# Why Stars Matter*

Ajay Agrawal
University of Toronto and NBER
ajay.agrawal@rotman.utoronto.ca

John McHale
National University of Ireland, Galway
john.mchale@nuigalway.ie

Alexander Oettl
Georgia Institute of Technology
alex.oettl@scheller.gatech.edu

October 7, 2013

**Abstract**

We use a rich longitudinal dataset on department-level productivity in a contemporary field of science to identify and decompose the causal impact of hiring a star on local knowledge production. Specifically, we estimate the relative roles of knowledge spillovers versus recruiting externalities as they affect co-located researchers who are related or unrelated to the star in idea space. Hiring a star does not increase overall incumbent productivity, but this aggregate effect hides offsetting effects on colleagues who are related (positive) versus unrelated (negative). Star hires improve subsequent joiner quality for both related and unrelated scientists, although the effect is significantly larger for related scientists. The overall positive impact of the star on department-level productivity is mainly due to joiner-quality effects. Furthermore, the productivity impact is more pronounced at mid- and lower-ranked institutions, suggesting implications for the optimal spatial organization of science and university strategies aimed at ascending departmental rankings.

**JEL Classifications:** O31, J24, I23.

*We have greatly benefited from thoughtful comments and suggestions from Aneil Agrawal, Francine Lafontaine, and seminar participants at Emory, Passau, and NBER. We gratefully acknowledge funding support from the Centre for Innovation and Entrepreneurship at the Rotman School of Management, the Martin Prosperity Institute, and the Social Sciences and Humanities Research Council of Canada. Errors remain our own. 

*Certainly in our own profession, the benefits of colleagues from whom we hope to learn are tangible enough to lead us to spend a considerable fraction of our time fighting over who they shall be, and another fraction travelling to talk with those we wish we could have as colleagues but cannot. We know this kind of external effect is common to all the arts and sciences - the "creative professions." All of intellectual history is the history of such effects.*

Robert Lucas (1988)

# 1 Introduction

An influential strand of modern growth theory emphasizes the importance of combining existing ideas to produce new knowledge (Romer, 1990; Jones, 1995; Weitzman, 1998). As Mokyr (2002, p. 7) notes: "[w]hat makes knowledge a cultural entity . . . is that it is distributed to, shared with, and acquired from others; if that acquisition becomes too difficult, . . . knowledge will not be accessible to those who do not have it but are seeking to apply it." The challenges of accessing knowledge and cooperating to produce new knowledge highlight the importance of the spatial organization of science. However, in a modern market economy with free movement, the ultimate location of scientific activity is largely unplanned, resulting from individual utility-maximizing decisions, raising questions about the efficiency of the spatial allocation of scientists.

The efficient allocation is also likely to have changed over time. One reason is that the extent and nature of scientific collaboration is itself evolving. Benjamin Jones (2009) develops a "knowledge burden" theory that the depth and breadth of knowledge required to work at the outward shifting research frontier is increasing, raising the returns to collaboration. Agrawal et al. (2013) report data that support the knowledge burden hypothesis using the collapse of the Soviet Union as a natural experiment; the sudden shock to the knowledge frontier caused by the release of previously hidden research was more significant in some fields than others. They show that fields that undergo a greater outward shift in the knowledge frontier subsequently experience a disproportionate increase in collaboration. Furthermore, the rising knowledge burden may increase the importance of co-location since proximity facilitates collaboration.

Pulling in the opposite direction, however, is evidence that evolving communications technologies reduce the distance-related costs of collaboration (Agrawal and Goldfarb, 2008; Kim et al., 2009). These forces have the potential to alter the spatial organization of science, including the tendency (and desirability) of leading scientists to concentrate at top departments.

Another well-known feature of science is that the distribution of output is highly skewed across scientists. Almost a century ago, Lotka (1926) observed that 6% of physicists produced more than 50% of all papers. The relative importance of scientists in the right tail of the output distribution – *stars* – has endured (Rosen, 1981; Narin and Breitzman, 1995; Ernst et al., 2000). As our opening quote makes clear, however, the impact of stars on productivity goes well beyond their own publications. The presence of star scientists could directly affect their colleagues' productivity; their presence could also affect subsequent recruitment through a desire of others to be near them for productivity, reputational, or consumption reasons.

Thus, given the widely acknowledged importance of stars in science, it is surprising that limited evidence exists on the consequences of recruiting them. An exception is Waldinger (2012), who, utilizing the dismissal of scientists in Nazi Germany as a natural experiment, does not find evidence of productivity effects on peers. This is curious since the broader peer-effects literature documents significant productivity effects that are highly sensitive to the micro-geography of interactions (Sacerdote, 2001; Mas and Moretti, 2009).[1] Furthermore, Azoulay et al. (2010) and Oettl (2012) both report significant star-specific peer effects, utilizing data on unexpected star deaths as a natural experiment, although they both focus on coauthoring peers as opposed to co-located peers. Moreover, the broader research on spillovers emphasizes spatial relationships as a key determinant of knowledge flow patterns (Jaffe et al., 1993; Agrawal et al., 2006; Singh and Agrawal, 2011; Catalini, 2013), although the focus in these papers concerns the effect of co-location on the *direction* of research, as reflected in citation patterns as opposed

[1] Other studies focus on different benefits of stars, such as Zucker et al. (1998), who identify the location of star scientists as a key determinant of the timing and location of the birth of biotechnology firms.

to productivity.

In terms of recruiting externalities, the existing evidence is more uniform. Notwithstanding his earlier finding of an absence of peer productivity effects, Waldinger (2013) uncovers evidence of long-lasting effects on the quality of recruits of star dismissals in Nazi Germany. Roach and Sauermann (2010) report a strong preference of scientists to work with the best scientists possible.

Little is known, however, about the factors that influence the relative roles of these local knowledge spillovers and recruiting externalities. The distinction has important implications. For example, if recruiting externalities play a significant role, then a department with resources for further growth through additional hiring will enjoy higher returns from recruiting a star than an otherwise similar department that is not able to make additional hires and thus is unable to benefit from those externalities. However, these departments would experience similar returns from recruiting a star if the benefits are instead primarily due to knowledge spillovers.

Thus, we examine the question of *why* stars matter. To generate testable hypotheses, we first develop a model of how the hiring of a star affects incumbent productivity and the quality of subsequent recruitment. We assume Romer-style, knowledge-production functions, where incumbent productivity depends on local knowledge stocks. The impact of these stocks is allowed to differ depending on whether the knowledge is related or unrelated to the research of the incumbent scientist.

Hiring a star has direct positive impacts on incumbent productivity, and these effects are assumed to be larger for related incumbents. The proportional direct impact is also larger for lower-ranked institutions since the star's knowledge stock is a larger proportion of the total local stock. Critically, however, the star's impact on incumbent productivity is also conditioned on the impact of the star hire on subsequent recruitment. We introduce the idea of a recruitment function to capture this recruitment channel. For a given research area, this function shows how the quality of the applicant pool depends on existing local knowledge stocks, as well as on the speed with which the quality of the marginal hire declines with the number of hires in a particular research area.

We show that the average quality of subsequent joiners in both related and unrelated areas rises as a result of hiring a star. However, the star hire also shifts the optimal composition of hiring towards scientists working in areas related to the star. Overall, it is possible for the productivity of unrelated incumbents to decline relative to a no-star-hire baseline, notwithstanding a direct positive impact on their productivity.

The model suggests a number of testable hypotheses. A star hire will: 1) increase the productivity of related incumbents; 2) increase or decrease the productivity of unrelated incumbents, depending on the balance of the direct effect of the star's knowledge stock and the indirect effect through the composition of subsequent hiring; 3) increase the quality of both related and unrelated joiners; and 4) have larger proportional effects on incumbent productivity and joiner quality in lower-ranked institutions.

We use a rich longitudinal dataset on incumbent and joiner productivity in a contemporary field of science to identify the causal impact of hiring a star on department-level productivity. We examine the effect of hiring a star on incumbent productivity, distinguishing between incumbents who are related and unrelated to the star in "idea space." We also examine the effect of the star hire on the quality of subsequent recruits, again distinguishing between related and unrelated joiners. Finally, we examine how the incumbent and joiner effects are mediated by the rank of the hiring institution. Taken together, these results allow us to look inside the black box of how the location of stars affects scientific knowledge production and to better understand the forces driving the spatial organization of science.

We base our productivity estimates on a sample of 255 evolutionary biology departments that published 149,947 articles over the 29-year period 1980 to 2008. We employ a difference-in-differences estimation approach, comparing the productivity of "treated" to "control" departments before versus after the arrival of a star, to estimate the impact of a star hire on department productivity, where treatment refers to the recruitment of a star. Importantly, we distinguish between incumbent versus joiner scientists in the department and also between those whose work is related versus unrelated to the star.

We find evidence of a large overall star effect. On average, department-level output increases by 54% after the arrival of a star. A significant fraction of the star effect is indirect: after removing the direct contribution of the star, department level output still increases by 48%. In terms of department-level productivity, which we estimate by controlling for department size, we observe a 38% increase after excluding the star's contribution. This implies that much of the observed indirect output gains are due to increasing department quality, not just size. The effect does not seem to diminish even by the end of our sample period, eight years after the arrival of a star.

We next turn our attention to composition and distinguish between incumbent scientists who are in the department prior to the star and new recruits (or "joiners") who join the department after the arrival of the star. We further decompose the samples of incumbents and joiners into those who conduct research related to the star versus those who do not. We find that related incumbents increase their productivity after the arrival of the star by 69%, whereas the effect on unrelated incumbents is negative, perhaps due to resource shifting (negative point estimate, but statistically insignificant at standard levels). The overall star effect on incumbent productivity (related and unrelated combined) is neutral. Thus, we offer a first step towards reconciling the seemingly contradictory findings described above by reporting evidence that is on the one hand consistent with Waldinger (that is, no aggregate productivity effect on incumbents from hiring a star) and on the other hand consistent with others (that is, significant productivity effects on some) by disaggregating departments and distinguishing between co-located peers who are related versus unrelated to the star in terms of their position in idea space.

We then examine the impact of hiring a star on the quality of joiners. Since by definition joiners are not present in the department in the pre-star period, we shift our analytical approach to examining the quality of joiners (measured by the citation-weighted stock of their publications) who join the department in the years before versus after the arrival of the star. Overall, the quality of joiners jumps significantly (68%) after the arrival of a star. When we split the sample into related and unrelated joiners, the estimated increase in the quality of related joiners is a striking 434%. Interestingly, the quality of unrelated joiners also increases by 48%. Thus, although stars do not seem to generate production externalities (spillovers) for unrelated incumbents, they do appear to provide recruiting externalities for unrelated scientists that lead to attracting higher-quality joiners.

Reflecting on these results, we decompose the overall indirect star effect (38%) to determine the relative importance of production versus recruiting externalities. Overall, based on rough calculations that extrapolate from mean productivity changes in response to a star's arrival, we estimate that roughly 11% of this effect is due to a boost in related incumbent productivity, 0% is due to a boost in unrelated incumbent productivity, 42% is due to a quality increase in related joiners, and 47% is due to a quality increase in unrelated joiners.[2] The impact from unrelated joiners is high relative to related joiners, despite a significantly greater quality increase in related joiners, due to a larger average number of unrelated joiners.

We also examine the extent to which the star effect on department-level productivity is correlated with department rank. We assume that a star's share of their department's knowledge stock is greater at lower ranked institutions; thus, as per the model, we expect the direct pro-

[2]These calculations are crude. The mean output value prior to a star's arrival for treated departments is 42 citation-weighted publications. A 38% increase corresponds to an increase of 16 citation-weighted publications after the first star's arrival. We disaggregate this increase into the fractions from related versus unrelated peers. While the output of related scientists increases by 152% (exp(.924)-1), the mean is only 5.7 citation-weighted publications, corresponding to an increase in 8.6 citation-weighted publications, or 53% of the total 16 citation-weighted publications. Unrelated scientists experience a 21% increase in output from a baseline of 36, resulting in an increase in 7.6 citation-weighted publications, or 47% of the the total 16 citation-weighted publications. Since the output of unrelated incumbents never increased after the star's arrival, unrelated joiners account for 47% of the total increase. Related incumbents, however, experience a 70% increase in output after the star's arrival from a baseline of 2.6 citation-weighted publications prior to the star's arrival or 1.8 citation-weighted publications. Thus, if 1.8 citation-weighted publications (11% of 16) can be attributed to related incumbents, then the remaining 6.8 (42% of 16) citation-weighted publications of the total 8.6 related scientist increase can be attributed to joiners.

portional productivity effect of hiring a star to be larger at lower-ranked institutions. Indeed, we find that the star effect is significantly greater at lower-ranked institutions.

Finally, we explore the role of star engagement. Some stars engage with their new colleagues significantly more than others. Does engagement level influence the externalities stars generate for their department's productivity, or is their presence alone enough? We find that engagement through collaboration explains most of the increase in incumbent productivity but only a much smaller fraction of the increase in quality of new recruits.

Our analysis is subject to identification concerns. For example, it is possible that stars are attracted to moving to departments that are on the rise, rather than stars arriving at a department and causing the rise in productivity. In addition, it is possible that an omitted variable, such as a positive shock to department resources (e.g., philanthropic gifts, sharp increases in government funding, the construction of a new building), causes the department to both increase its chances of hiring a star and increase its overall productivity in terms of both incumbent productivity and the quality of subsequent recruits. Our difference-in-differences estimation method partially addresses these concerns by controlling for general productivity trends (treated versus control departments) and department-specific attributes (before versus after with department fixed effects).

To complement our initial empirical approach, we take three additional steps that, while not fully ruling out alternative explanations, give us further confidence that the relationship between the arrival of a star and department productivity is indeed causal. First, we employ a spline regression analysis for all results reported above (main star effect, star effect with star output netted out, star effect on incumbents, star effect on joiners, star effect on related versus unrelated for both incumbents and joiners). In all cases, we find: 1) the main effect persists over time (throughout the eight years examined after the arrival of the star), and 2) no evidence of a pre-trend in increasing productivity prior to the arrival of the star. These results help to rule out the alternate explanation (reverse causality) that stars in our sample move to departments because they are on the rise.

Second, we add controls for department- and university-level shocks that may influence both the hiring of a star and output by controlling for changes in the size, quality, and presence of a star in another subfield within biology (developmental biology, which is distinct from our focal subfield of evolutionary biology) as well as two additional unrelated departments at the focal university: mathematics and psychology. These results help to rule out the alternate explanation (omitted variable bias) that university- or even department-level shocks that may be correlated with both the recruiting of a star as well as the productivity of incumbents and quality of joiners are driving our result.

Third, we employ an instrumental variable analysis based on a count of the number of stars at other institutions who are at risk of moving to the focal institution in any given year, which is a function of the star's career age and work history (based on prior interactions with researchers from the focal university's region). This instrument is correlated with the probability of department $i$ hiring a star in year $t$ but is not correlated with department-level output. Our main results are robust to each of these extensions. While none of these individual tests are fully conclusive with respect to identification, together they provide further evidence that is consistent with our causal interpretation and inconsistent with alternative explanations.

The paper proceeds as follows. In Section 2, we set out the mechanisms through which incumbent productivity and recruitment externalities might operate in the context of a simple model with Romer-style local-knowledge production functions. We describe our data in Section 3 and our empirical strategy in Section 4. We report and interpret our basic difference-in-differences results in Section 5. In Section 6, we provide further evidence for a causal interpretation. Robustness checks are provided in Section 7 and we conclude with a discussion of the implications of our findings in Section 8.

# 2 A Simple Model

How does the hiring of a star scientist affect the performance of the hiring department? This section develops a simple model of the effects of hiring a star on both the productivity of incumbent scientists and the quality of subsequent hires.

## 2.1 Direct Productivity Effects on Incumbents

We begin with the direct effect of a star hire on the productivity of incumbents, ignoring initially any potential impacts through a changed composition of subsequent hires. We assume there are two types of scientists: type-1 and type-2. Type-1 scientists work on topic 1, and type-2 scientists work on topic 2. We further assume that the star is type-1. We measure individual scientist productivity by the flow of citation-weighted publications. For a given scientist of type-1, we model productivity by a Romer-style research production function:

$$\dot{A}_{1\iota} = \delta_{1i} A_1^{\theta_{11}} A_2^{\theta_{12}}, \tag{1}$$

where $\delta_{1i}$ is an individual productivity parameter for scientist $i$, $A_1$ is the total citation-weighted local knowledge stock of type-1 scientists, $A_2$ is the total citation-weighted local knowledge stock of type-2 scientists, and $\theta_{11}$ and $\theta_{12}$ are elasticities of individual productivity with respect to the local knowledge stocks of type-1 and type-2 scientists, respectively. We assume $\theta_{11} > \theta_{12}$, so that the knowledge spillover effect is greater within than across types. A similar productivity equation applies to type-2 scientists:

$$\dot{A}_{2\iota} = \delta_{2i} A_1^{\theta_{21}} A_2^{\theta_{22}}, \tag{2}$$

where $\theta_{22} > \theta_{21}$.

How does the hiring of a star type-1 scientist directly affect the productivity of the two scientist types? We assume that the knowledge stock of the star is $sA_1$, where $s$ is the star's knowledge stock as a share of the initial type-1 knowledge stock at the institution. Focusing first on type-1 scientists, the marginal productivity benefit of a one unit increase in the local knowledge stock of type-1 scientists is:

$$\frac{\partial \dot{A}_{1\iota}}{\delta A_1} = \theta_{11}\delta_{1i}A_1^{\theta_{11}-1}A_2^{\theta_{12}}. \tag{3}$$

We then represent the total impact on the productivity of type-1 scientists by the linear approximation:

$$d\dot{A}_{1\iota} \approx \frac{\partial \dot{A}_{1\iota}}{\delta A_1} dA_1 = \frac{\partial \dot{A}_{1\iota}}{\delta A_1} sA_1. \tag{4}$$

Using (1) and (3), we can write the proportional effect on type-1 productivity as:

$$\frac{d\dot{A}_{1i}}{\dot{A}_{1i}} \approx s\theta_{11}. \tag{5}$$

Similarly, we can write the proportional effect on type-2 scientists as:

$$\frac{d\dot{A}_{2i}}{\dot{A}_{2i}} \approx s\theta_{21}. \tag{6}$$

Thus, the direct productivity effect will be larger for type-1 scientists and also larger for institutions where the star represents a larger share of the initial type-1 knowledge stock (i.e., a large $s$). Assuming this share tends to rise with the rank of the institution, the direct proportional productivity effect of the hiring of a star will be larger at lower-ranked institutions.

## 2.2 Indirect Productivity Effects on Incumbents through Subsequent Hiring

In addition to these direct effects, the productivity of incumbents will also be affected by any impacts of the hiring of the star on subsequent recruitment. We therefore allow for the possibility of "recruitment externalities" in addition to the "knowledge spillover externalities" discussed above. We assume that the department has a fixed number of hiring slots, $H$ (not including the star). The hiring of a star may change the composition of the applicant pool for these slots and thus the composition of the subsequent hires. Letting $dA_{H1}$ be the change in the knowledge stock of type-1 scientists who are hired due the hiring of the type-1 star and $dA_{H2}$ be the change in the knowledge stock of type-2 scientists who are hired due to the type-1 star, the indirect effect on the productivity of type-1 scientists through the hiring channel is:

$$d\dot{A}_{1\iota} \approx \frac{\partial \dot{A}_{1\iota}}{\delta A_1} dA_{H1} + \frac{\partial \dot{A}_{1\iota}}{\delta A_2} dA_{H2}. \tag{7}$$

We can in turn rewrite this in terms of the proportional change in the productivity of type-1 scientists as:

$$\frac{d\dot{A}_{1i}}{\dot{A}_{1i}} \approx \alpha_1 \theta_{11} + \alpha_2 \theta_{12}, \tag{8}$$

where $\alpha_1 = \frac{dA_{H1}}{A_1}$ and $\alpha_2 = \frac{dA_{H2}}{A_2}$.

Similarly, the proportional indirect effect for type-2 scientists is:

$$\frac{d\dot{A}_{2i}}{\dot{A}_{2i}} \approx \alpha_1 \theta_{21} + \alpha_2 \theta_{22}. \tag{9}$$

We next consider how the hiring of the type-1 star affects the composition of hiring. We assume the institution hires the best scientists from the applicant pool for its open positions, where we measure quality by the citation-weighted knowledge stocks of the applicants. To solve for the optimal composition of hiring, we introduce the idea of a recruitment function. For type-1 scientists, the recruitment function gives the quality of the applicant in the $j$th position in the quality ranking, where we rank the applicants from best to worst. Letting $H_1$ represent the number of type-1 scientists hired, we give the quality of the marginal hire by:

$$A_{j1} = \phi_{11}(1+s)A_1 + \phi_{12}A_2 - \beta_1 H_1, \tag{10}$$

where the parameter $\beta_1$ measures how the quality of the marginal recruit falls with additional hires. In Figure 1, we graph from left to right the relationship between the quality of the marginal hire and the number of hires. Critically, the quality of the existing scientists (including the star scientist) is a shift factor for the recruitment function. An increase in the quality of incumbents will shift the recruitment curve upwards in Figure 1. Thus, the initial recruitment of the star scientist can support the hiring of better quality scientists for the additional available positions through a recruitment externality. Note that we allow for the possibility that potential recruits are attracted by the quality of existing scientists of the other type, though we assume $\phi_{11} > \phi_{12}$. A similar recruitment function applies for type-2 hires:

$$A_{j2} = \phi_{21}(1+s)A_1 + \phi_{22}A_2 - \beta_2 H_2, \tag{11}$$

where $\beta_2$ measures the rate of decline in the quality of the marginal type-2 recruit and $\phi_{22} > \phi_{21}$.

Assuming the institution seeks to maximize the total quality of recruits, the marginal quality of recruits will be equalized at the optimal composition of hires. We show this optimal composition in Figure 1. Imposing the condition $H_1 + H_2 = H$, we yield the optimal number of type-1 hires by:

$$H_1 = \left(\frac{\phi_{11} - \phi_{21}}{\beta_1 + \beta_2}\right)(1+s)A_1 + \left(\frac{\phi_{12} - \phi_{22}}{\beta_1 + \beta_2}\right)A_2 + \left(\frac{\beta_2}{\beta_1 + \beta_2}\right)H. \tag{12}$$

We next identify the *change* in the number of type-1 hires that results from the hiring of the star. From (12), we denote this change by:

$$dH_1 = \left(\frac{\phi_{11} - \phi_{21}}{\beta_1 + \beta_2}\right) sA_1. \tag{13}$$

The change in type-1 hires will be positive, provided that $\phi_{11} > \phi_{21}$. This will be the case if a given improvement in the quality of type-1 scientists has a greater positive impact on the recruitment of type-1 scientists than type-2 scientists. We assume this condition holds. Using similar reasoning, we denote the change in type-2 hires by:

$$dH_2 = -\left(\frac{\phi_{11} - \phi_{21}}{\beta_1 + \beta_2}\right) sA_1. \tag{14}$$

Thus, the hiring of the type-1 star will also shift the composition of subsequent hires towards type-1.

We next examine the impact of hiring the star on the average quality of hires. To determine the impact on average quality, we first note that the total quality of type-1 hires (measured by the total citation-weighted publications) is:

$$\begin{aligned} A_{H1} &= \int_0^{H_1} A_{j1} dj1 \\ &= \phi_{11}(1+s)A_1H_1 + \phi_{12}A_2H_1 - \frac{\beta_1}{2}H_1^2. \end{aligned} \tag{15}$$

We then represent the average quality of type-1 hires by:

$$\frac{A_{H1}}{H_1} = \phi_{11}(1+s)A_1 + \phi_{12}A_2 - \frac{B_1}{2}H_1. \tag{16}$$

The change in the average quality due to the hiring of the star is:

$$\begin{aligned} d\left(\frac{A_{H1}}{H_1}\right) &= \left(\phi_{11} - \frac{\beta_1}{2}\left(\frac{\phi_{11} - \phi_{21}}{\beta_1 + \beta_2}\right)\right) sA_1 \\ &= \left(\frac{(\phi_{11} + \phi_{21})\beta_1 + 2\phi_{11}\beta_2}{2(\beta_1 + \beta_2)}\right) sA_1 > 0. \end{aligned} \tag{17}$$

Therefore, the average quality of type-1 hires rises as a result of hiring the type-1 star. The average quality of type-2 hires also rises as a result of hiring the type-1 star. This is the result of both an upward shift in the recruitment function for type-2 scientists and a move along the curve due to the reduced hiring of these scientists.

$$\begin{aligned} d\left(\frac{A_{H2}}{H_2}\right) &= \left(\phi_{21} - \frac{\beta_2}{2}\left(\frac{\phi_{11}-\phi_{21}}{\beta_1+\beta_2}\right)\right) sA_1 \\ &= \left(\frac{(\phi_{11}+\phi_{21})\beta_2 + 2\phi_{21}\beta_1}{2(\beta_1+\beta_2)}\right) sA_1 > 0. \end{aligned} \tag{18}$$

What then is the overall indirect effect on the productivity of incumbents due to a changed composition of hiring compared to the case where no star is hired? For type-1 scientists, the effect on the local type-1 knowledge stock is positive. This positive effect comes through more type-1 scientists being hired and a higher average quality of those scientists. For type-1 scientists, the effect on the local type-2 knowledge stock is ambiguous: fewer type-2 scientists are hired but they are of higher average quality. However, given that the marginal productivity benefit for type-1 scientists of co-located type-1 scientists is greater than for type-2, these scientists will benefit from any shift in composition towards type-1. Overall, the indirect effect through hiring will be positive for type-1 scientists, which reinforces the direct productivity effect.

The overall indirect productivity effect is ambiguous for incumbent type-2 scientists. Productivity is enhanced as a result of more and higher-quality type-1 scientists being hired. However, they may lose from a possible negative effect on the total quality of type-2 hires, which will depend on the balance between fewer hires and a higher average quality of those hires. If the impact of fewer type-2 hires dominates, then it is possible that the overall indirect effect on the productivity of type-2 incumbents could be negative.

### 2.3 Summary of Testable Propositions

The model yields a number of testable propositions:

- A type-1 star hire will increase the productivity of type-1 incumbents. This is the result of a positive direct productivity effect from the star and a positive indirect effect through a star-related reputation effect on hiring.
- A type-1 star hire has an ambiguous effect on the productivity of type-2 incumbents. This is the result of a positive productivity direct effect and an ambiguous indirect productivity effect.
- Hiring a type-1 star will increase the average quality of type-1 and type-2 hires relative to the no-star-hire baseline.
- The productivity effects will be larger at lower-ranked institutions; that is, the productivity effects are increasing in $s$, the star's citation weighted knowledge stock expressed as a share of the initial type-1 knowledge stock.

## 3 Empirical Setting and Data

Our study focuses on the field of evolutionary biology, a sub-field of biology concerned with the processes that generate diversity of life on earth (e.g., the origin of species). Research in evolutionary biology consists of both theoretical and experimental contributions. While experimental evolutionary biology can be capital intensive due to the costs of running experiments in a lab, productivity within the discipline is not predicated on access to very specific facilities, as is the case in experimental particle physics and empirical astronomy. Evolutionary biology's mix of theoretical and experimental research activities makes it a good test subject for an initial exploration of the star effect on department growth.

## 3.1 Defining Evolutionary Biology

We use bibliometric data from the ISI Web of Science to calculate output at the department level and to identify the locations of evolutionary biologists. A critical first step is to define the field of evolutionary biology. We impute department membership using the following approach:

- We collect data on all articles published in the four main society journals of evolutionary biology: *Evolution*, *Systematic Biology*, *Molecular Biology and Evolution*, and *Journal of Evolutionary Biology*. We focus on these four society journals since every article published here concerns evolutionary biology and is relevant to evolutionary biologists. This yields 15,256 articles.
- We next collect all 149,947 articles that are referenced at least once by these 15,526 society journal articles. We call this set the corpus of influence since all of these referenced articles have had some impact on an evolutionary biology article. These 149,946 will serve as the basis of evolutionary biology knowledge for the purposes of our study.
- We then weight this corpus of influence by how many times each article has been cited by an article published in the set of 15,256 evolutionary biology society journal articles within five years of publication. There are 501,952 references from the 15,256 society journal articles to the 149,946 corpus of influence articles. We use the 501,952 references to construct our citation-weighted publication measure.

The key benefit of this approach, as opposed to simply using the ISI Journal Citation reports field definitions, is that it allows us to include general journals that evolutionary biologists are likely to publish in, such as *Science*, *Nature*, and *Cell* (among others).

## 3.2 Identifying Authors

We next attempt to attribute the 149,946 articles in the corpus of influence to individual authors. One problem with the ISI Web of Science data is that until recently it listed only the first initial, a middle initial (if present), and the last name for each author. Since our empirical objective is to trace the movement of evolutionary biologists across departments, it is first necessary to disambiguate authors (that is, to distinguish J Smith from JA Smith). We rely on heuristics developed by Tang and Walsh (2010) to disambiguate between authors who share the same name. The heuristic considers backward citations of two focal papers. If two papers reference similar papers (weighted by how many times the paper has been cited, i.e., how obscure or popular it is), then the likelihood of the papers belonging to the same author increases, and we link the two papers to the same author. We repeat this process for all papers with authors who have the same first initial and last name. We exclude scientists who do not have more than two publications linked to their name.

### 3.3 Identifying Scientist Locations

Using the generated unique author identifiers for each evolutionary biology paper, we next attribute each scientist to a particular institution for every year they are active. A scientist is active from the year they publish their first paper to the year they publish their last paper. Here again, we must overcome a data deficiency inherent within the ISI Web of Science data. Until recently, the Web of Science did not link institutions listed on an article to the authors. Instead, we impute author location using reprint information that provides a one-to-one mapping between the reprint author and the scientist's affiliation. In addition, we take advantage of the fact that almost 57% of evolutionary biology papers are produced with only a single institution listing. We thus are able to directly attribute the location of all authors on these papers to the focal institution.

We note that this method of location attribution is more effective within evolutionary biology than many other science disciplines since article production within evolutionary biology is not characterized by large teams (2.55 average authors per paper).

## 3.4 Unit of Analysis

Our unit of analysis is the department-year. We include all evolutionary biology departments that had at least one scientist present in 1980 and at least one scientist present in 2008. This criterion ensures that we are not simply counting new entrants or other idiosyncratic details of the data. Furthermore, this ensures that for any given department-year, a department is at risk of hiring a star scientist. Two-hundred-fifty-five departments fit this criterion. As such, we have 7,395 department-year observations.

## 3.5 Dependent Variables

We use three key dependent variables: 1) $Output_{it}$: the sum of the citation-weighted papers published by scientists present at department $i$ in year $t$; 2) $IncumbentOutput_{it}$: the sum of the citation-weighted papers published by scientists present the year prior to the star's arrival at department $i$ in year $t$; and 3) $JoinerQuality_{it}$: the mean citation-weighted stock of papers published up until year $t-1$ of all scientists who join department $i$ in year $t$.

We only use citations from articles published in the four evolutionary biology society journals that are made within five years of the focal paper's publication. In the majority of our specifications, we also exclude the publications of the arriving star.

## 3.6 Independent Variables

Our key independent variable is $Star_{it-1}$, which equals 1 if the year is greater than or equal to the year a star scientist (above the 90th percentile of citation-weighted stock of papers published up until year $t-1$) joins department $i$ and 0 otherwise. To ensure we observe adequate pre-treatment observations, we only examine the arrival of stars starting in 1985. Furthermore, we only examine the impact of the first arrival of a star. We provide a histogram of the variation in year of first star arrival in Figure2. As the figure illustrates, the timing of first star arrival varies significantly across institutions with approximately two thirds of the universities that recruit a star doing so during the first 10 years (1985-1995) and the remainder doing so in the second ten years (1995-2005).

### 3.7 Descriptive Statistics

We provide summary statistics of our dataset in Table 1. The average department in our sample produces just over 80 citation-weighted publications per year. When we exclude the contributions of the star, this number is reduced to just under 77 citation-weighted publications per year. While it initially may appear that the star is not contributing much to the department, we should note that this is the mean across all department-years and as such includes departments that never receive a star as well as the output of departments prior to the arrival of a star. Just under 22 scientists are active in each department in a given year on average, and incumbent scientists produce fewer than 18 citation-weighted publications a year.

## 4 Empirical Strategy

We examine the relationship between the arrival of a star scientist and the subsequent output of the department. The main empirical model we estimate is:

$$E[Y_{it}] = \exp(\alpha * Star_{it-1} + \beta * Scientists_{it} + \delta_t + \mu_i), \tag{19}$$

where $Y_{it}$ is one of our three dependent variables. As previously mentioned, we remove the arriving stars contributions to $Y_{it}$ in most specifications.

Of the 255 departments, 178 are treated (receive a star). The untreated departments act as control departments, allowing us to perform a difference-in-differences type estimation. The traditional post-treatment and treated cross-sectional unit coefficients are subsumed by the time dummies ($\delta_t$) and department fixed effects ($\mu_i$), respectively. Since the dependent variable is a count variable, we estimate our key specification using poisson quasi maximum-likelihood methods and adopt "Wooldridge" robust standard errors clustered at the department-level, which allows for arbitrary serial correlation Wooldridge (1999).

We also estimate our main specification with a full set of leading and lagging indicators of the star arrival variable in the following form:

$$\begin{aligned} E[Y_{it}] = \exp(&\alpha_{-10} Star_{it-10} + \alpha_{-9} Star_{it-9} + ... + \alpha_{-2} Star_{it-2} \\ &+ \alpha_0 Star_{it} + ... + \alpha_8 Star_{it+8} + \beta Scientists_{it} + \delta_t + \mu_i). \end{aligned} \tag{20}$$

The leading indicators help discern the extent to which reverse-causality influences our coefficients; that is, whether changes in department output influence the likelihood of recruiting a star. In addition, the lagged indicators allow us to explore temporal dynamics, in particular the duration of the star effect.

# 5 Difference-in-Differences Results

## 5.1 Department Output Increases after the Arrival of a Star

We begin by examining the relationship between the arrival of a star and the productivity of the department. The estimated coefficient on *Star* (Table 2, Column 1) implies that after a star arrives, department-level output increases by 53.7%, on average, per year (exp(0.430)-1 = 0.537). This is not surprising since the department now has a star who, by definition, is prolific. However, even after we remove the star's contribution, we still find a department-level increase in output of 48% per year on average (Column 2).

Recognizing that recruiting a star may coincide with an overall expansion of the department, we add a control for the number of scientists present in the department in the focal year. The estimated coefficient on *Star* indicates that a department's productivity (output per scientist) increases by 38%, on average, after the arrival of a star, still excluding the star's contribution to department output (Column 3). This estimate is both economically and statistically significant (1% level).

We next distinguish between incumbent scientists, who are in the department before the star arrives, and subsequent recruits ("joiners"). We begin by focusing on incumbents. Specifically, we drop joiners from the sample and estimate the prior equation based solely on incumbent data, controlling for the number of incumbents (as defined by their presence the year prior to the star's arrival) present in year $t$. The arrival of a star does not seem to have an economically or statistically significant relationship with incumbent output (Column 4). Since we define incumbents as scientists present the year prior to a star's arrival, we are only able to examine changes to incumbent output for departments that are "treated" by recruiting a star.

## 5.2 Star Effect on Joiner Quality

We turn next to examining joiners. We are not able to estimate joiner output the way we do for incumbents since by construction joiners have no output at the focal department prior to their arrival. Therefore, it is impossible to estimate a change in joiner productivity between the periods pre- and post-arrival of the star using our prior approach. However, we are able to observe variation in the *quality* of joiners before versus after the arrival of a star. To do this, we calculate the mean annual citation-weighted stock of papers published during the period prior to $t_{0-1}$ for each scientist joining department $i$ in year $t$. Significant variation exists in the quality of joining cohorts (mean = 37, standard dev. = 78, min. = 1, max = 2348, Table 1). Thus, we estimate the relationship between joiner quality (dependent variable) and the presence of a star (Table 3). As before, we use the department as the unit of analysis and employ both department and year fixed effects. The estimated coefficient on star indicates that after the arrival of a star, the mean quality of joining scientists increases by more than 70% (Column 1).

Next, we examine whether this boost in joiner quality applies across all levels of recruits (rookie, mid-career, senior). A number of studies document variation in productivity of scientists over their professional lifecycle (Lillard and Weiss, 1978; Levin and Stephan, 1991; Jones, 2010). Furthermore, Weinberg (2006) reports evidence that the extent to which a researcher is influenced by their co-located peers varies with age. To explore this issue in our setting in terms of how star impact on the quality of joiners varies with joiner vintage, we split the sample according to career age: 1) early-career (up to 10 years of publishing experience), 2) mid-career (10-20 years), and 3) late-career (more than 20 years). We report results in Columns 2, 3, and 4. The largest increase in quality appears to come from mid-career joiners, although the point estimates are not statistically distinguishable from those of early- and late-career.

### 5.3 Star Effect on Related Incumbents

We further dissect our main result by examining the difference between scientists who are working on topics related to the star versus those who are not. We classify a scientist as related if they cite at least one of the star's papers in any year prior to $t_{0-1}$ and unrelated otherwise. We split the sample accordingly. On average, 9% of incumbents and an equal fraction of joiners (9%) are related to the star. We find that the portion of the department that does research in areas related to that of the star experiences a significantly greater increase in output than the unrelated portion (Table 4, Column 1 versus 3). In fact, after the arrival of a star, the output of related scientists increases by more than 150% compared to 21% for unrelated.

In contrast to our earlier "no effect" result on incumbents, we find that incumbents who are related increase their productivity by 68% on average (Column 2). This result is hidden in the aggregate result reported earlier concerning incumbents since related incumbents represent a small fraction of overall incumbents (9%). Furthermore, the arrival of a star may adversely affect the level of resources allocated to unrelated incumbents, shifting resources from unrelated to related areas (e.g., future hires, department funds), which may result in a decrease in their productivity. The negative, albeit insignificant at conventional levels, point estimate may reflect that (Column 4). The negative effect on unrelated incumbents counteracts the positive effect on related incumbents such that, in the aggregate, the overall effect on incumbents is neutral, as reported above (Table 2, Column 4) and consistent with the aggregate findings reported in Waldinger (2012).

### 5.4 Star Effect on Related Joiners

We combine our analyses on joiner quality and relatedness in the analysis we report in Table 5. We classify joiners as related or unrelated following the procedure described above. We split the sample according to relatedness and, following the procedure described in Section 5.2 above, we estimate the relationship between joiner quality and the presence of a star. Although the quality of both types of joiners increases after the arrival of the star, the increase is much greater for joiners who work in related areas of research: 434% compared to 48% (Columns 1 and 2, respectively). Still, it is interesting to note that the quality of unrelated joiners increases after the arrival of a star, in contrast to the productivity of unrelated incumbents, which does not increase.

### 5.5 Department Rank

Next, we examine the extent to which the star effect on department-level productivity is influenced by the rank of the institution. In Table 6, we report the point estimates of $Star_{it-1}$ for regressions using our three main dependent variables (*Output w/o Star*, *Incumbent Output*, and *Joiner Quality*) split by institutions in the top 25 at the time of the star's arrival versus not-top 25. The rank splits reveal large heterogeneity in effects across institution types. Top departments experience less of a gain after star arrival of their first star compared to institutions outside of the top 25. These results are robust to different cutoffs for top institutions (e.g., top 10, top 50).

### 5.6 Collaboration

Next, we examine the extent to which star engagement with their new colleagues is associated with the observed department-level productivity gains. We employ co-production of new knowledge (i.e., coauthorship) as a proxy measure for star engagement. We report the results in Table 7. First, we focus on the sample that includes all scientists (Columns 1-3). The variable *Collaborations w/Star* is a count of the number of collaborations between the star and a colleague in the same department. An additional collaboration with the star is associated with a 3.1% increase in overall department-level productivity. The effect is approximately 50% greater when we focus only on related peers (4.5%). Star engagement is not correlated with the productivity of unrelated peers.

Although star collaboration accounts for some of the variation in department-level productivity (as compared to the point estimates in Table 2, Column 3 and Table 4, Column 1) it does not fully account for the increase in productivity after the star's arrival. While co-production between stars and their department peers is important, it does not fully explain the productivity increase post-star arrival. That said, collaboration is only one channel through which stars may engage with their peers. However, star collaboration does seem to account for all of the productivity boost for incumbents (Columns 4-6). As with the results we report in Columns 1 through 3, more star collaboration is associated with a greater increase in incumbent productivity, but in contrast to Columns 1 through 3, in Columns 4 through 6 the inclusion of the collaboration variable causes the main effect of the star's arrival to disappear. This stands in stark contrast to the large and statistically significant effect from the arrival of a star on related incumbent productivity that we report in Table 4, Column 2.

# 6 Is the Estimated Star Effect Causal?

The previous section has documented an economically and statistically significant star effect on department productivity (excluding the output of the star), related incumbent productivity, and post-star joiner quality. However, the suspicion remains that these effects might not reflect the causal impact of the star. Star recruitment might just be a manifestation of a broader strategy to improve department size and quality (omitted variable bias). Moreover, the successful recruitment of the star might itself be the result of independently improving department performance. We adopt a three-strand approach to further support a causal interpretation of the Section 5 results.

First, we use spline regressions and associated graphics to examine pre-trends in productivity and joiner quality. This allows us to examine whether the improvement in performance pre-dates the arrival of the star. The absence of a pre-trend would help rule out a broader department-improvement strategy or reverse causality from performance to star recruitment. Second, we add controls for department- and university-level shocks that might influence both the hiring of a star and output. We add controls for changes in the size, quality, and presence of a star in another subfield within biology (developmental biology, which is distinct from our focal subfield of evolutionary biology) as well as two additional unrelated departments at the focal university: mathematics and psychology. Third, we introduce an instrument for star recruitment based on a time-varying measure of move risk for stars in evolutionary biology who have a well-defined pre-existing connection with the focal department.

## 6.1 Spline Regressions and Pre-Trends

It is plausible that departments are better able to recruit a star because their output is increasing or that the recruitment of the star reflects a pre-existing department-improvement strategy. To explore this possibility, we estimate regressions that include a full set of leading and lagging indicators for the star variable in line with the specification outlined in Equation 20.

We present the results for department output in graphical form in Panel A of Figure 3. Department-level output remains reasonably constant in the years leading up to recruiting the star. Specifically, output in years $t_{-10}$ to $t_{-2}$ is statistically indistinguishable from output in the year prior to the star's arrival ($t_{-1}$), the omitted category. The bars correspond to 95% confidence intervals. Output increases sharply the year of the star's arrival relative to $t_{-1}$. Thus, we find no evidence of a pre-trend. In other words, stars do not appear to be moving in order to join departments "on the rise."

We then repeat this estimation on the sample that drops the contributions of the arriving star in Panel B of Figure 3. Thus, this sample includes only incumbents and joiners. Once again, we find no evidence of an uptick in department-level output in the years leading up to the arrival of the star. Furthermore, with this sample, we do not observe an increase in post-arrival output until two years after the star's arrival, likely since this is driven by new recruits who may be more likely to join due to the presence of the star. Moreover, the increase in output relative to $t_{-1}$ persists for the full period for which we have data (up to $t_8$).

Next, we repeat this estimate on the sample that drops the contributions of both the star and joiners, such that only incumbent data remains. This corresponds to the regression results we report in Table 2, Column 4. We find no evidence of a change in the output of incumbents either in the years leading up to the arrival of the star or in the years following the star's arrival (Figure 4).

Turning to the joiner quality findings, the spline regression reveals that joiner quality is not rising prior to the arrival of a star. We find no discernible evidence of a pre-trend (Figure 5). However, the average quality of joiners increases significantly and almost immediately after a star is hired. This suggests that much of the observed increase in department-level output is attributable to an increase in the quality of recruits following the star's arrival.

Figures 6 and 7 each show separate splines for related and unrelated scientists. Figure 6 shows the pre- and post-arrival productivity effects excluding the star but including joiners. Figure 7 shows these same effects for incumbents only. We observe no pre-trends.

## 6.2 Additional Department and University Controls

In Table 7, we control for department- and university-level shocks that may influence both the hiring of a star and department-level output. We do this by controlling for the presence of a star and the department size at the focal institution's developmental biology, mathematics, and psychology departments. We construct our developmental biology sample in a similar fashion to the one outlined in Section 3.1 by drawing upon all articles cited at least once in the following main developmental biology journals: *Development*, italDevelopmental Biology, *Developmental Cell*, and italGenes & Development. We construct our mathematics and psychology departments by drawing upon all articles published in journals classified as "Mathematics" or "Psychology" in the ISI Journal Citation Reports. Controlling for these effects only slightly diminishes the magnitude of the reported effects.

## 6.3 An Instrument for Star Recruitment

The splines and controls help to rule out pre-existing strategies to improve evolutionary biology department performance and strategies that coincide with the recruitment of the star that are also present in other biology disciplines and the wider university. However, the recruitment of the star might still be coincident with a new strategy of department improvement that is specific to evolutionary biology. This suggests the use of an instrument for star recruitment that is plausibly uncorrelated with any change in departmental strategy. Our instrument, $MoveRisk$, is a dummy set to 1 if the cumulative count of the number of star scientists (90th percentile) who are at risk of moving in year $t$ to department $i$ is above the median level across all years and institutions, and 0 otherwise.[3]

[3]The median number of stars at risk of moving is 14.

Our instrumental-variable strategy relies on within-department analysis; consequently, we require a time-varying instrument. The instrument consists of two components: 1) the timing of a move, and 2) the choice of destination. We model the timing of a move on the exogenous increase in career age. As can be seen from the plot overlays in Figure 8, a scientist has the highest probability of moving between the career ages (years since first publication) of 5 and 10. The red line (y) plots the results from Table 9, Column 1.

In Table 10 we present results from a location choice regression, where the decision of scientist $j$ to move to department $i$ is a function of a dummy if the scientist has any coauthors (prior to move) which are at the focal department and a dummy if the scientist has worked at a department in the same state as the focal department (if in the US) or if the scientist has worked at a department in the same country as the focal department (if not in the US). The results in Table 10 show that all three factors positively correlate with the probability of moving to the focal department. An f-test of excluded instruments indicates that we do not have a weak instrument problem.

The instrument $MoveRisk$ is the cumulative count of scientists who have coauthors at the focal department and have either been at a department in the same state as the focal department (if in the US) or in the same country as the focal department (if outside of the US) and who have a career age between 5 and 10. The instrument is cumulative since our endogenous variable is a dummy that stays "on" once treatment has occurred.

Table 11 presents the two-stage least squares (2SLS) estimates instrumenting the arrival of a 90th percentile star with $MoveRisk$. Column 1 presents the results of the first-stage regression regressing $StarArrival$ on the instrument, $MoveRisk$. The excluded instrument is both economically and statistically meaningful: when there are more than 14 scientists (above the median number of risk) at risk of moving to institution $i$ in year $t$, institution $i$ is 10% more likely to hire a star. As can be seen in the remainder of the specifications (Columns 2-5), the point estimates are qualitatively similar to those generated in earlier tables.[4] The point estimates are larger, but the differences are not statistically significant.

Table 12 presents instrumental variables (IV) results for our analysis of department-level output, incumbent output, and joiner quality by the scientist's topic-relatedness to the star. Once again, the IV results provide results qualitatively in line with the poisson results previously presented: the arrival of a star positively increases a department's output, incumbents' output, and joiner quality for scientists who work in related areas, does not increase output of aggregate department output, and decreases incumbent output of those working in unrelated areas but still increases joiner quality for scientists working in unrelated topic areas.

# 7 Additional Robustness Checks

We next conduct three additional robustness tests for our main results. We first examine the effect of star departures (rather than star arrivals). We report the results for our three main dependent variables in Table 13. Not surprisingly, star departures are associated with a decline in department output, even after excluding the output of the star. Perhaps more surprisingly, the negative effect on incumbent productivity of star departures is larger in magnitude than the positive effect of star arrival. A possible explanation is that departing stars have developed relationships with incumbents (e.g., collaborations, mentoring, or simply knowledge exchange) leading to adverse impacts on the productivity of those left behind. As Agrawal et al. (2006) emphasize, relationship capital built during periods of co-location endures, at least in part, post separation. Nonetheless, prior co-location is likely to be less effective in supporting incumbent productivity than current co-location. The final column in Table 13 shows a positive effect of star departure on the quality of subsequent hires, although the coefficient is not statistically significant at conventional levels. The positive coefficient may reflect the freeing up of resources

[4] We log transform the dependent variables (ln +1) to allow for easier comparison with the log-linear poisson model we present throughout.

as a result of the star departure. However, another possibility is suggested by the model in Section 2. In the model, the positive effect of a star arrival comes partly through the effect on subsequent hiring. These effects tend to be positively reinforcing, as successful recruitment supports further successful recruitment. The positive effect of star departure may reflect this dynamic to the extent that star departure is correlated with prior star arrival in our data.

Second, we examine the robustness of our results to an alternative method of identifying stars (Table 14). Rather than identifying stars based on their ranking in the distribution of citation-weighted cumulative output, we do so based on their membership in the National Academy of Sciences (NAS). This advantage of this method is that it is not directly related to any measures of output that we use as dependent variables in our regressions. One disadvantage is that it reduces the number of observed star arrivals in our data from 178 to 31 scientists. We find that the arrival of an NAS scientist has no statistically significant impact on the output of the institution or on subsequent joiner quality but is associated with a decrease in the output of incumbents. To check whether the arrival of a NAS scientist has a greater impact on lower-ranked departments, we next interact our arrival variable with department ranking dummies. We report estimated coefficients on indicator variables for the arrival of a NAS scientists at a Top 25 and not-Top 25 institution in Table 15. The results indicate that the arrival of NAS scientist at a Top 25 institution diminishes incumbent productivity and has no effect on subsequent joiner quality while the arrival of a NAS scientist at a lower ranked institution has no effect on incumbent productivity but increases the quality of subsequent joiners.

Third, we examine the effect of an incumbent scientist being elected to the National Academy of Sciences (Table 16 ). Comparing these results with our prior results allows us to distinguish between a physical move and a change in status. The change in status could have an effect on the quality of subsequent recruitment due to reputation effects. However, we do not observe statistically significant associations for any of our three dependent variables. Separating the relationship between an incumbent's election to the National Academy of Sciences by department tier once again reveals heterogeneity in outcomes. Table 17 shows that irrespective of department rank, the election of an incumbent scientist to the National Academy of Sciences is unrelated to changes in incumbent output and subsequent joiner quality. On the other hand, non-Top 25 departments with an incumbent that becomes a member of the National Academy of Science experience an increase in output net of the scientist's own output. These results suggest that the reputations of stars that are elected to the National Academy are already established prior to their election. We did not have a strong prior on how election to the National Academy would affect incumbents. On the one hand, the election could help the star access funding or improve publication prospects, with positive spillovers to incumbent colleagues. On the other hand, the election could create additional external demands on the time of the star, reducing their capacity to support the productivity of their departmental colleagues.

# 8 Discussion and Conclusion

We explore how the hiring of a star scientist affects incumbent productivity and the quality of subsequent recruitment. We find that the effects of star location are economically significant but subtle. To illuminate the causal channels, we apply a simple model that allows for both differentiated knowledge and recruitment spillovers. We base differentiation on the relatedness of work of the star to incumbents and potential joiners. The model's prediction that related incumbents should benefit from a star hire is strongly supported in the data, with the effect being strongest where there is evidence of actual collaboration by the star with incumbents. For unrelated incumbents, the model shows how a star hire can actually harm incumbent productivity through hiring composition effects, despite positive direct knowledge spillovers. Empirically, we find evidence of modest negative adverse impacts, which also explains the failure to find evidence of productivity effects for incumbents in the aggregate. The model's prediction that a star will improve the quality of both related and unrelated joiners also finds strong support in the data. Finally, we also uncover evidence to support the model's prediction of larger proportional productivity and recruitment effects in lower-ranked institutions.

The main empirical challenge is to demonstrate that the observed star-related associations are at least in part causal. We adopt a three-part approach to support a causal interpretation: an examination of pre-trends (to rule out a pre-existing department-improvement trend), controls for university- and department-level shocks (e.g., surge in resources), and use of an instrument that is correlated with star recruitment but plausibly uncorrelated with broader department improvement strategies. While none of these approaches provides perfectly clean identification on its own, together they give evidence that is consistent with a casual explanation of the observed star effects and inconsistent with the plausible alternative explanations.

What are possible normative implications of our findings on why stars matter? In general, our findings on the impact of stars on colleague productivity and the dynamics of recruitment suggest that the location decisions of stars are important for the efficient organization of science. The evidence that highly productive scientists are drawn to one another for reputational as well as productivity reasons raises a concern that there may be excessive positive sorting of scientists from an efficiency perspective at top-ranked institutions.

Such sorting might lead to missed opportunities for the development of strong clusters of related scientists to form around a star at less highly-ranked institutions. On the other hand, certain such institutions should have a strong incentive to pursue star-focused strategies to ascend the rankings. Our findings suggest that star-recruitment strategies may be most effective where a cadre of related incumbents is already present and the department has a flow of new hiring slots sufficient to take advantage of the improved quality of potential new recruits. Our findings thus have possible lessens for public/private funding and endowment strategies for seeding dynamic research clusters.

As noted in the introduction, the efficient spatial distribution of scientists is also likely to shift with increased incentives for collaboration on one side and reduced costs of distance-related communication on the other. Although a university department is a rather special local knowledge economy, our findings on the relative importance of knowledge- and recruitment-related externalities is also suggestive of a broader role of "stars" – scientists, CEOs, entrepreneurs, and the like – in the dynamics of local agglomeration and growth. We plan to further probe these potential social welfare implications in future research.

# References


Agrawal, A., I. Cockburn, and J. McHale (2006). Gone but not forgotten: Knowledge flows, labor mobility, and enduring social relationships. *Journal of Economics Geography 6*(5), 571–591.

Agrawal, A. and A. Goldfarb (2008, September). Restructuring research: Communication costs and the democratization of university innovation. *American Economic Review 98*(4), 1578–90.

Agrawal, A. K., A. Goldfarb, and F. Teodoridis (2013). Does knowledge accumulation increase the returns to collaboration? evidence from the collapse of the soviet union. University of Toronto mimeo.

Azoulay, P., J. Graff Zivin, and J. Wang (2010). Superstar extinction. *Quarterly Journal of Economics 125*(2), 549–589.

Catalini, C. (2013). Microgeography and the direction of inventive activity. Massachusetts Institute of Technology mimeo.

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.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993, August). 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? *The Review of Economic Studies 76*(1), 283–317.

Jones, B. F. (2010). Age and great invention. *The Review of Economics and Statistics 92*(1), 1–14.

Jones, C. I. (1995, August). R & d-based models of economic growth. *Journal of political Economy 103*(4), 759–784.

Kim, E. H., A. Morse, and L. Zingales (2009). Are elite universities losing their competitive edge? *Journal of Financial Economics 93*(3), 353–381.

Levin, S. and P. Stephan (1991, March). Research productivity over the life cycle: Evidence from american scientists. *American Economic Review 81*(1), 114–132.

Lillard, L. A. and Y. Weiss (1978). Experience, vintage, and time effects in the growth of earnings: American scientists 1960-1970. *Journal of Political Economy 86*, 427–448.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Science 16*, 317–325.

Lucas, R. E. (1988, July). On the mechanisms of economic development. *Journal of Monetary Economics 22*(1), 3–42.

Mas, A. and E. Moretti (2009, September). Peers at work. *American Economic Review 99*(1), 112–45.

Mokyr, J. (2002). *The gifts of Athena: Historical origins of the knowledge economy*. Princeton and Oxford: Princeton University Press.

Narin, F. and A. Breitzman (1995, July). Inventive productivity. *Research Policy 24*(4), 507–519.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science 58*(6), 1122–1140.

Roach, M. and H. Sauermann (2010). A taste for science? phd scientists' academic orientation and self-selection into research careers in industry. *Research Policy 39*(3), 422–434.

Romer, P. M. (1990, October). Endogenous technological change. *Journal of Political Economy 98*(5), S71–S102.

Rosen, S. (1981, December). The economics of superstars. *The American Economic Review 71*(5), 845–858.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of Economics 116*(2), 681–704.

Singh, J. and A. Agrawal (2011). Recruiting for ideas: How firms exploit the prior inventions of new hires. *Management Science 57*(1), 129–150.

Tang, L. and J. P. Walsh (2010). Bibliometric fingerprints: Name disambiguation based on approximate structure equivalence of cognitive maps. *Scientometrics 84*(3), 763–784.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in nazi germany. *The Review of Economic Studies 79*(2), 838–861.

Waldinger, F. (2013). Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge. University of Warwick mimeo.

Weinberg, B. A. (2006). Which labor economists invested in human capital? geography, vintage, and participation in scientific revolutions. Ohio State University mimeo.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics 113*(2), 331–360.

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

Zucker, L. G., M. R. Darby, and M. B. Brewer (1998, March). Intellectual human capital and the birth of u.s. biotechnology enterprises. *The American Economic Review 88*(1), 290–306.

Figure 1: Impact of a Type-1 Star Hire on Subsequent Recruitment

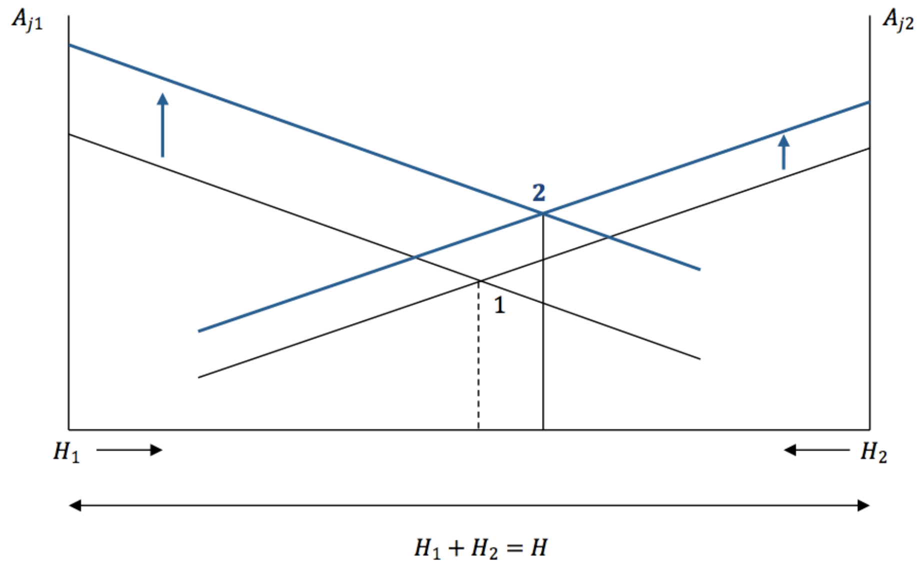
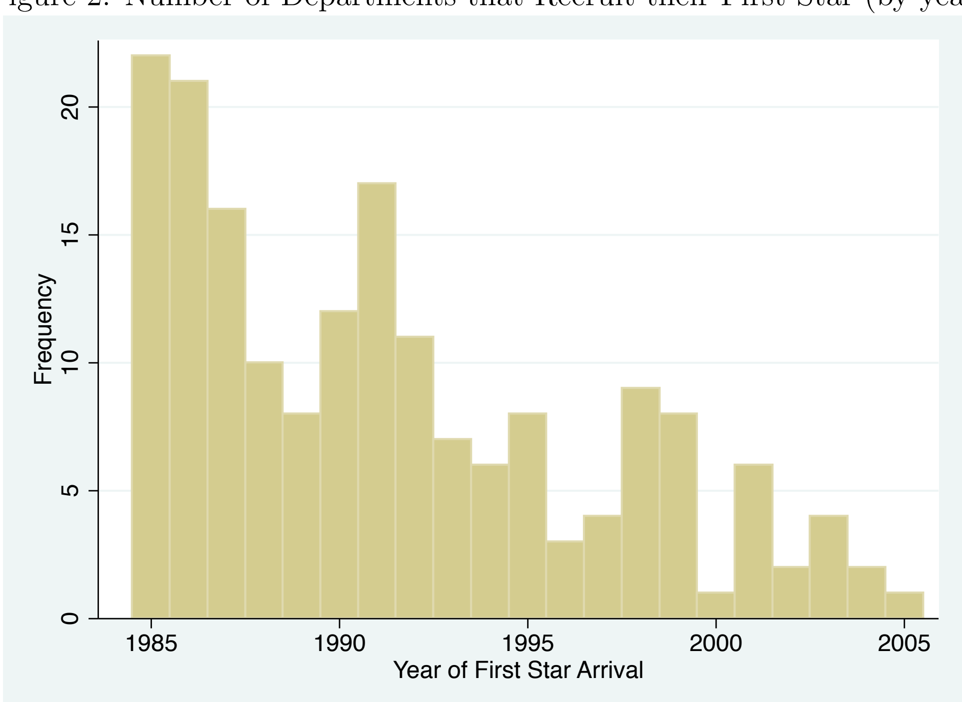
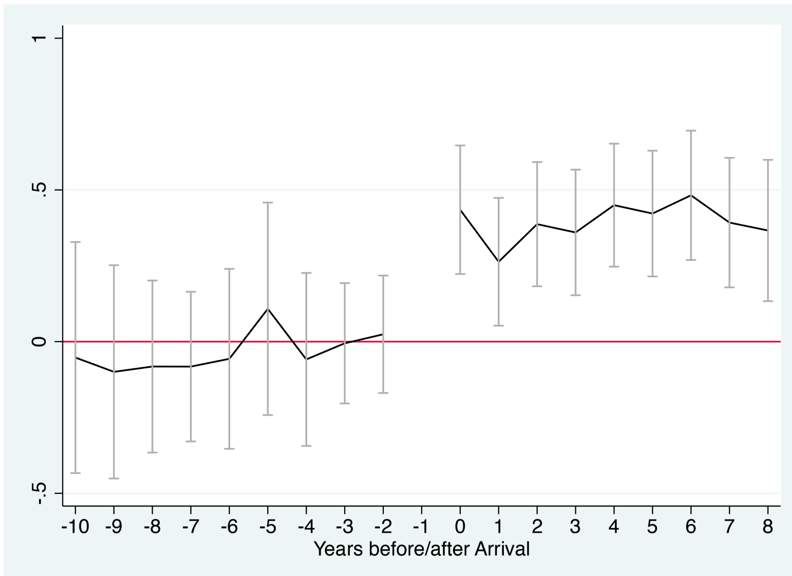



Figure 2: Number of Departments that Recruit their First Star (by year)

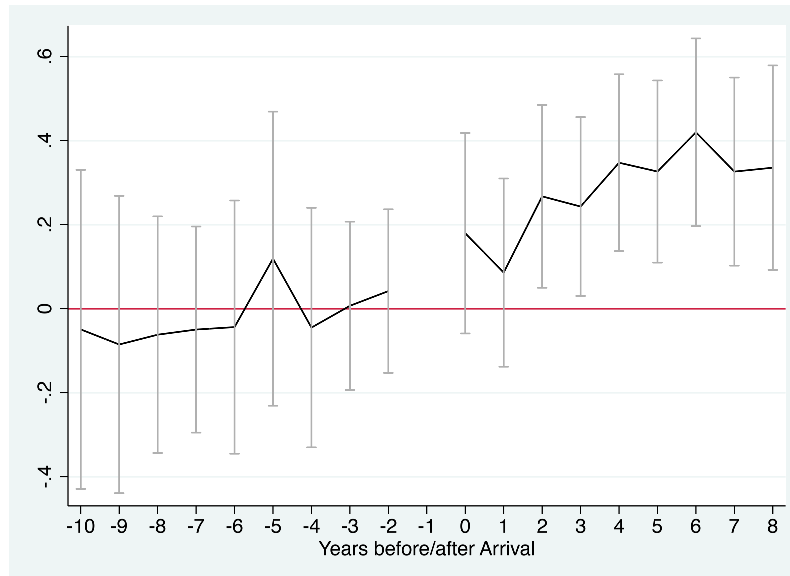



Notes: The above histogram displays the year in which departments recruit their first star.

Figure 3: Department Output



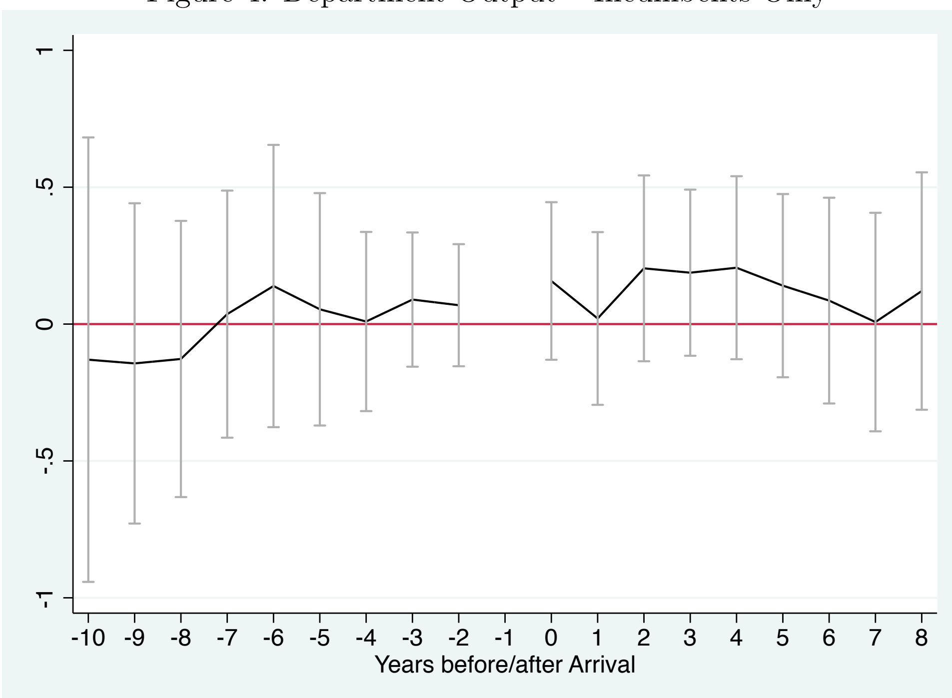

(a) Including Star




(b) Excluding Star

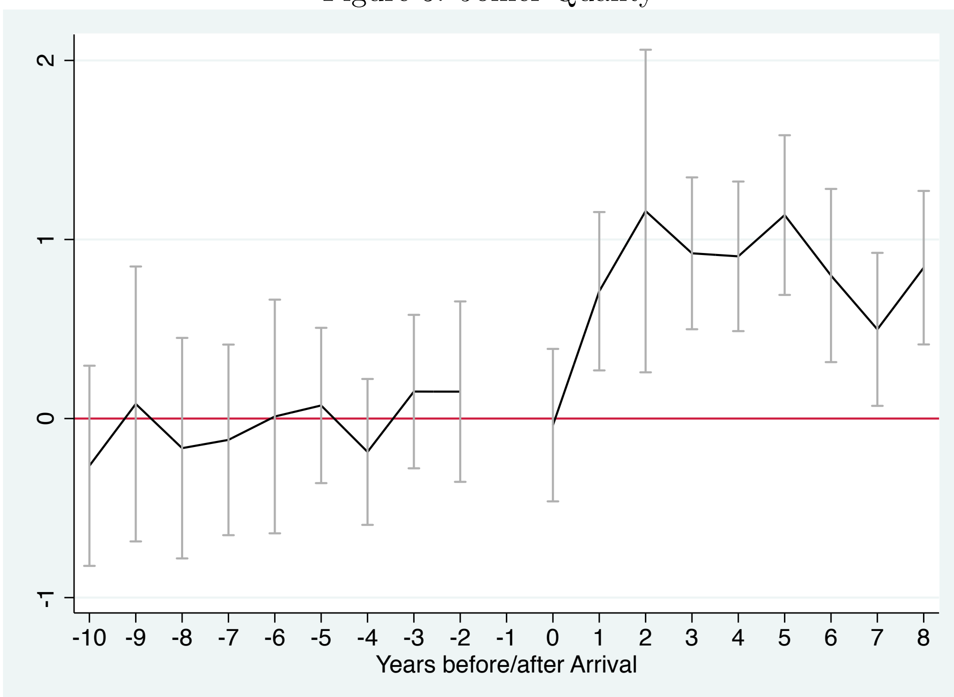
Notes: This figure plots point estimates for leading and lagging indicators for the arrival of a department's first star. Both panels plot the point estimates of the following specification: $E[Y_{it}] = exp(\alpha_{-10}\ Star_{it-10} + \alpha_{-9}\ Star_{it-9} + \ldots + \alpha_{-2} Star_{it-2} + \alpha_0 Star_{it} + \ldots + \alpha_8 Star_{it+8} + \beta Scientists_{it} + \delta_t + \mu_i)$. $E[Y_{it}]$ is the output of department $i$ in year $t$. $Star_{it-10}$ is set to 1 for years up to and including 10 years prior to the arrival of the star and 0 otherwise. $Star_{it+8}$ is set to 1 for all years eight years after the arrival of the star and 0 otherwise. The omitted category is one year prior to the star's arrival. The vertical bars correspond to 95% confidence intervals with department-clustered standard errors. Both panels include controls for the number of scientists present in year $t$. Panel A includes all scientists, while Panel B excludes the focal star.

Figure 4: Department Output – Incumbents Only



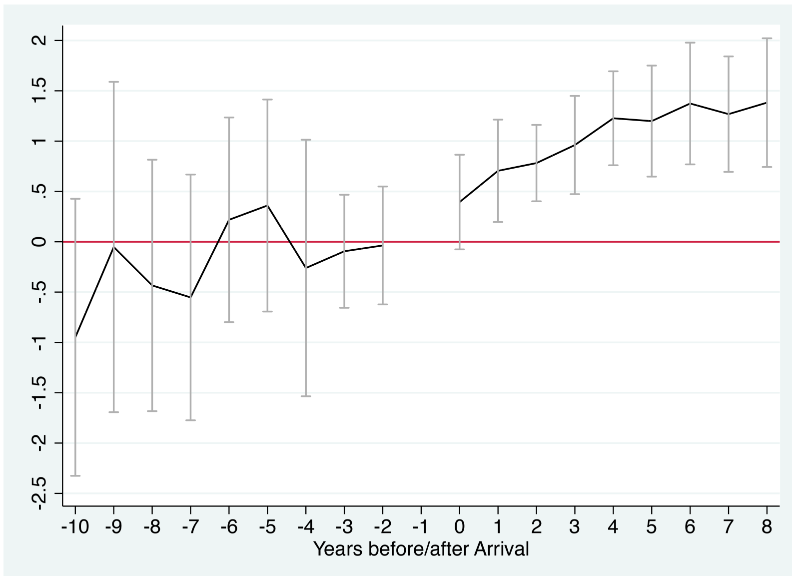

Notes: This figure plots point estimates for leading and lagging indicators for the arrival of a department's first star. The figure plots the point estimates of the following specification:
$E[Y_{it}] = exp(\alpha_{-10}\ Star_{it-10} + \alpha_{-9}\ Star_{it-9} + \ldots + \alpha_{-2} Star_{it-2} + \alpha_0 Star_{it} + \ldots + \alpha_8 Star_{it+8} + \beta Incumbents_{it} + \delta_t + \mu_i)$. $E[Y_{it}]$ is the incumbent output of department $i$ in year $t$. $Star_{it-10}$ is set to 1 for years up to and including 10 years prior to the arrival of the star and 0 otherwise. $Star_{it+8}$ is set to 1 for all years eight years after the arrival of the star and 0 otherwise. $Incumbents_{it}$ controls for the number of incumbents present in year $t$ at department $i$. We define incumbents as scientists who are present in department $i$ the year prior to the star's arrival. The omitted category is one year prior to the star's arrival. The vertical bars correspond to 95% confidence intervals with department-clustered standard errors.

Figure 5: Joiner Quality

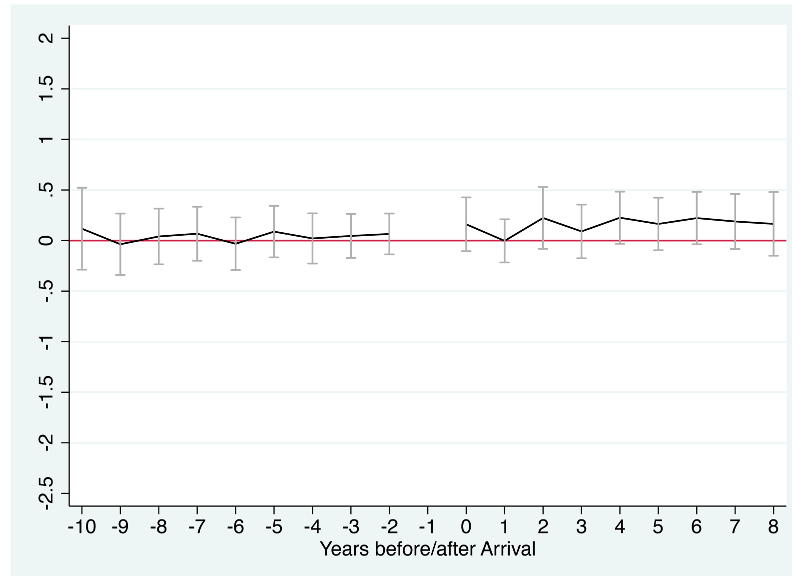



Notes: This figure plots point estimates for leading and lagging indicators for the arrival of a department's first star. The figure plots the point estimates of the following specification: $E[Y_{it}] = exp(\alpha_{-10}\ Star_{it-10} + \alpha_{-9}\ Star_{it-9} + \ldots + \alpha_{-2} Star_{it-2} + \alpha_0 Star_{it} + \ldots + \alpha_8 Star_{it+8} + \delta_t + \mu_i)$. $E[Y_{it}]$ is the mean quality of scientists who join department $i$ in year $t$. $Star_{it-10}$ is set to 1 for years up to and including 10 years prior to the arrival of the star and 0 otherwise. $Star_{it+8}$ is set to 1 for all years eight years after the arrival of the star and 0 otherwise. The omitted category is one year prior to the star's arrival. The vertical bars correspond to 95% confidence intervals with department-clustered standard errors.

Figure 6: Department Output Excluding Star: Related versus Unrelated



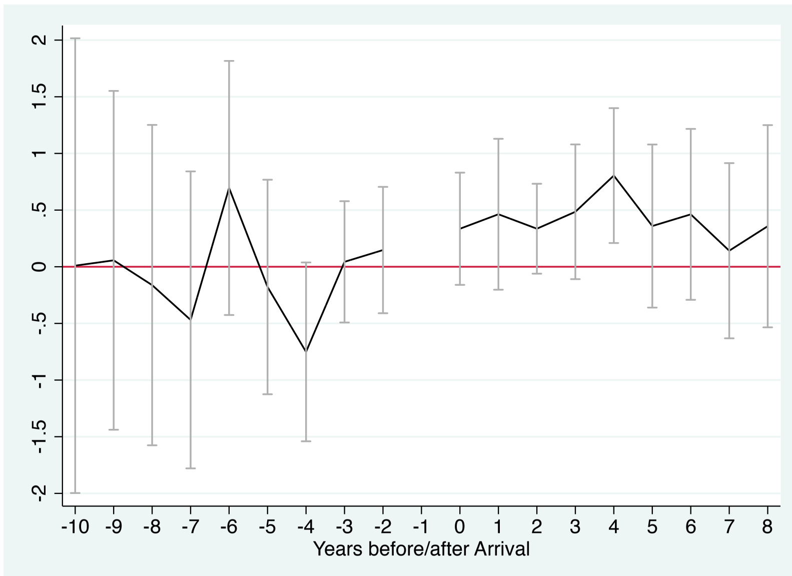

(a) Related Scientists



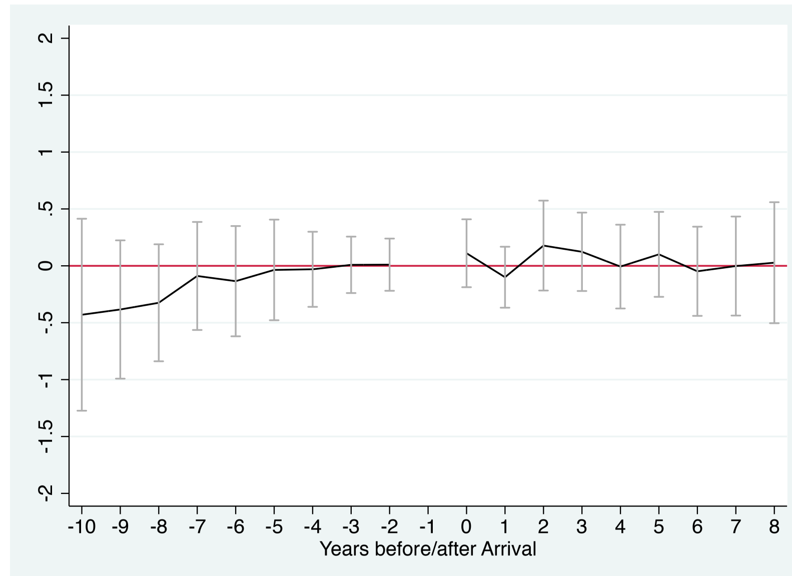

(b) Unrelated Scientists

Notes: This figure plots point estimates for leading and lagging indicators for the arrival of a department's first star. Both panels plot the point estimates of the following specification: $E[Y_{it}] = exp(\alpha_{-10}\ Star_{it-10} + \alpha_{-9}\ Star_{it-9} + \ldots + \alpha_{-2} Star_{it-2} + \alpha_0 Star_{it} + \ldots + \alpha_8 Star_{it+8} + \beta Scientists_{it} + \delta_t + \mu_i)$. In Panel A, $E[Y_{it}]$ is the output of department $i$ in year $t$ of related scientists. In Panel B, $E[Y_{it}]$ is the output of department $i$ in year $t$ of unrelated scientists. A scientist is related (to the star) if the scientist has cited at least one of the star's work in earlier years. $Star_{it-10}$ is set to 1 for years up to and including 10 years prior to the arrival of the star and 0 otherwise. $Star_{it+8}$ is set to 1 for all years eight years after the arrival of the star and 0 otherwise. The omitted category is one year prior to the star's arrival. The vertical bars correspond to 95% confidence intervals with department-clustered standard errors. Both panels include controls for the number of scientists present in year $t$.

Figure 7: Department Output – Incumbents Only: Related versus Unrelated




(a) Related Incumbents




(b) Unrelated Incumbents

Notes: This figure plots point estimates for leading and lagging indicators for the arrival of a department's first star. Both panels plot the point estimates of the following specification: $E[Y_{it}] = exp(\alpha_{-10}\ Star_{it-10} + \alpha_{-9}\ Star_{it-9} + \ldots + \alpha_{-2} Star_{it-2} + \alpha_0 Star_{it} + \ldots + \alpha_8 Star_{it+8} + \beta Incumbents_{it} + \delta_t + \mu_i)$. In Panel A, $E[Y_{it}]$ is the output of department $i$ in year $t$ of related incumbent scientists. In Panel B, $E[Y_{it}]$ is the output of department $i$ in year $t$ of unrelated incumbent scientists. A scientist is related (to the star) if the scientist has cited at least one of the star's papers in earlier years. $Incumbents_{it}$ controls for the number of incumbents present in year $t$ at department $i$. We define incumbents as scientists who are present in department $i$ the year prior to the star's arrival. $Star_{it-10}$ is set to 1 for years up to and including 10 years prior to the arrival of the star and 0 otherwise. $Star_{it+8}$ is set to 1 for all years eight years after the arrival of the star and 0 otherwise. The omitted category is one year prior to the star's arrival. The vertical bars correspond to 95% confidence intervals with department-clustered standard errors.

Figure 8: Kernel Density of Move Age and Fifth Order Age Polynomial Plot of Coefficients from Move Regression in Table 9

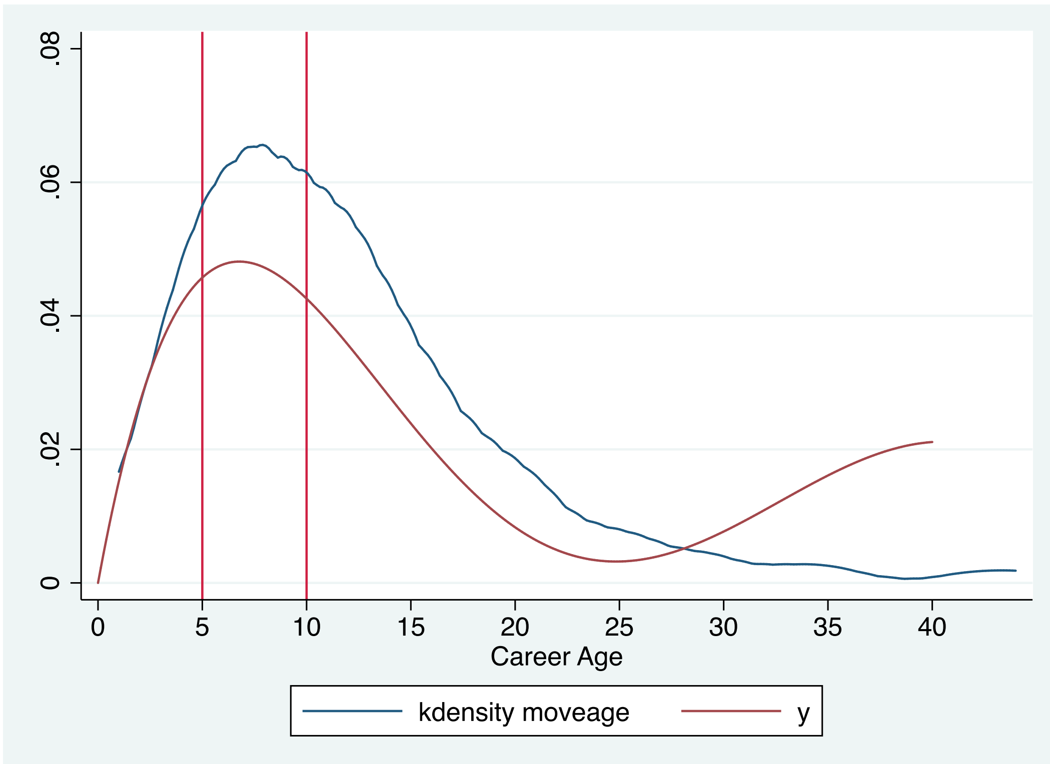


Table 1: Summary Statistics; N = 7,395

| Variables | Mean | Median | Std. Dev. | Min. | Max. |
|---|---|---|---|---|---|
| Output | 80.90 | 26 | 155.32 | 0 | 2500 |
| Output w/o Star | 76.75 | 24 | 151.60 | 0 | 2498 |
| Scientists | 21.67 | 14 | 24.23 | 1 | 175 |
| Incumbent Output | 17.61 | 2 | 53.83 | 0 | 1650 |
| Incumbents | 6.60 | 3 | 9.88 | 0 | 93 |
| Star | 0.43 | 0 | 0.49 | 0 | 1 |
| Joiner Quality | 36.53 | 14 | 78.34 | 1 | 2348 |
| Joiner Quality - Early Career | 27.97 | 11.5 | 56.16 | 1 | 1137 |
| Joiner Quality - Mid Career | 72.15 | 19 | 163.79 | 1 | 2925 |
| Joiner Quality - Late Career | 108.65 | 23 | 296.53 | 1 | 3242 |
| Output w/o Star - Related | 14.62 | 0 | 49.89 | 0 | 1687 |
| Output w/o Star - Unrelated | 62.13 | 20 | 127.91 | 0 | 2498 |
| Incumbent Output - Related | 3.93 | 0 | 18.61 | 0 | 719 |
| Incumbent Output - Unrelated | 13.68 | 1 | 48.96 | 0 | 1650 |
| Joiner Quality - Related | 21.21 | 0 | 94.29 | 0 | 1766 |
| Joiner Quality - Unrelated | 29.71 | 0 | 73.26 | 0 | 2348 |
| MoveRisk | 51.15 | 26 | 67.67 | 0 | 589 |

Table 2: Main Results

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Variable: | *Output* | *Output w/o Star* | *Output w/o Star* | *Incumbent Output* |
| $\text{Star}_{t-1}$ | $0.430^{**}$ | $0.392^{**}$ | $0.324^{**}$ | $-0.013$ |
| | (0.077) | (0.082) | (0.097) | (0.092) |
| Scientists | | | $0.011^{**}$ | |
| | | | (0.004) | |
| Incumbents | | | | $0.042^{**}$ |
| | | | | (0.007) |
| Department Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| Observations | 7140 | 7140 | 7140 | 4984 |
| Number of Departments | 255 | 255 | 255 | 178 |
| Log-Likelihood | -155577 | -151447 | -146349 | -55496 |

Notes: This table reports coefficients for four Poisson quasi-maximum likelihood (QML) regressions. Observations are at the $\text{department}_i$-$\text{year}_t$ level. *Output* refers to Citation-Weighted Publications. Columns 2 and 3 remove the Output of the arriving star. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who were present the year prior to the star's arrival. The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival and 0 otherwise. The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 3: Characteristics of Joining Scientists

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Variable: | *Joiner Quality* | *Joiner Quality* | *Joiner Quality* | *Joiner Quality* |
| Sample: | | *Early Career* | *Mid Career* | *Late Career* |
| $\text{Star}_{t-1}$ | $0.541^{**}$ | $0.678^{**}$ | $1.044^{**}$ | $0.853^{+}$ |
| | (0.124) | (0.132) | (0.260) | (0.467) |
| Department Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| Observations | 3629 | 3051 | 1539 | 735 |
| Number of Departments | 250 | 244 | 215 | 155 |

Notes: This table reports coefficients for four Poisson quasi-maximum likelihood (QML) regressions. Observations are at the $\text{department}_i$-$\text{year}_t$ level. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. The dependent variables in Columns 2, 3, and 4 are the mean stock of all scientists hired by department $i$ in year $t$ who have a career age of less than 10, between 10 and 20, and more than 20, respectively. The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival and 0 otherwise. Robust standard errors clustered at the department are in parentheses.

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 4: Output by Topically Related and Unrelated Scientists

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Variable: | | *Output w/o Star* | | |
| Sample: | *All* | *Incumbents* | *All* | *Incumbents* |
| SubSample: | *Related* | | *Unrelated* | |
| $\text{Star}_{t-1}$ | $0.924^{**}$ | $0.529^{**}$ | $0.191^{+}$ | $-0.143$ |
| | (0.240) | (0.182) | (0.107) | (0.107) |
| Scientists | $0.015^{*}$ | | $0.010^{*}$ | |
| | (0.007) | | (0.004) | |
| Incumbents | | $0.031^{*}$ | | $0.042^{**}$ |
| | | (0.013) | | (0.007) |
| Department Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| Observations | 4704 | 3472 | 7140 | 4984 |
| Number of Departments | 168 | 124 | 255 | 178 |

Notes: This table reports coefficients for four Poisson quasi-maximum likelihood (QML) regressions. Observations are at the $\text{department}_i$-$\text{year}_t$ level. The dependent variable, *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions split by the characteristics of the scientist. Columns 1-2 only include scientists who are topically related to the arriving star (make at least one reference in their papers to the arriving star), while Columns 3-4 only include scientists who are topically unrelated to the star (do not make any references to the papers of the arriving star). Columns 1 and 3 include all scientists, and Columns 2 and 4 include all incumbents present the year prior to the star's arrival. The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival and 0 otherwise. The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 5: Joiner Quality by Topically Related and Unrelated

| | (1) | (2) |
|---|---|---|
| Dependent Variable: | *Joiner Quality* | |
| SubSample: | *Related* | *Unrelated* |
| $\text{Star}_{t-1}$ | 1.676** | 0.390** |
| | (0.378) | (0.120) |
| Department Fixed Effects | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ |
| Observations | 2663 | 3629 |
| Number of Departments | 151 | 250 |

Notes: This table reports coefficients for two Poisson quasi-maximum likelihood (QML) regressions. Observations are at the $\text{department}_i\text{-year}_t$ level. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. The *Related* subsample consists of scientists who are topically related to the arriving star (make at least one reference in their papers to the arriving star) and the *Unrelated* subsample consists of scientists who are not topically related to the arriving star (do not make any references to the papers of the arriving star). The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival and 0 otherwise. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 6: Results Split by Rank

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | *Output w/o Star* | | *Incumbent Output* | | *Joiner Quality* | |
| Sample: | *Top 25* | *Non-top 25* | *Top 25* | *Non-top 25* | *Top 25* | *Non-top 25* |
| $\text{Star}_{t-1}$ | −0.094 | 0.221** | 0.031 | −0.021 | 0.431 | 0.907** |
| | (0.221) | (0.068) | (0.298) | (0.072) | (0.264) | (0.172) |
| Scientists | 0.006 | 0.022** | | | | |
| | (0.005) | (0.003) | | | | |
| Incumbents | | | 0.007 | 0.061** | | |
| | | | (0.005) | (0.006) | | |
| Department Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 2828 | 6468 | 672 | 4312 | 2772 | 6412 |
| Number of Departments | 101 | 231 | 24 | 154 | 99 | 229 |

Notes: This table reports coefficients for six Poisson quasi-maximum likelihood (QML) regressions. Observations are at the department$_i$-year$_t$ level. The dependent variable, *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who were present the year prior to the star's arrival. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. Odd columns include institutions that are ranked in the top 50 (as measured by citation-weighted publications), while even includes institutions outside the top 50. The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival and 0 otherwise. All specifications include institutions that never receive a star as control institutions in addition to department and year fixed effects. The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 7: Star Coauthorships

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | *Output w/o Star* | | | *Incumbent Output* | |
| Sample: | *Full* | *Related* | *Unrelated* | *Full* | *Related* | *Unrelated* |
| $\text{Star}_{t-1}$ | $0.259^{**}$ | $0.813^{**}$ | 0.151 | −0.095 | 0.284 | −0.142 |
| | (0.093) | (0.245) | (0.102) | (0.098) | (0.199) | (0.114) |
| Collaborations w/ Star | $0.031^{**}$ | $0.045^{*}$ | 0.021 | $0.161^{**}$ | $0.213^{**}$ | −0.005 |
| | (0.012) | (0.018) | (0.013) | (0.045) | (0.052) | (0.057) |
| Scientists | $0.011^{**}$ | $0.015^{*}$ | $0.010^{*}$ | | | |
| | (0.004) | (0.007) | (0.004) | | | |
| Incumbents | | | | $0.040^{**}$ | $0.028^{*}$ | $0.042^{**}$ |
| | | | | (0.006) | (0.013) | (0.007) |
| Department Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 7140 | 4704 | 7140 | 4984 | 3472 | 4984 |
| Number of Departments | 255 | 168 | 255 | 178 | 124 | 178 |

Notes: This table reports coefficients for six Poisson quasi-maximum likelihood (QML) regressions. Observations are at the $\text{department}_i\text{-year}_t$ level. The dependent variable, *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions split by the characteristics of the scientist. Columns 1 and 4 include all scientists. Columns 2 and 5 only include scientists who are topically related to the arriving star (make at least one reference in their papers to the arriving star). Columns 3 and 6 only include scientists who are topically unrelated to the star (do not make any references to the papers of the arriving star). Columns 1-3 include all scientists, and Columns 4-6 include all incumbents present the year prior to the star's arrival. The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival and 0 otherwise. The independent variable *Collaborations w/Star* is a count of the number of collaborations the star had with a scientist at institution $i$ in year $t$. The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 8: Main Results with Developmental Biology Controls

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| $\text{Star}_{t-1}$ | $0.294^{**}$ | 0.030 | $0.518^{**}$ |
| | (0.072) | (0.084) | (0.123) |
| Scientists | $0.016^{**}$ | | |
| | (0.004) | | |
| Incumbents | | $0.039^{**}$ | |
| | | (0.006) | |
| Devel. Biology $\text{Star}_{t-1}$ | $-0.134$ | $-0.158$ | 0.069 |
| | (0.091) | (0.138) | (0.177) |
| Devel. Biology $\text{Scientists}_{t-1}$ | $-0.003^{*}$ | $-0.002^{+}$ | $-0.001$ |
| | (0.001) | (0.001) | (0.001) |
| Department Fixed Effects | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ |
| Math and Psychology Controls | ✓ | ✓ | ✓ |
| Observations | 7140 | 4984 | 3629 |
| Number of Departments | 255 | 178 | 250 |

Notes: This table reports coefficients for three Poisson quasi-maximum likelihood (QML) regressions. Observations are at the $\text{department}_i$-$\text{year}_t$ level. *Output* refers to Citation-Weighted Publications. Columns 2 and 3 remove the output of the arriving star. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who are present the year prior to the star's arrival. The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival, and 0 otherwise. There are four control variables. *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. *Development Star* is a value of 1 if the year is greater than or equal to the year of a developmental biology star arriving at institution $i$ and 0 otherwise. *Development Scientists* is a count of the number of developmental biology scientists present at institution $i$ in year $t$. All specifications include controls for the arrival of a star and the number of scientists in the focal department's Math and Psychology departments. Robust standard errors clustered at the department are in parentheses.

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 9: IV - Mobility as a Function of Age

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Variable: | *Move* | *Move* | *Move* | *Move* |
| Estimation | OLS | Logit | OLS | Logit |
| Age | $0.017^{**}$ | $0.655^{**}$ | $0.019^{**}$ | $0.725^{**}$ |
| | (0.002) | (0.145) | (0.002) | (0.151) |
| $\text{Age}^2$ | $-0.002^{**}$ | $-0.070^{**}$ | $-0.002^{**}$ | $-0.070^{**}$ |
| | (0.000) | (0.019) | (0.000) | (0.019) |
| $\text{Age}^3$ | $0.000^{**}$ | $0.003^{**}$ | $0.000^{**}$ | $0.003^{**}$ |
| | (0.000) | (0.001) | (0.000) | (0.001) |
| $\text{Age}^4$ | $-0.000^{**}$ | $-0.000^{+}$ | $-0.000^{**}$ | $-0.000^{+}$ |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| $\text{Age}^5$ | $0.000^{**}$ | 0.000 | $0.000^{**}$ | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Scientist Fixed Effects | | | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| Observations | 7872 | 5397 | 7872 | 7872 |
| $R^2$ | 0.05 | | 0.05 | |
| Log-Likelihood | | -970 | | -781 |
| F-stat of excluded instruments | 22.24 | | 22.14 | |
| $\chi^2$ of excluded instruments | | 77.82 | | 40.67 |

Notes: This table reports coefficients for five specifications estimated by OLS. Observations are at the $\text{scientist}_i$-$\text{year}_t$ level. The dependent variable *Move* is equal to 1 if scientist $i$ moves in year $t$ and 0 otherwise. *Age* is the career age of the scientist (the number of years elapsed since the author's first publication). Robust standard errors clustered at the scientist are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 10: IV - Location Choice as a Function of Legacy

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable: | *Move* | *Move* | *Move* | *Move* | *Move* | *Move* |
| Estimation | OLS | Logit | OLS | OLS | Logit | Logit |
| Prior Coauthorships | $0.005^{**}$ | $2.850^{**}$ | $0.006^{**}$ | $0.005^{**}$ | $3.757^{**}$ | $2.689^{**}$ |
| | (0.001) | (0.208) | (0.001) | (0.001) | (0.224) | (0.219) |
| Prior State | $0.001^{**}$ | $0.655^{**}$ | $0.001^{*}$ | $0.001^{**}$ | $0.415^{+}$ | $0.983^{**}$ |
| | (0.000) | (0.208) | (0.000) | (0.000) | (0.229) | (0.239) |
| Prior Country | $0.002^{+}$ | $0.731^{*}$ | $0.002^{+}$ | $0.002^{*}$ | $0.658^{+}$ | $0.886^{**}$ |
| Scientist Fixed Effects | | | ✓ | | ✓ | |
| Department Fixed Effects | | | | ✓ | | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 150960 | 150960 | 150960 | 150960 | 25755 | 59200 |
| $R^2$ | 0.01 | | 0.01 | 0.01 | | |
| Log-Likelihood | | -809 | | | -463 | -606 |
| F-stat | 20.30 | | 22.25 | 24.09 | | |
| $\chi^2$ | | 223.74 | | 328.45 | 212.41 | |

Notes: Observations are at the scientist$_i$-department$_j$ level. The dependent variable, *Move*, is equal to 1 if scientist $i$ ever moves to department $j$. *Prior Coauthorships* is equal to 1 if scientist $i$ has at least one coauthor (formed prior to the scientist's move) at department $j$ and 0 otherwise. *Prior State* is equal to 1 if scientist $i$ has previously lived in the state that department $j$ is in and 0 otherwise. *Prior Country* is equal to 1 if scientist $i$ has previously lived in the same country as department $j$ (excluding USA) and 0 otherwise. Robust standard are in parentheses. The standard errors are clustered at the level of the department in Columns 4 and 6 and at the scientist level in all other columns. F-test and $\chi^2$ are tests of the key instruments (excluding all fixed effects).

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 11: Instrumental Variable Results

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Estimations: | *OLS* | *2SLS* | *2SLS* | *2SLS* | *2SLS* |
| Dependent Variable: | *Star* | *Output w/o Star* | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| MoveRisk | $0.105^{**}$ | | | | |
| | (0.017) | | | | |
| $\text{Star}_{t-1}$ | | $1.434^{**}$ | $0.981^{*}$ | 0.392 | $2.189^{**}$ |
| | | (0.452) | (0.468) | (0.336) | (0.726) |
| Scientists | | | $0.024^{**}$ | | |
| | | | (0.004) | | |
| Incumbents | | | | $0.081^{**}$ | |
| | | | | (0.003) | |
| Department Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 7395 | 7140 | 7140 | 7140 | 7140 |
| Number of Departments | 255 | 255 | 255 | 255 | 255 |
| Angrist-Pischke F-test | 62.12 | 62.12 | 49.19 | 71.70 | 62.12 |

Notes: Observations are at the $\text{department}_i$-$\text{year}_t$ level. All dependent variables except for Star have had a 1 added to them and converted to natural logarithms. Column 1 is the first-stage regression of *MoveRisk* on the endogenous variable *Star*. The variable *MoveRisk* is a dummy set to 1 if the number of star scientist's who are at risk of moving to department $j$ (have prior coauthorship and either prior state or country experience from Table 10 and are between the career ages of 5 and 10 [see Figure 8 and Table 9]) is greater than or equal to the median number of 15, and 0 otherwise. The independent variable *Star* is a value of 1 if the year is greater than or equal to the year of the star's arrival and 0 otherwise. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 12: Instrumental Variable Results - II

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable: | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| SubSample: | | *Related* | | | *Unrelated* | |
| $\text{Star}_{t-1}$ | 1.508** | 1.380** | 0.824* | 0.397 | −0.047 | 1.660** |
| | (0.491) | (0.313) | (0.396) | (0.499) | (0.333) | (0.569) |
| Scientists | 0.028** | | | 0.027** | | |
| | (0.004) | | | (0.004) | | |
| Incumbents | | 0.025** | | | 0.081** | |
| | | (0.003) | | | (0.003) | |
| Department Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 7140 | 7140 | 7140 | 7140 | 7140 | 7140 |
| Number of Departments | 255 | 255 | 255 | 255 | 255 | 255 |
| Angrist-Pischke F-test | 49.19 | 71.70 | 62.12 | 49.19 | 71.70 | 62.12 |

Notes: This table reports coefficients for six IV regressions estimated by two-stage least squares (2SLS). All dependent variables have had a 1 added to them and converted to natural logarithms. Observations are at the $\text{department}_i$-$\text{year}_t$ level. The *Related* subsample consists of scientists who are topically related to the arriving star (make at least one reference in their papers to the arriving star) and the *Unrelated* subsample consists of scientists who are not topically related to the arriving star (do not make any references to the papers of the arriving star). The variables *Star* is treated as endogenous and instrumented with the variable *MoveRisk* (see Column 1 of Table 11 for first-stage estimates). The variable *MoveRisk* is a dummy set to 1 if the number of star scientist's who are at risk of moving to department $j$ (have prior coauthorship and either prior state or country experience from Table 10 and are between the career ages of 5 and 10 [see Figure 8 and Table 9]) is greater than or equal to the median number of 15, and 0 otherwise. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 13: Star Departure Results

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| Star Depart$_{t-1}$ | $-0.214^{*}$ | $-0.286^{**}$ | 0.241 |
| | (0.105) | (0.102) | (0.152) |
| Scientists | $0.012^{**}$ | | |
| | (0.004) | | |
| Incumbents | | $0.038^{**}$ | |
| | | (0.005) | |
| Department Fixed Effects | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ |
| Observations | 7140 | 3416 | 3835 |
| Number of Departments | 255 | 122 | 250 |
| Log-Likelihood | -150674 | -61313 | -121149 |

Notes: This table reports coefficients for three Poisson quasi-maximum likelihood (QML) regressions. Observations are at the department$_i$-year$_t$ level. *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions split by the characteristics of the scientist. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who were present the year prior to the star's arrival. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. The independent variable *Depart Star* is a value of 1 if the year is greater than or equal to the year of the star's departure and 0 otherwise. The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 14: Arrival of a National Academies Scientist Results

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| Arrive National Academies Scientist$_{t-1}$ | 0.093 | $-0.334^{*}$ | 0.409 |
| | (0.098) | (0.156) | (0.318) |
| Scientists | $0.009^{+}$ | | |
| | (0.005) | | |
| Incumbents | | $0.024^{**}$ | |
| | | (0.005) | |
| Department Fixed Effects | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ |
| Observations | 868 | 868 | 868 |
| Number of Departments | 31 | 31 | 31 |
| Log-Likelihood | -33994 | -19354 | -25520 |

Notes: This table reports coefficients for three Poisson quasi-maximum likelihood (QML) regressions. Observations are at the department$_i$-year$_t$ level. *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions split by the characteristics of the scientist. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who were present the year prior to the star's arrival. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. The independent variable *Arrive National Academies Scientist* is a value of 1 if the year is greater than or equal to the year of the arrival of a scientist who is a member of the National Academies of Sciences and 0 otherwise. The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 15: Arrival of a National Academies Scientist Rank Results

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| Arrive National Academies Scientist$_{t-1}$ | $-0.190$ | $-0.553^{**}$ | 0.002 |
| X Top 25 | (0.153) | (0.174) | (0.258) |
| Arrive National Academies Scientist$_{t-1}$ | $0.713^{**}$ | 0.205 | $0.857^{*}$ |
| X Non-top 25 | (0.154) | (0.213) | (0.376) |
| Scientists | $0.008^{+}$ | | |
| | (0.004) | | |
| Incumbents | | $0.023^{**}$ | |
| | | (0.005) | |
| Department Fixed Effects | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ |
| Observations | 868 | 868 | 868 |
| Number of Departments | 31 | 31 | 31 |
| Log-Likelihood | -31059 | -18672 | -25015 |

Notes: This table reports coefficients for three Poisson quasi-maximum likelihood (QML) regressions. Observations are at the department$_i$-year$_t$ level. *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions split by the characteristics of the scientist. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who were present the year prior to the star's arrival. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. The independent variable *Arrive National Academies Scientist* is a value of 1 if the year is greater than or equal to the year of the arrival of a scientist who is a member of the National Academies of Sciences and 0 otherwise. This variable is interacted with two indicators each set to 1 if the department the scientist arrived at a top 25 department (at the year of arrival) or a non-top 25 department (at the year of arrival). The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 16: Becoming a National Academies Scientist Results

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| Became National Academies Scientist$_{t-1}$ | 0.172 | $-0.153$ | 0.167 |
| | (0.160) | (0.207) | (0.195) |
| Scientists | 0.013$^{**}$ | | |
| | (0.004) | | |
| Incumbents | | 0.025$^{**}$ | |
| | | (0.006) | |
| Department Fixed Effects | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ |
| Observations | 896 | 896 | 896 |
| Number of Departments | 32 | 32 | 32 |
| Log-Likelihood | -32875 | -18817 | -33429 |

Notes: This table reports coefficients for three Poisson quasi-maximum likelihood (QML) regressions. Observations are at the department$_i$-year$_t$ level. *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions split by the characteristics of the scientist. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who were present the year prior to the star's arrival. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. The independent variable *Became National Academies Scientist* is a value of 1 if the year is greater than or equal to the year an incumbent scientist becomes a member of the National Academies of Sciences and 0 otherwise. The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$

Table 17: Becoming a National Academies Scientist Rank Results

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | *Output w/o Star* | *Incumbent Output* | *Joiner Quality* |
| Became National Academies Scientist$_{t-1}$ | 0.109 | $-0.198$ | 0.213 |
| X Top 25 | (0.185) | (0.239) | (0.222) |
| Became National Academies Scientist$_{t-1}$ | 0.385$^{*}$ | 0.020 | 0.126 |
| X Non-top 25 | (0.156) | (0.175) | (0.234) |
| Scientists | 0.013$^{**}$ | | |
| | (0.004) | | |
| Incumbents | | 0.025$^{**}$ | |
| | | (0.006) | |
| Department Fixed Effects | ✓ | ✓ | ✓ |
| Year Fixed Effects | ✓ | ✓ | ✓ |
| Observations | 896 | 896 | 896 |
| Number of Departments | 32 | 32 | 32 |
| Log-Likelihood | -32660 | -18768 | -33423 |

Notes: This table reports coefficients for three Poisson quasi-maximum likelihood (QML) regressions. Observations are at the department$_i$-year$_t$ level. *Output w/o Star*, is the Citation-Weighted Publications in year $t$ net of the arriving star's contributions split by the characteristics of the scientist. *Incumbent Output* is a count of the Citation-Weighted Publication of all scientists at department $i$ who were present the year prior to the star's arrival. *Joiner Quality* is the mean stock of all scientists hired by department $i$ in year $t$. The independent variable *Became National Academies Scientist* is a value of 1 if the year is greater than or equal to the year an incumbent scientist becomes a member of the National Academies of Sciences and 0 otherwise. This variable is interacted with two indicators each set to 1 if the institution the scientist arrived at a top 25 department (at the year of arrival) or a non-top 25 department (at the year of arrival). The two control variables *Scientists* and *Incumbents* are a count of the number of scientists present at department $i$ in year $t$ and the number of incumbent scientists (who are present the year prior to the star's arrival) who are present at department $i$ in year $t$, respectively. Robust standard errors clustered at the department are in parentheses. $^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$