(3.48)	(163.75)	(16.19)	(48.92)
Died = 0				
All-Star	2.28	85.47	10.46	27.21
N = 3,867	(4.05)	(186.21)	(19.42)	(52.18)
Lone Wolf	2.08	90.60	10.65	31.32
N = 11,029	(4.17)	(205.38)	(19.11)	(67.79)
Maven	1.84	88.38	9.57	33.05
N = 2,734	(2.25)	(179.38)	(13.49)	(75.18)
Non-Star	2.33	71.09	9.65	21.82
N = 531,885	(4.22)	(173.17)	(18.48)	(55.72)

Table 6: Poisson QML Baseline Model - Dyads with a Death

Dependent Variable:	*Coauthor Impact Factor Weighted Publication Counts*				
	(1)	(2)	(3)	(4)	(5)
Death	-0.173^{+}		0.036	-0.044	0.098
	(0.092)		(0.098)	(0.076)	(0.107)
Death X High Productivity		-0.191^{*}	-0.227		
		(0.094)	(0.139)		
Death X Average Productivity		0.036			
		(0.098)			
All-Star Death					-0.488^{+}
					(0.257)
Lone Wolf Death					-0.269^{+}
					(0.143)
Maven Death					-0.489^{**}
					(0.130)
Dyad Fixed Effects	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓
Star Age Cohort FE	✓	✓	✓	✓	✓
Coauthor Age Cohort FE	✓	✓	✓	✓	✓
Observations	8671	25968	25968	25968	25968
Number of Dyads	258	816	816	816	816
Log Likelihood	-55362	-151631	-151631	-151830	-151225

QML robust star cluster adjusted standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 7: Poisson QML Baseline Model - Full Sample

Dependent Variable:	*Coauthor Impact Factor Weighted Publication Counts*				
	(1)	(2)	(3)	(4)	(5)
Death	-0.076		-0.188^{+}	-0.264^{**}	-0.115
	(0.112)		(0.110)	(0.084)	(0.117)
Death X High Productivity		-0.381^{**}	-0.194		
		(0.125)	(0.160)		
Death X Average Productivity		-0.188^{+}			
		(0.110)			
All-Star Death					-0.510^{+}
					(0.271)
Lone Wolf Death					-0.241
					(0.165)
Maven Death					-0.494^{**}
					(0.119)
Dyad Fixed Effects	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓
Star Age Cohort FE	✓	✓	✓	✓	✓
Coauthor Age Cohort FE	✓	✓	✓	✓	✓
Observations	23567	575483	575483	575483	575483
Number of Dyads	861	24175	24175	24175	24175
Log Likelihood	-147400	-3055558	-3055558	-3055758	-3055102

QML robust star cluster adjusted standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 8: OLS - Linear Probability Model - First Stage Estimates of Exit

Dependent Variable:	Exit	All-Star Exit	Lone Wolf Exit	Maven Exit	Non-Star Exit
	(1)	(2)	(3)	(4)	(5)
Death	0.429**				
	(0.082)				
All-Star Death		0.557**			
		(0.100)			
Lone Wolf Death			0.488**		
			(0.161)		
Maven Death				0.613**	
				(0.039)	
Non-Star Death					0.530**
					(0.075)
Constant	0.388**	−0.007	0.023	0.002	0.361**
	(0.077)	(0.006)	(0.018)	(0.005)	(0.075)
Dyad Fixed Effects	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓
Star Age Cohort FE	✓	✓	✓	✓	✓
Coauthor Age Cohort FE	✓	✓	✓	✓	✓
Observations	575483	575483	575483	575483	575483
Number of Dyads	24175	24175	24175	24175	24175
Log Likelihood	180066	1245618	822523	1114320	217141
R^2	0.39	0.07	0.15	0.13	0.39

Star cluster adjusted standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 9: OLS and 2SLS Model with Deaths and Exits I

Dependent Variable:	*Log of Coauthor Impact Factor Weighted Publication Counts*					
Dyads Included	Death	All	Death	All	Death	All
Estimation	OLS	OLS	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Death	-0.035	-0.091				
	(0.045)	(0.061)				
All-Star Death	-0.223^{*}	-0.274^{**}				
	(0.103)	(0.097)				
Lone Wolf Death	-0.082	-0.090				
	(0.067)	(0.069)				
Maven Death	-0.192^{*}	-0.243^{**}				
	(0.088)	(0.077)				
Exit			-0.006	-0.148^{**}	-0.240^{+}	-0.194^{+}
			(0.066)	(0.024)	(0.124)	(0.108)
All-Star Exit			0.007	-0.005	-0.341	-0.546^{*}
			(0.077)	(0.089)	(0.285)	(0.260)
Lone Wolf Exit			-0.132	-0.140^{**}	-0.174	-0.256
			(0.087)	(0.043)	(0.130)	(0.158)
Maven Exit			-0.171	-0.189^{+}	-0.434^{**}	-0.445^{**}
			(0.115)	(0.103)	(0.145)	(0.083)
Constant	1.078	0.744^{**}	1.003	0.804^{**}		
	(0.660)	(0.144)	(0.664)	(0.144)		
Dyad Fixed Effects	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Star Age Cohort FE	✓	✓	✓	✓	✓	✓
Coauthor Age Cohort FE	✓	✓	✓	✓	✓	✓
Observations	25968	575483	25968	575483	25968	575483
Number of Dyads	816	24175	816	24175	816	24175
Log Likelihood	-29755	-636510	-29753	-636045	-29928	-636336
Adjusted R^2	0.52	0.53	0.52	0.53	0.50	0.51

Star cluster adjusted standard errors in parentheses.
For Specifications 5 and 6, death instruments for Exit.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 10: OLS Model with Deaths and Exits II

Dependent Variable:	*Log of Coauthor Citation Weighted Publication Counts*					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV	(6) IV
Death	0.018	−0.002				
	(0.042)	(0.032)				
All-Star Death	−0.092	-0.096^{*}				
	(0.109)	(0.045)				
Lone Wolf Death	-0.103^{+}	−0.091				
	(0.060)	(0.068)				
Maven Death	−0.186	-0.179^{+}				
	(0.111)	(0.093)				
Exit			−0.058	-0.159^{**}	−0.065	−0.011
			(0.059)	(0.023)	(0.132)	(0.062)
All-Star Exit			−0.039	−0.019	−0.107	-0.176^{**}
			(0.081)	(0.054)	(0.168)	(0.050)
Lone Wolf Exit			-0.228^{**}	-0.083^{*}	-0.177^{*}	-0.186^{**}
			(0.069)	(0.036)	(0.076)	(0.071)
Maven Exit			-0.097^{+}	-0.086^{*}	-0.350^{+}	-0.292^{*}
			(0.049)	(0.037)	(0.207)	(0.133)
Constant	0.246	0.434^{*}	0.255	0.504^{**}		
	(0.616)	(0.177)	(0.657)	(0.176)		
Dyad Fixed Effects	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Star Age Cohort FE	✓	✓	✓	✓	✓	✓
Coauthor Age Cohort FE	✓	✓	✓	✓	✓	✓
Observations	25968	575483	25968	575483	25968	575483
Number of Dyads	816	24175	816	24175	816	24175
Log Likelihood	-32805	-718489	-32776	-718109	-32799	-718363
Adjusted R^2	0.66	0.67	0.66	0.67	0.65	0.65

Star cluster adjusted standard errors in parentheses.
For Specifications 5 and 6, death instruments for Exit.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 11: Robustness: Negative Binomial Fixed Effects

Dependent Variable:	*Coauthor Impact Factor Weighted Publication Counts*	
	(1)	(2)
Death	0.023	-0.052^{**}
	(0.024)	(0.019)
All-Star Death	-0.237^{*}	-0.260^{**}
	(0.095)	(0.091)
Lone Wolf Death	-0.090^{*}	-0.070^{*}
	(0.036)	(0.031)
Maven Death	-0.144^{*}	-0.224^{**}
	(0.058)	(0.057)
Constant	-1.712^{**}	-1.854^{**}
	(0.333)	(0.075)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Star Age Cohort FE	✓	✓
Coauthor Age Cohort FE	✓	✓
Observations	25968	575483
Number of Dyads	816	24175
Log Likelihood	-73415	-1477864

Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 12: Robustness: Continuous Measures of Productivity and Helpfulness

Dependent Variable:	*Log of Coauthor Impact Factor Weighted Publication Counts*			*Log of Coauthor Citation Weighted Publication Counts*		
	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) OLS	(6) IV
Exit	0.144^{*}		-0.143	0.221^{**}		0.363
	(0.062)		(0.263)	(0.077)		(0.230)
Exit X Helpfulness	-0.058		-0.187^{*}	0.017		-0.097
	(0.045)		(0.094)	(0.030)		(0.067)
Exit X Productivity	-0.052^{**}		-0.017	-0.069^{**}		-0.062^{*}
	(0.011)		(0.047)	(0.012)		(0.030)
Death		-0.074			0.198^{+}	
		(0.141)			(0.111)	
Death X Helpfulness		-0.134^{**}			-0.075^{+}	
		(0.046)			(0.045)	
Death X Productivity		-0.004			-0.033^{+}	
		(0.020)			(0.017)	
Constant	0.806^{**}	0.741^{**}		0.511^{**}	0.433^{*}	
	(0.146)	(0.144)		(0.176)	(0.177)	
Dyad Fixed Effects	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Star Age Cohort FE	✓	✓	✓	✓	✓	✓
Coauthor Age Cohort FE	✓	✓	✓	✓	✓	✓
Observations	575483	575483	575483	575483	575483	575483
Number of Dyads	24175	24175	24175	24175	24175	24175
Log Likelihood	-635901	-636513	-636222	-717926	-718487	-718294
Adjusted R^2	0.53	0.53	0.51	0.67	0.67	0.65

Star cluster adjusted standard errors in parentheses.
For Specifications 5 and 6, death instruments for Exit.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$