Paper to be presented at
the DRUID16 20th Anniversary Conference
Copenhagen, June 13-15, 2016

Knowledge Diversification and Stars: Implications for the Knowledge Creation Process

Florenta Teodoridis
University of Southern California
Marshall School of Business
florenta.teodoridis@marshall.usc.edu

Keyvan Vakili
London Business School
Strategy and Entrepreneurship
kvakili@london.edu

Michael Bikard
London Business School
Strategy and Entrepreneurship
mbikard@london.edu

Abstract

Stars – researchers in the right tail of the productivity distribution – play a significant role in the creation of new knowledge. This paper explores a differential role of stars driven by variation in their breadth of knowledge and motivated by the cumulative nature of innovation that increases returns to specialization but also increases demand for coordination across the knowledge frontier. Specifically, we exploit the collapse of the Soviet Union as a natural experiment that led to an unexpected forward movement of the scientific knowledge frontier in some fields of theoretical mathematics but not in others. We find that "specialist" stars (and their collaborators) were able to leverage the opportunity to increase their productivity whereas "generalist" ones were hurt by the sudden change in scientific landscape (though their collaborators were not). These results point to breadth of knowledge as a source of heterogeneity influencing the role of stars in knowledge production. Furthermore, the results suggest an under-recognized downsides of knowledge brokerage in creative work.

Jelcodes:O31,O33

Knowledge Diversification and Stars: Implications for the Knowledge Creation Process

February 2016

Abstract

Stars – researchers in the right tail of the productivity distribution – play a significant role in the creation of new knowledge. This paper explores a differential role of stars driven by variation in their breadth of knowledge and motivated by the cumulative nature of innovation that increases returns to specialization but also increases demand for coordination across the knowledge frontier. Specifically, we exploit the collapse of the Soviet Union as a natural experiment that led to an unexpected forward movement of the scientific knowledge frontier in some fields of theoretical mathematics but not in others. We find that "specialist" stars (and their collaborators) were able to leverage the opportunity to increase their productivity whereas "generalist" ones were hurt by the sudden change in scientific landscape (though their collaborators were not). These results point to breadth of knowledge as a source of heterogeneity influencing the role of stars in knowledge production. Furthermore, the results suggest an under-recognized downsides of knowledge brokerage in creative work.

1. Introduction

Human capital is essential to the production of new scientific knowledge. As early as 1926, Lotka observed that 6% of physicists produced more than 50% of all academic publications, motivating a subsequent enduring focus on highly productive individuals and their role in the knowledge creation process. Researchers in the right tail of the output distribution— i.e., stars — contribute tremendously to science, not only through their extraordinary productivity, but also because they influence the productivity of their peers (Azoulay et al., 2010; Waldinger, 2011; Oettl, 2013; Agrawal et al., 2014).

The type of skills that stars leverage to push science forward remains unclear, however. Highlighting the importance of brokerage and recombination, prior work on creativity has suggested that individuals who are able to bridge across scientific specialties (here called "generalists") should be at an advantage (e.g., Hargadon and Sutton 1997; Fleming, Mingo, and Chen 2007). This should be especially true in scientific research considering the growing "burden of knowledge"—i.e., the increasing fragmentation of the scientific frontier into narrower niches (Jones, 2009; Teodoridis, 2016). Another, less tested, stream of work has however emphasized the importance of knowledge depth for creativity (Kaplan and Vakili 2015). According to this view, individuals with deep expertise (here called "specialists") might reach superior levels of creative performance because they have access to insights that are invisible to most. The literature preoccupied with the role of stars in knowledge production doesn't distinguish between the two types of scientists. Yet, the distinction has important implications. It might inform theory by explaining the role of the two types of skills in scientific research in particular and more broadly in creative work. Also, more practically, if brokering knowledge is not always a successful strategy for scientists, then current incentives fostering inter-disciplinary work (e.g., Stephan 2012) might negatively impact scientific productivity.

In this paper, we distinguish between "specialists" and "generalist" stars to explore the relative importance of expertise in the knowledge creation process. First, we set to establish that variation in stars' breadth of expertise leads to a differential outcome in the knowledge production process. Second, and related, we take a first step in exploring how the variation in breadth of expertise and, hence, brokerage influence the knowledge creation process. We focus on the continuous forward movement of the knowledge frontier that opens up opportunities for further discoveries. It is unclear ex-ante how stars adjust to the emergence of new scientific opportunities contingent on variation in their breadth and depth of expertise. Prior literature has overwhelmingly considered that breadth of knowledge (knowledge brokers) might be disproportionately able to spot new opportunities because of their wider exposure to a larger number of knowledge niches (Hargadon and Sutton 1997; Fleming 2001; Burt 2004; Uzzi and Spiro 2005; Audia and Goncalo 2007; Teodoridis, 2016). In contrast, we argue that an under-recognized advantage of specialization and deep field-specific expertise is that it allows scientists to take advantage of developments that occur at the knowledge frontier. The assertion is motivated by the "knowledge burden"[1] hypothesis (Jones 2009, 2010) which suggests increasing returns to specialization. Correspondingly, brokerage is likely to be associated with shallower field-specific expertise, and hence generalist star scientists might be at a greater risk of seeing some of their knowledge be rendered obsolete with forward movements of the knowledge frontier.

Testing these hypotheses empirically is difficult because movements in the scientific frontier tend to be endogenous to the work of stars and invisible to the empiricist. In other words, when observing the relative performance of specialist and generalist stars, how can we adjudicate whether it is due to differences in opportunities, in abilities, or in some other broader social change? To get around this identification challenge, we propose an empirical strategy based on exploiting a sudden and unexpected forward movement of the knowledge frontier. Our empirical strategy therefore complements prior studies of star scientists that have used other "exogenous shocks" such as college student random assignment to dorm rooms (Sacerdote, 2001; Mas and Moretti, 2009) and unexpected star deaths (Azoulay et al, 2012; Oettl, 2013).

We focus on the sudden collapse of the Soviet Union, and its impact in theoretical mathematics. Before 1989, the Soviet Union was at the forefront of research in theoretical mathematics. Despite their lead, Communist government officials forced their researchers to work in isolation from the rest of the world. For example, with few exceptions, scholars were prohibited from traveling,

[1] Knowledge accumulation due to forward movements of the frontier place a knowledge burden on scientists leading to narrower specialization.

publishing outside of the Soviet Union, and from accessing foreign publications without case-by-case government approval. Thus, when the Iron Curtain fell and Soviet science became widely available, the knowledge frontier in mathematics outside the USSR experienced a shock. Furthermore, Soviet mathematicians focused their contributions on certain areas of theoretical mathematics more than on others due to reason uncorrelated to the Soviet regime, but rather as a result of historical path dependency. For example, Soviet mathematics community was very advanced relative to the rest of the world in some subfields of theoretical mathematics, such as "partial differential equations" and "operator theory," and much less so in others, such as "abstract harmonic analysis" and "sequences, series, summability."

To test our hypotheses, we exploit the variation in the degree of knowledge shock across subfields using a difference-in-differences type of analysis using world-wide data on academic publications in theoretical mathematics. Specifically, we compare the academic output of star mathematicians working outside the USSR in areas of theoretical mathematics where Soviets made strong advancements relative to areas of theoretical mathematics where Soviet made comparably weaker contributions before and after the collapse of the Soviet Union. To be clear, Soviet mathematics made great advancements along all subfields of theoretical mathematics. In our empirical analysis, we exploit the variation in strength of these contributions across subfields of theoretical mathematics. We use 21 years of publication data in theoretical mathematics covering the period 1980-2000, 10 years before and after the fall of the Iron Curtain.

To identify fields of mathematics influenced by Soviet expertise to various degrees, we categorize our set of academic publications using the internationally recognized Mathematics Subject Classification codes developed and assigned by the Mathematical Reviews division of the American Mathematical Society. We follow the Soviet-rich versus –poor subfield classification in Agrawal, Goldfarb and Teodoridis (2016) which is based on the fraction of publications produced by Soviet researchers relative to mathematicians from US and robust against considering a classification based on publications of Soviet mathematicians relative to the rest of the world.

Our dataset of individual researchers focuses on mathematicians outside the USSR and drops observations of Soviet researchers. We do so because we are interested in the effects of the sudden forward movements of the frontier on stars, rather than in the effect on the researchers fueling the shift. In our analysis, we take into account the presence of Soviet mathematicians in labor markets to the extent they influence output of non-Soviet mathematicians through, for example, collaboration on academic publications. Our results remain robust.

Our findings broadly confirm that variation in stars' breadth of expertise has a differential impact on knowledge output of stars contingent on their knowledge breadth. First, we find that the Soviet shock had no aggregate effect on the productivity of stars in Soviet-rich versus –poor subfields of mathematics. However, when taking into account the variation in breadth of expertise among stars, we find evidence of a differential impact on the productivity of specialist and generalist stars. Specifically, we find evidence of a disproportionate increase in productivity of specialist stars in Soviet-rich subfields relative to Soviet-poor specialist stars, after the collapse of the Soviet Union. In contrast, we find evidence of a disproportionate decrease in the productivity of generalist stars in Soviet-rich subfields relative to Soviet-poor, after the collapse.

We further explore consequences on stars' collaborators, motivated by literature findings suggesting positive spillover effects on stars' peers (e.g., Azoulay et al., 2010; Waldinger, 2011; Oettl, 2013; Agrawal et al., 2014). Aligned with this literature, we find evidence of a disproportionate increase in academic output of specialist stars' collaborators in Soviet-rich subfields, after the collapse of the Soviet Union, when compared to collaborators of specialist stars in Soviet-poor subfields. However, we don't find evidence of a similar impact on the productivity of generalist stars' collaborators. This finding is aligned with the hypothesized knowledge brokerage role of generalist stars that is distinct from the knowledge creation role of specialist stars.

Next, we explore the impact of the sudden advance of the knowledge frontier on scientists' rate of collaboration, as a mechanism of identified effects on productivity of scientists. We do so since the theory suggests that specialists ought to collaborate more after the shock because the deeper knowledge base requires increasing collaboration as researchers specialize on increasingly narrow niches (Jones 2009, Jones 2010, Agrawal et al., 2016). Furthermore, specialist stars' deep knowledge is likely to be particularly attractive to collaborators. Hence, we expect only the increase in specialist stars' and their collaborators' productivity to be associated with increases in collaboration rates. Our empirical analysis brings support to this prediction.

This paper attempts to contribute to prior literature in three ways. First, our results contribute to the literature on stars in science by showing that the exclusive focus on stars' high-levels of productivity conceals important differences in their role in the knowledge production process, driven by variation in their breadth of knowledge. Second, we theorize about the benefits of specialization in creative work and test our predictions empirically. Our results show that specialization provides unique and hitherto under-recognized benefits in creative work. Third, we document an important downside to brokering ideas across fields in science. Those scientists in our data that broker ideas across specialties were not only unable to benefit from the opportunities stemming by the sudden shock to the landscape of scientific opportunities, but that they were in fact disproportionately hurt by the displacement of the knowledge frontier. Presumably, their knowledge is more at risk of becoming obsolete.

We structure the remainder of the paper as follows. In Section 2, we describe the historical context of our empirical setting as it pertains to our identification strategy. In Section 3, we discuss our data collection and variable construction process, followed by a description of our empirical strategy in Section 4. We present results in Section 5 and conclude in Section 6.

2. Mathematics and the Soviet Union

Our empirical strategy relies on the assertion that the collapse of the Soviet Union caused an outward shift in the knowledge frontier in theoretical mathematics and that it did so more for some subfields than others. We base this claim on three main observations: 1) the Soviet Union's effect on the knowledge frontier in theoretical mathematics was significant for scientific advancements in mathematics, 2) the Soviet Union's effect on the knowledge frontier was greater in some subfields than others and the reason for this differential impact is not correlated with active efforts to focus advancements on areas of research away from the rest of the world, and 3) Soviet mathematicians conducted their advancements in secrecy from the outside world due to reasons related to communism government ruling.

Agrawal, Goldfarb and Teodoridis (2016) provide extensive discussion attesting to all these points. First, the Soviet Union was, and Russia continues to be, a world-renowned center of scientific research, with mathematics holding a prominent position. Scholarly research in mathematics attracted great minds as it was uniquely detached from politics, conferred status and prestige, and offered financial rewards superior to many other occupations. Second, while Soviet mathematics was strong across the entire spectrum of mathematics, Soviet mathematicians made the greatest advancements more in some subfields than others (Graham, 1994). Moreover, these differences reflect historical path dependency. Specifically, some subfields of theoretical mathematics built on strong mentorship from early 1900s and thus continued to attract bright minds later on (Borjas and Doran, 2012). For example, the success of Moscow mathematics can be traced back to Ergorov and his student N. N. Luzin (Tikhomirov, 2007) whose famous work was mainly focused on the theory of functions. The same didn't hold true for other areas of theoretical mathematics, such as algebraic geometry (Borjas and Doran, 2012).

Last, Soviet knowledge in theoretical mathematics was kept secret from the outside world due to communist government rules and regulations. The Communist government kept strict control on international travel. Academics who wished to attend foreign conferences had to go through a stringent and lengthy approval process, with many researchers blacklisted because of "tainted" backgrounds. The few approvals granted were typically for travel in Eastern Europe (Ganguli, 2012). Additionally, Soviet researchers were prevented from publishing their findings, traveling to conferences, communicating or collaborating with non-Soviets, and even accessing non-Soviet references. As such, Soviet advancements in mathematics remained relatively unknown to the outside world until the collapse of the Iron Curtain (Graham and Dezhina, 2008) when they were suddenly made available.[2]

[2] The following quote, from an article published on May 8, 1990 in the New York Times, provides an indication of the sudden outward shift of the knowledge frontier: *Persi Diaconis, a mathematician at Harvard, said: ''It's been fantastic. You just have a totally fresh set of insights and results.'' Dr. Diaconis said he recently asked Dr. Reshetikhin for help with a problem that had stumped him for 20 years. ''I had asked everyone in America who had any chance of knowing'' how to solve a problem of determining how organized sets become disorganized, Dr. Diaconis*

All in all, the fall of the Iron Curtain provides a plausible natural experiment differentially affecting the forward movement of the knowledge frontier across subfields of theoretical mathematics. This historical event was exogenous and unexpected to the mathematics research community. Furthermore, it is important to note that we don't require a full seclusion of Soviet knowledge before collapse, but rather that enough knowledge was suddenly made available to move the knowledge frontier forward and unexpectedly.

3. Data

We collect data on every academic publication in theoretical mathematics published during the 21-year period 1980 – 2000, 10 years before and after the collapse of the Soviet Union in 1989. We follow Borjas and Doran's (2012) and Agrawal, Goldfarb and Teodoridis' (2016) interpretation of historical events in focusing on 1990 as the first year when academic seclusion was significantly lessened. Our results remain robust to choosing neighboring years as the cut-offs for our estimations.

We collect our academic publication data from the Mathematical Reviews (MR) division of the American Mathematical Society (AMS). The MR database includes all worldwide academic publications in mathematics covering the three main categories of mathematics: mathematical foundations (including history and biography), pure or theoretical mathematics, and applied mathematics. Our focus is on theoretical mathematics, which includes analysis, algebra, and geometry.

In our empirical estimation we rely on the variation in the degree to which the knowledge frontier moved forward, predicated on the observation that Soviet mathematicians made grater contributions to some subfields of theoretical mathematics but not to others. We rely on the careful and exhaustive work of the Mathematical Reviews division, which classifies each paper in

said. No one could help. But Dr. Reshetikhin told Dr. Diaconis that Soviet scientists had done a lot of work on such problems. ''It was a whole new world I had access to,'' Dr. Diaconis said.

mathematics using Mathematics Subject Classification (MSC) codes. The MSC schema is internationally recognized and facilitates targeted searches on research subjects across all subfields of mathematics. The MR team assigns precisely one primary MSC code to each academic publication uploaded to the MR database. We follow the ranking in Agrawal, Goldfarb and Teodoridis (2016) of the 33 primary MSC codes of theoretical mathematics indicating the degree to which Soviets contributed to a particular subfield before the collapse of the Soviet Union. We list the 33 subfields and their rank in Table 1.

Next, we convert our dataset to the individual unit of analysis and proceed to identify three indexes: a star indicator, an index of Soviet knowledge exposure and an index of breadth of knowledge. First, we identify stars through a two-step process. We start by identifying scientists who won at least one prestigious prize in theoretical mathematics between 1980 and 1989. We follow the list of prestigious prizes identified by Borjas and Doran (2015): Fields Medal, Wolf Prize, Cole Algebra Prize, Bocher Prize, Veblen Prize and Salem Prize. We identify 84 scientists who won one of the prizes during the period of interest before the collapse of the Soviet Union (1980-1989). Next, we supplement our list with scientists identified as having a productivity level in the top 5% of the distribution based on counts of academic publications in the period before the collapse of the Soviet Union. We identify 2,268 such scientists, with 33 of them also being captured in our group of prize winning stars. Our results remain robust to considering definitions of stars with productivity levels in the 10% of the productivity distribution.

Second, we construct an index of Soviet exposure for each scientist in our dataset who publishes between 1980 and 1989. The index is calculated as a sum of percentages of publications in each of the 33 subfields of theoretical mathematics, weighted by the ranking of the 33 subfields, per individual, for the entire period before the collapse of the Soviet Union. The higher the percentage of academic publications in one's publication portfolio in subfields where Soviets made higher contributions, the higher the Soviet impact index. Formally, we calculate:

$$SovietImpactIndex_i = \sum_{s=1}^{33} \frac{PubCount_{si}}{PubCount_i * SubfieldRankOrder_s}$$

where $PubCount_{si}$ is the total count of publications of scientist i and subfield s, $PubCount_i$ is the total count of publications of scientist i and $SubfieldRankOrder_s$ is the rank order of the corresponding s subfield of theoretical mathematics. The calculation takes into account the full publication portfolio during the period before the collapse (1980-1989). In our sample of stars, the minimum value of the Soviet impact index is 0.0303 and the maximum value is 0.9275. Scientists in the bottom 1% of the Soviet impact distribution have an index of 0.0313 and below. Scientists in the top 1% of the Soviet impact distribution have an index of 0.4560 and above. The average is 0.0714, the mean is 0.1077 and the standard deviation is 0.1008. In our main specification, we define stars least affected by the Soviet shock as scientists with a Soviet impact index in the bottom 5% of the distribution (below 0.033) and stars most affected by the shock as scientists with a Soviet impact index in the top 5% of the distribution (above 0.333). Our results remain robust to considering different cut-off values that capture the variation in Soviet impact.

Last, we construct and index of diversification at the individual level capturing the heterogeneity in breadth of knowledge based on individual scientists' publication portfolio during the period before the collapse of the Soviet Union (1980-1989). The index is calculated as one minus the Euclidian distance in the multidimensional space of 33 subfields of theoretical mathematics and is based on percentages of publications in each of the 33 subfields, per scientist. By definition, the Euclidian distance is equal to the square root of the Herfindahl index. Our results remain robust when considering a diversification index based on the Herfinadahl. Formally, we calculate:

$$DiversificationIndex_i = 1 - \sqrt{\sum_{s=1}^{33}(\frac{PubCount_{s,i}}{PubCount_i})^2}$$

By construction, the higher the value, the higher the diversity of research portfolio areas at the individual level i. Furthermore, the diversification measure is higher or equal to 0 and never 1. The highest possible value of the diversification index is equal to 0.83 and would characterize researchers who publish an equal percentage of their publication portfolio across the 33 subfields of theoretical mathematics. The lowest diversification index is equal to 0, and characterizes researchers who exclusively publish in one subfield of theoretical mathematics.

In our main sample of stars, the highest diversification index is 0.5419 and the lowest is 0. Scientists in the bottom 1% of the diversification distribution have an index value of 0 and those in the top 1% have an index value of 0.4922 and above. In our main specification, we define top diversified stars as scientists with a diversification index in the top 5% of the distribution (above 0.3919) and we define specialist stars as scientists with a diversification index in the bottom 5% of the diversification distribution (values of 0). Our results remain robust to considering alternative cut-off points that continue to capture the variation in breadth of expertise.

4. Estimation Strategy

Our main estimation strategy is a difference-in-difference estimation which compares productivity of stars most ("treated") and least ("controls") affected by the sudden forward movement of the knowledge frontier in theoretical mathematics, before and after the collapse of the Soviet Union. In other words, we examine the difference between treated and control stars in two periods, before and after the treatment. Thus, we distinguish between the change in productivity of stars that is directly attributable to the shift in the knowledge frontier from the underlying differences between treated and control stars as well as the underlying changes in publication patterns of stars in theoretical mathematics over time. We measure productivity as a count of academic publications per scientist per year, from 1980 until 2000, 10 years before and 10 years after the collapse of the Soviet Union. Formally, we estimate:

$$PubCount_{i,t} = \alpha(SovietRichStar_i * AfterIronCurtain_t) + I_i + \gamma_t + \varepsilon_{i,t} \quad (1)$$

$PubCount_{i,t}$ is a count of academic publications of author i in year t. $SovietRichStar_i$ is an indicator variable equal to 1 if scientist i belongs to the treated group and 0 otherwise. $AfterIronCurtain_t$ is an indicator variable equal to 1 if year of observation t is after 1989 and 0 otherwise. This applies to scientists in both treated and control groups. We include individual and time fixed effects, hence the main effects $SovietRichStar_i$ and $AfterIronCurtain_t$ drop out of the estimating equation.

We are interested in the estimated coefficient on the interaction between $SovietRich_i$ and $AfterIronCurtain_t$, which equals 1 for scientists in the treated group after the knowledge shock and equals 0 for all others. We interpret a positive estimated value of this coefficient as implying that the average productivity of stars most affected by the forward movement of the knowledge frontier increased disproportionately relative to the productivity of stars least affected by Soviet work, after the knowledge shock. We estimate this relationship separately for stars with a high and low diversity index. Next, we repeat the estimation for stars' coauthors, identified as scientists who published at least once with a star in the period before the collapse of the Soviet Union (1980-1989). As with our stars sample, we exclude Soviet collaborators from our sample of star collaborators.

After establishing the effect of the forward moving frontier on the productivity of stars and their coauthors, while taking into account the variation directly attributable to stars' heterogeneity in breadth of expertise, we turn our attention to collaboration as a mechanism of these observed effects. To do so, we repeat our main estimating equation where we replace the dependant variable with a measure of collaboration. We consider both the extensive margin of collaboration (number of distinct collaborators per year) as well as the intensive margin of collaboration (number of coauthorship instances per year).

Because all of our dependent variables are count variables, we use conditional fixed-effect panel Possion model with robust standard errors clustered at the individual level in all of our regressions.

5. Results

We start with a baseline result estimating changes in productivity of treated and control stars, after the collapse of the Soviet Union, without taking into account the hypothesized effect of variation in breadth of knowledge. We find no evidence of a differential impact on productivity of treated and control stars driven by the sudden forward movement of the frontier (Table 2, Column 1). The result remains robust to considering potential changes in productivity due to changes in labor market as observed through collaboration with Soviet mathematicians (Table 2, Column 2).

Next, we consider the hypothesized role of variation in breadth of knowledge, as captured by our diversification index. We turn to our main estimation of the differential impact on productivity of specialist stars (stars with a diversification index in the bottom 5% of the diversification distribution) and generalist stars (stars with a diversification index in the top 5% of the diversification distribution). We present results in Table 3. In Columns 1 and 3 we consider a measure of productivity that excludes academic publications that with at least one Soviet collaborator. In Columns 2 and 4 we include both publications with and without Soviet collaborators. We do so to ensure that our results are not driven by the presence of Soviet mathematicians in labor markets. Columns 1 and 2 indicate a disproportionate increase in the productivity of treated relative to control specialist stars, after the collapse of the Soviet Union. In contrast, Columns 3 and 4 indicate a disproportionate decrease in the productivity of treated relative to control generalist stars, after the collapse. This main fining remains robust to different cut-offs of treated and control stars based on the risk of Soviet knowledge influence as measured by our Soviet index as well as specialist and generalist definitions based on our diversification index.

To ensure that our results are not driven by underlying trends towards increased productivity before the collapse of the Soviet Union, we examine the timing of these effects. Specifically, we run a similar regression as equation (1), however we replace the single interaction $SovietRichStar_i * AfterIronCurtain_t$ with a sequence of dummy variables representing each year before and after the collapse interacted with $SovietRichStar_i$. We present the estimation results in Figures 1 to 4. Each point represents the coefficient value of the covariate $SovietRichStar_i * Year$ and thus describes the relative difference in productivity of treated and control stars in that year. The bars surrounding each point represents the 95% confidence interval and all values are relative to the omitted base year of 1989. We include four figures, one for each equivalent estimation in each of the four columns of Table 3. It is important to note that all figures show no significant difference between the productivity of treated and control stars before the collapse of the Soviet Union. Then, starting in 1990, the difference in productivity changes towards positive values for specialist stars and negative values for non-specialist stars, in line with our findings and interpretation of Table 3 results. This trend is in line with our hypothesized differential role stars' breadth of knowledge plays in knowledge production as evidenced by the difference in slopes between the effect on the productivity of specialist and generalist stars.

Next, we extend our analysis on productivity consequences for stars' coauthors. We identify stars' collaborators as scientists who published as least once with a star in the period before the collapse of the Soviet Union. As before, we eliminate Soviet scientists from the pool of collaborators. We distinguish between collaborators of specialist stars and collaborators of generalist stars. As hypothesized, we find evidence of a disproportionate increase in productivity of specialist stars' collaborators (Table 4, Columns 1 and 2). This result is aligned with our hypothesized positive role of spillovers from specialist stars. Also in line with our hypotheses, we find evidence of no impact on the productivity of generalist stars' collaborators (Table 4, Columns 3 and 4). As before, we consider a measure of productivity that excludes academic publications with Soviet collaborators (Table 4, Columns 1 and 3) and one that include such publications (Table 4, Columns 2 and 4). We do so to ensure that our results are not driven by the presence of Soviet mathematicians in labor markets. We interpret these results as strengthening evidence of a differential role of stars driven by heterogeneity in breadth of knowledge. The positive result on the productivity of specialist stars' coauthors is aligned with the spillovers mechanism emphasized in the literature (Azoulay et al., 2012). The result on the productivity of generalist stars' coauthors provide additional robustness to this interpretation as the role of generalist stars is hypothesized to work through different mechanisms than spillovers.

We build on this interpretation by turning to collaboration as a mechanism. First, increases in returns to specialization due to forward movements of the knowledge frontier are shown to require increases in rates of collaboration (Jones 2009, Jones 2010, Agrawal et al., 2016). Since we are investigating the role of heterogeneity in returns to specialization, we expect a level of heterogeneity to reflect in rates of collaboration. Furthermore, the role of stars with broad expertise is predicated on their ability to facilitate idea recombination across knowledge domains by bringing individuals together in collaborative projects. This suggests an ex-ante heterogeneity in levels of collaboration driven by heterogeneity in breadth of expertise. Taken together, we expect collaboration rates to increase faster for treated specialist stars, relative to control specialist stars, while we expect the reverse for collaboration rates of treated non-specialist stars relative to control non-specialist stars. Following the same rationale, we expect the same effects to extend on stars' collaborators.

We focus on two measures of collaboration: extensive margin of collaboration calculated as the number of distinct collaborators per year and intensive margin of collaboration calculated as the number of co-authorship instances per year. We present results of changes in collaboration rates of stars in Tables 5a (intensive margin of collaboration) and 5b (extensive margin of collaboration). In line with our hypotheses, we find evidence of a disproportionate increase in collaboration rates of treated specialist stars relative to control specialist stars (Table 5a and 5b, Columns 1 and 2), and a decrease in collaboration rates of treated generalist stars relative to control generalist stars (Tables 5a and 5b, Columns 3 and 4), after the collapse of the Soviet Union. Furthermore, we find evidence of same effects extending to the respective stars' coauthors (Tables 6a and 6b).

6. Discussion

We distinguish between stars exhibiting variation in their breadth of expertise and present evidence consistent with a differential role in the knowledge creation process. Furthermore, we take a first step in exploring how stars' variation in breadth of expertise and, hence, brokerage influence the knowledge creation process. We present evidence consistent with an under-recognized benefit of specialization for creative work. We focus on a setting of a sudden forward movement of the knowledge frontier and find evidence on a differential impact on the productivity of stars contingent on their breadth of expertise. Furthermore, we provide evidence of the subsequent impact of these effects on the productivity of stars' coauthors. We also investigate changes in rates of collaboration as a mechanism supporting the observed differential effects on productivity rates.

Our estimations are not without limitations. First, we are confined by the research behaviour and norms of mathematics that might influence the dynamics of opportunities for knowledge creation arising. Second, we exploit a within effect variation to estimate our results. Hence the magnitude of effects is to be interpreted relative to this variation and while taking into account that the variation might, again, be field specific. Third, sudden shocks to the knowledge frontier might be different from incremental forward movements in ways that further influence the observed effects. For example, perhaps the negative effect on productivity of generalist stars can be mitigated faster under conditions of incremental movement of the frontier.

Our attempted contributions to our understanding of the knowledge creation process are three-fold. First, our results contribute to the literature that focuses on stars in science by showing that stars' common high-levels of productivity conceals important differences in their role in the knowledge production process, driven by variation in their breadth of expertise. Second, we show results consistent with an interpretation that specialization provides unique and hitherto under-recognized benefits in creative work. Third, we document an important downside to brokering ideas across fields in science. Those star scientists characterized by wider breadth of knowledge, who are theorized to broker ideas across specialties, appear disproportionately affected by a displacement of the knowledge frontier.

Our results invite further investigation of the role of breadth of knowledge in the process of knowledge creation. While we focus on a setting that increased returns to specialization, it was shown that technological advancements provide opportunities for knowledge advancements that increase returns to breadth of knowledge (Wuchty et al., 2007; Teodoridis, 2016). This suggests that future research should unpack situations that differentially benefit stars contingent on their breadth of knowledge and quantify that impact to enrich our understating of the knowledge production process.

References

Ajay Agrawal, Avi Goldfarb, Florenta Teodoridis "Understanding the Changing Structure of Scientific Inquiry" (with Ajay Agrawal and Avi Goldfarb), American Economic Journal: Applied Economics, 2016 8(1): 100-128

Ajay Agrawal, John McHale and Alexander Oettl "Why stars matter", NBER working paper #20012, 2014.

Azoulay, Pierre, Joshua Graff Zivin, and Jialan Wang, "Superstar Extinction," Quarterly Journal of Economics, 2010, 125 (2), 549–589.

Borjas, George J., and Kirk B. Doran. 2012. "The Collapse of the Soviet Union and the Productivity of American Mathematicians." Quarterly Journal of Economics 127 (3): 1143–1203.

Borjas, George J., and Kirk B. Doran. 2015. "Prizes and productivity: How Winning the Fields Medal Affects Scientific Output." Journal of Human Resources 50 (3): 728-758.

Burt, R. S. 2004. "Structural Holes and Good Ideas." American Journal of Sociology 110 (2): 349–99.

Fleming, Lee. 2001. "Recombinant Uncertainty in Technological Search." Management Science 47 (1): 117–32.

Fleming, Lee, Santiago Mingo, and David Chen. 2007. "Collaborative Brokerage, Generative Creativity, and Creative Success." Administrative Science Quarterly 52 (3): 443–75.

Franzoni, C., Scelatto, G., Stephan, P. 2011. Changing Incentives to Publish, Science, 333: 702-703.

Ganguli, Ina. 2012. "Saving Soviet Science: The Impact of Grants When Government R&D Funding Disappears." Unpublished.

Graham, Loren R. 1994. Science in Russia and the Soviet Union: A Short History. Cambridge: Cambridge University Press

Graham, Loren R., and Irina Dezhina. 2008. Science in the New Russia: Crisis, Aid, Reform. Bloomington: Indiana University Press.

Hargadon, Andrew, and Robert Sutton. 1997. "Technology Brokering and Innovation in a Product Development Firm." Administrative Science Quarterly 42 (4): 716–49. doi:10.2307/2393655.

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

Jones, Benjamin. 2010. "As Science Evolves, How Can Science Policy?" National Bureau of Economic Research (NBER) Working Paper 16002.

Lotka, Alfred J., “The Frequency Distribution of Scientific Productivity,” Journal of the Washington Academy of Science, 1926, 16, 317–325.

Mas, Alexandre and Enrico Moretti, “Peers at Work,” American Economic Review, September 2009, 99 (1), 112–45.

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

Roach, M., Sauermann, H. “A taste for science? PhD scientists’ academic orientation and self-selection into research careers in industry,” Research Policy, 2010, 39 (3), 422–434.
Sacerdote, Bruce, “Peer effects with random assignment: Results for Dartmouth roommates,” The Quarterly Journal of Economics, 2001, 116 (2), 681–704.

Stephan, P. 2012. How Economics Shapes Science, Harvard University Press. Cambridge, MA.

Teodoridis, Florenta “Understanding Team Knowledge Production: Generalists and Specialists”, working paper 2016.

Tikhomirov, Vladimir M. 2007. “On Moscow Mathematics: Then and Now.” In Golden years of moscow mathematics, 2nd ed., edited by Smilka Zdravkovska and Peter Duren, 273–83. Ann Arbor: American Mathematical Society.

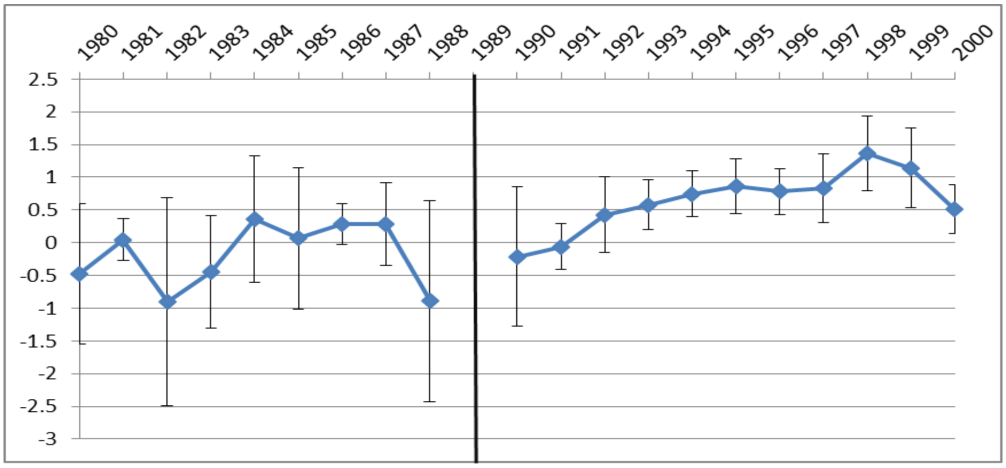
Kaplan, S, Vakili, K. 2015. “The Double-Edged Sword of Recombination in Breakthrough Innovation.” Strategic Management Journal 36 (10): 1435–57. doi:10.1002/smj.2294.

Waldinger, Fabian, “Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany,” The Review of Economic Studies, 2012, 79 (2), 838–861.

Wuchty, S., Jones, B., Uzzi, B. 2007. The Increasing Dominance of Teams in the Production of Knowledge. Science, 316: 1036-1039.

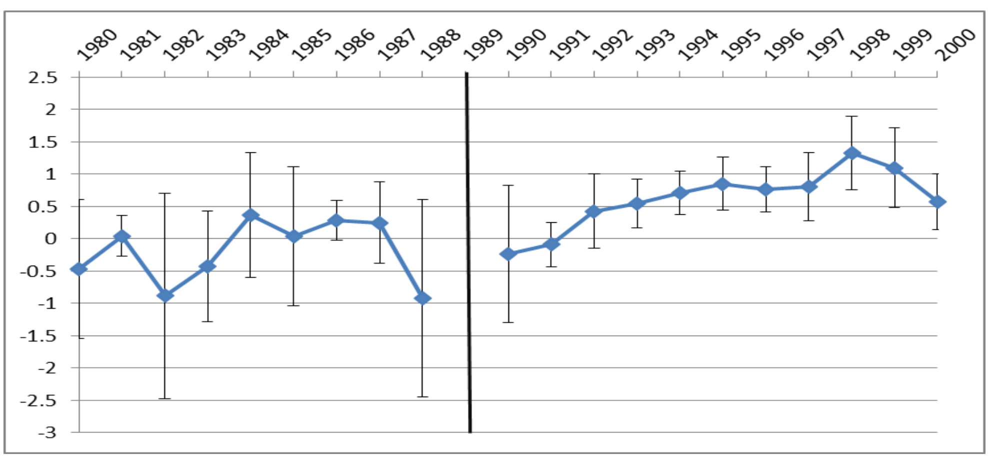
Uzzi, B., Spiro, J. 2005. “Collaboration and Creativity: The Small World Problem.” American Journal of Sociology 111 (2): 447–504.

Figure 1: Estimated difference in productivity between treated and control specialist stars (bottom 5% of the diversification index) per year; productivity calculated as count of publications without Soviet collaborators



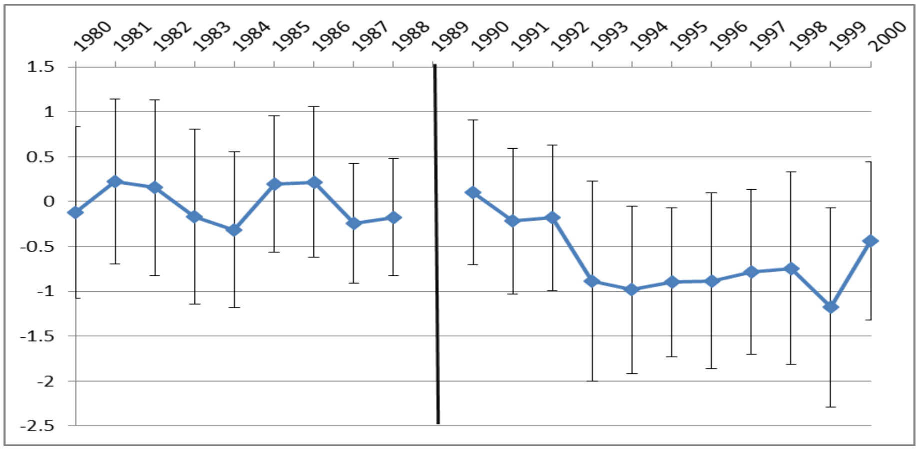
Notes: We base this figure on publications data between 1980 and 2000. Each point on the graph represents the coefficient value on the covariate SovietRichStarxYear and thus describes the relative difference in productivity between the productivity of treated and control stars in that year. The bars surrounding each point represent the 95% confidence interval. All values are relative to the base year of 1989.

Figure 2: Estimated difference in productivity between treated and control specialist stars (bottom 5% of the diversification index) per year; productivity calculated as count of publications with Soviet collaborators



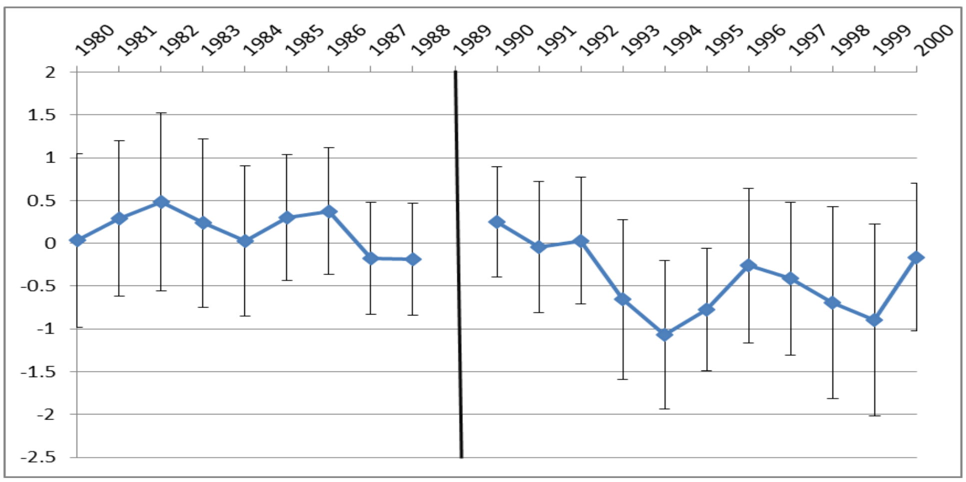
Notes: We base this figure on publications data between 1980 and 2000. Each point on the graph represents the coefficient value on the covariate SovietRichStarxYear and thus describes the relative difference in productivity between the productivity of treated and control stars in that year. The bars surrounding each point represent the 95% confidence interval. All values are relative to the base year of 1989.

Figure 3: Estimated difference in productivity between treated and control generalist stars (top 5% of the diversification index) per year; productivity calculated as count of publications without Soviet collaborators



Notes: We base this figure on publications data between 1980 and 2000. Each point on the graph represents the coefficient value on the covariate SovietRichStarxYear and thus describes the relative difference in productivity between the productivity of treated and control stars in that year. The bars surrounding each point represent the 95% confidence interval. All values are relative to the base year of 1989.

Figure 3: Estimated difference in productivity between treated and control generalist stars (top 5% of the diversification index) per year; productivity calculated as count of publications with Soviet collaborators



Notes: We base this figure on publications data between 1980 and 2000. Each point on the graph represents the coefficient value on the covariate SovietRichStarxYear and thus describes the relative difference in productivity between the productivity of treated and control stars in that year. The bars surrounding each point represent the 95% confidence interval. All values are relative to the base year of 1989.

Table 1: Subfield rank of Soviet contributions to theoretical mathematics

Subfield Rank	MSC	Theoretical mathematics category	Description
1	45	Analysis	Integral equations
2	42	Analysis	Fourier analysis
3	35	Analysis	Partial differential equations
4	40	Analysis	Sequences, series, summability
5	31	Analysis	Potential theory
6	49	Analysis	Calculus of variations and optimal control; optimization
7	44	Analysis	Integral transforms, operational calculus
8	30	Analysis	Functions of a complex variable
9	8	Algebra	General algebraic systems
10	39	Analysis	Difference equations and functional equations
11	47	Analysis	Operator theory
12	17	Algebra	Non-associative rings and non-associative algebras
13	41	Analysis	Approximations and expansions
14	58	Geometry	Global analysis, analysis on manifolds
15	32	Analysis	Several complex variables and analytic spaces
16	33	Analysis	Special functions
17	22	Algebra	Topological groups, lie groups, and analysis upon them
18	54	Geometry	General topology
19	20	Algebra	Group theory and generalizations
20	28	Algebra	Measure and integration
21	18	Algebra	Category theory; homological algebra
22	55	Analysis	Algebraic topology
23	26	Algebra	Real functions, including derivatives and integrals
24	52	Geometry	Convex geometry and discrete geometry
25	14	Algebra	Algebraic geometry
26	43	Analysis	Abstract harmonic analysis
27	15	Algebra	Linear and multilinear algebra; matrix theory
28	6	Algebra	Order theory
29	12	Algebra	Field theory and polynomials
30	5	Algebra	Combinatorics
31	51	Geometry	Geometry
32	57	Geometry	Manifolds
33	13	Algebra	Commutative rings and algebras

Notes: We follow the ranking in Agrawal, Goldfarb and Teodoridis (2016)

Table 2: Changes in productivity of non-Soviet stars after the fall of the Soviet Union, not taking into account breadth of knowledge

Dependent variable: count of academic publications per year		
	No Soviet Collaborators	With Soviet Collaborators
SovietRich x AfterIronCurtain	**-0.0806**	**-0.0522**
	(0.1174)	**(0.1126)**
Year FE	Yes	Yes
Author FE	Yes	Yes
LL	-16,618.29	-17,040.95
Observations	10,458	10,458

The data is a panel at the author level based on publication data between 1980 and 2000. All models are Poisson with robust standard errors. *significant at 10%, **significant at 5%, ***significant at 1%

Table 3: Changes in productivity of non-Soviet stars after the fall of the Soviet Union

Dependent variable: count of academic publications per year				
	(1) Specialist Stars (bottom 5% of diversification index)		(2) Generalist Stars (top 5% of diversification index)	
	No Soviet Collaborators	With Soviet Collaborators	No Soviet Collaborators	With Soviet Collaborators
SovietRich x AfterIronCurtain	**0.7689***	**0.7537***	**-0.5490****	**-0.5032***
	(0.4158)	**(0.4182)**	**(0.2688)**	**(0.2861)**
Year FE	Yes	Yes	Yes	Yes
Author FE	Yes	Yes	Yes	Yes
LL	-4,783.99	-4,864.56	-820.82	-893.43
Observations	3,045	3,045	609	609

The data is a panel at the author level based on publication data between 1980 and 2000. All models are Poisson with robust standard errors. *significant at 10%, **significant at 5%, ***significant at 1%

Table 4: Changes in productivity of non-Soviet star collaborators after the fall of the Soviet Union

Dependent variable: count of academic publications per year				
	(1) Specialist Stars Collaborators (bottom 5% of diversification index)		(2) Generalist Star Collaborators (top 5% of diversification index)	
	No Soviet Collaborators	With Soviet Collaborators	No Soviet Collaborators	With Soviet Collaborators
SovietRich x AfterIronCurtain	**1.2884*** (0.3581)**	**1.2696*** (0.3584)**	**-0.2295 (0.2396)**	**-0.3521 (0.2385)**
Year FE	Yes	Yes	Yes	Yes
Author FE	Yes	Yes	Yes	Yes
LL	-7,280.27	-7,441.85	-1,505.93	-2,578.13
Observations	9,576	9,576	1,953	1,953

The data is a panel at the author level based on publication data between 1980 and 2000. All models are Poisson with robust standard errors. *significant at 10%, **significant at 5%, ***significant at 1%

Table 5a: Changes in the intensive margin of collaboration of non-Soviet stars after the fall of the Soviet Union

Dependent variable: count of collaboration instances per year				
	(1) Specialist Stars (bottom 5% of diversification index)		(2) Generalist Stars (top 5% of diversification index)	
	No Soviet Collaborators	With Soviet Collaborators	No Soviet Collaborators	With Soviet Collaborators
SovietRich x AfterIronCurtain	**1.1584****	**1.1397****	**-0.7460****	**-0.8698****
	(0.4726)	**(0.4743)**	**(0.2956)**	**(0.3984)**
Year FE	Yes	Yes	Yes	Yes
Author FE	Yes	Yes	Yes	Yes
LL	-7,926.64	-8,136.27	-1,242.88	-1,524.74
Observations	3,045	3,045	609	609

The data is a panel at the author level based on publication data between 1980 and 2000. All models are Poisson with robust standard errors. *significant at 10%, **significant at 5%, ***significant at 1%

Table 5b: Changes in the extensive margin of collaboration of non-Soviet stars after the fall of the Soviet Union

Dependent variable: count of collaboration instances per year				
	(1) Specialist Stars (bottom 5% of diversification index)		(2) Generalist Stars (top 5% of diversification index)	
	No Soviet Collaborators	With Soviet Collaborators	No Soviet Collaborators	With Soviet Collaborators
SovietRich x AfterIronCurtain	**0.9357****	**0.9168****	**-0.6821****	**-0.7911****
	(0.4158)	**(0.4184)**	**(0.2939)**	**(0.3802)**
Year FE	Yes	Yes	Yes	Yes
Author FE	Yes	Yes	Yes	Yes
LL	-7,115.85	-7,267.40	-1,143.33	-1,368.20
Observations	3,045	3,045	609	609

The data is a panel at the author level based on publication data between 1980 and 2000. All models are Poisson with robust standard errors. *significant at 10%, **significant at 5%, ***significant at 1%

Table 6a: Changes in the intensive margin of collaboration of non-Soviet star collaborators after the fall of the Soviet Union

Dependent variable: count of collaboration instances per year				
	(1) Specialist Stars (bottom 5% of diversification index)		(2) Generalist Stars (top 5% of diversification index)	
	No Soviet Collaborators	With Soviet Collaborators	No Soviet Collaborators	With Soviet Collaborators
SovietRich x AfterIronCurtain	**0.9111****	**0.8940****	**-0.3233**	**-0.8601*****
	(0.3630)	**(0.3640)**	**(0.2571)**	**(0.2833)**
Year FE	Yes	Yes	Yes	Yes
Author FE	Yes	Yes	Yes	Yes
LL	-13,069.79	-13,631.01	-2,669.34	-3,979.43
Observations	9,576	9,576	1,953	1,953

The data is a panel at the author level based on publication data between 1980 and 2000. All models are Poisson with robust standard errors. *significant at 10%, **significant at 5%, ***significant at 1%

Table 6b: Changes in the extensive margin of collaboration of non-Soviet star collaborators after the fall of the Soviet Union

Dependent variable: count of collaboration instances per year				
	(1) Specialist Stars (bottom 5% of diversification index)		(2) Generalist Stars (top 5% of diversification index)	
	No Soviet Collaborators	With Soviet Collaborators	No Soviet Collaborators	With Soviet Collaborators
SovietRich x AfterIronCurtain	**0.8686****	**0.8503****	**-0.2729**	**-0.7693*****
	(0.3503)	**(0.3508)**	**(0.2473)**	**(0.2666)**
Year FE	Yes	Yes	Yes	Yes
Author FE	Yes	Yes	Yes	Yes
LL	-12,185.38	-12,660.83	-2,489.24	-3,695.13
Observations	9,576	9,576	1,953	1,953

The data is a panel at the author level based on publication data between 1980 and 2000. All models are Poisson with robust standard errors. *significant at 10%, **significant at 5%, ***significant at 1%