American Economic Journal: Applied Economics 2016, 8(1): 100–128
http://dx.doi.org/10.1257/app.20140135

Understanding the Changing Structure of Scientific Inquiry†

By Ajay Agrawal, Avi Goldfarb, and Florenta Teodoridis*

The fall of the Iron Curtain led to an influx of new mathematical ideas into western science. We show that research teams grew disproportionately in size in subfields of mathematics in which the Soviets were strongest. This is consistent with the knowledge burden hypothesis that an outward shift in the knowledge frontier increases the returns to collaboration. We also report additional evidence consistent with this interpretation: (*i*) *The effect is present in countries outside the United States and is not correlated with the local population of Soviet scholars*, (*ii*) *Researchers in Soviet-rich subfields disproportionately increased their level of specialization.* (*JEL* I23, O31, O33, P36)

A growing body of evidence indicates that research teams are increasing in size across almost every discipline (Adams et al. 2005; Wuchy, Jones, and Uzzi 2007; Jones 2009, 2008). Several theories explain the rise in collaboration, including the accumulation of knowledge (Jones 2009), declining communication costs (Agrawal and Goldfarb 2008; Kim, Morse, and Zingales 2009), increasing capital intensity, shifting authorship norms, and increasing returns to research portfolio diversification (Stephan 2012). However, a paucity of empirical research focuses on disentangling these different mechanisms. This is a shortcoming of the literature because these different explanations yield distinct policy implications regarding, for example, the role of poverty traps in economic development, subsidies to higher education, and the composition of research evaluation committees (Jones 2010).

In this paper, we examine changes in collaboration patterns in theoretical mathematics after the fall of the Soviet Union and find evidence most consistent with an increase in collaboration due to the accumulation of knowledge. The fall of the Iron Curtain led to a sudden and unexpected increase in knowledge of theoretical mathematics because ideas that had previously been kept secret were subsequently

*Agrawal: Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6, Canada and National Bureau of Economic Research (NBER) (e-mail: ajay.agrawal@rotman.utoronto.ca.); Goldfarb: Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6, Canada and NBER (e-mail: agoldfarb@rotman.utoronto.ca); Teodoridis: Marshall School of Business, University of Southern California, 701 Exposition Blvd., HOH 503, Los Angeles, CA 90089 (e-mail: Florenta.Teodoridis@marshall.usc.edu). All authors contributed equally to this work. This research was funded by 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 (grant numbers 486971 and 7204438). We thank Pierre Azoulay, Kirk Doran, Danielle Li, and seminar participants at the Productivity, Innovation, and Entrepreneurship meeting at the NBER Summer Institute; the Organization, Economics, and Policy of Scientic Research workshop; REER; the Workshop on Scholarly Communication and Open Science; Carnegie Mellon University; IN University; and the University of Toronto for valuable feedback.

†Go to http://dx.doi.org/10.1257/app.20140135 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

translated and disseminated. The USSR had been a world leader in various subfields of mathematics, but 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, publishing outside of the Soviet Union, and accessing foreign publications without case-by-case government approval. Thus, when the Iron Curtain fell, the knowledge frontier experienced an outward-shifting shock in mathematics outside the USSR as Soviet science suddenly became widely available.

The degree of the knowledge shock across subfields of mathematics varied. We focus on theoretical mathematics, in which there was significant variance across subfields in terms of the degree to which mathematicians in the Soviet Union were particularly advanced relative to those in the west. Employing an identification strategy pioneered by Borjas and Doran (2012) to study labor markets, we categorize as "Soviet-rich" those subfields of theoretical mathematics where Soviet mathematicians made a higher contribution relative to mathematicians from other nations before the collapse of the Soviet Union.[1] We exploit this variation across subfields using a difference-in-differences type of analysis. Specifically, we compare the propensity of mathematicians working outside the USSR to collaborate in Soviet-rich versus Soviet-poor subfields before and after the shock. We do this using 41 years of publication data in theoretical mathematics covering the period 1970–2010, 20 years before and after the collapse of the Soviet Union.

We find that team size—the number of coauthors on a paper—increased after the fall of the Iron Curtain, in both Soviet-rich and -poor subfields. However, team size increased disproportionately in Soviet-rich relative to Soviet-poor subfields after the shock. Specifically, we calculate the mean team size before and after the collapse of the Soviet Union. For the *treated* subfields (Soviet-rich), the mean team size for the 20-year period before 1990 was 1.34 compared to 1.78 for the 20-year period after. By comparison, for the control subfields (Soviet-poor), the mean team size was 1.26 before compared to 1.55 after. These differences in means suggest a disproportionate increase in team size for Soviet-rich subfields after the collapse of the Soviet Union (Figure 1). The mean team size for Soviet-rich subfields was just 6 percent higher than for Soviet-poor before 1990, but 15 percent higher after. We present the same data on a yearly basis in Figure 2.

However, there may be systematic differences between Soviet-rich and Soviet-poor subfields that are not accounted for when comparing these simple means. Therefore, we turn to our regression estimates to study the relationship further. We find evidence of an 8 percent disproportionate increase in collaboration for Soviet-rich subfields after the collapse of the Soviet Union. This result is robust to various definitions of Soviet-rich versus Soviet-poor subfields. Preexisting time trends do not drive these results: We show that the disproportionate increase in team size in Soviet-rich fields did not begin until shortly after the treatment year, 1990.

[1] Our specific identification strategy exploits the same political shock as Borjas and Doran (2012, forthcoming, 2013). In using a political shock as an instrument, we follow in a long tradition of papers that employ such shocks to understand changes in knowledge production, knowledge dissemination, and growth. Other recent papers that use this empirical strategy include Waldinger (2010, 2012), Fons-Rosen (2012); Stuen, Mobarak, and Maskus (2012); Jones and Olken (2005); and Acemoglu, Hassan, and Robinson (2011).

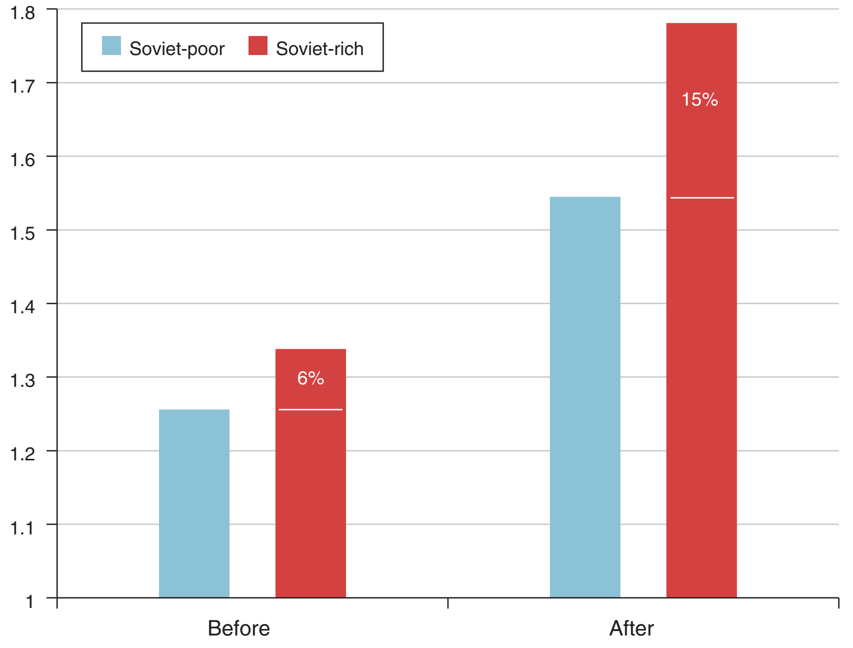

Figure 1. Disproportionate Increase in Mean Team Size in Soviet-Rich Subfields after 1990

Notes: We base this figure on 40 years of publication data for the 3 top and 3 bottom ranked subfields of theoretical mathematics. We plot the average team size for Soviet-rich and Soviet-poor subfields, before and after 1990.

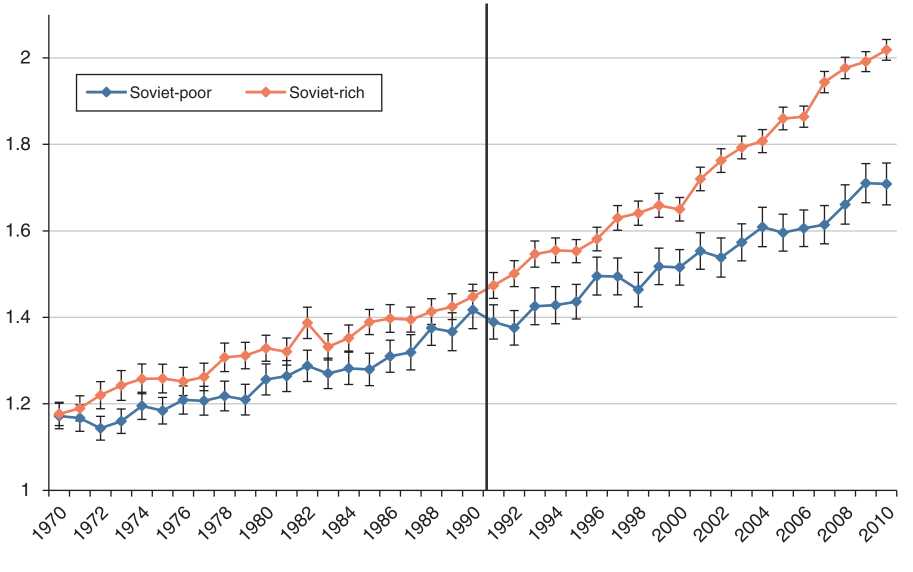

Figure 2. Disproportionate Increase in Yearly Mean Team Size in Soviet-Rich Subfields after 1990

Notes: We base this figure on 40 years of publication data for the 3 top and 3 bottom ranked subfields of theoretical mathematics. We plot the yearly average team size for Soviet-rich and Soviet-poor subfields, before and after 1990.

Although we focus on mathematics, other studies document that the size of research teams has increased across fields (Adams et al. 2005; Wuchy, Jones, and Uzzi 2007; Jones 2009). For example, Wuchy, Jones, and Uzzi (2007) show that over the latter half of the twentieth century, team size increased in 170 out of 171 fields in science and engineering, 54 out of 54 fields in the social sciences, and 24 of 27 fields in the arts and humanities. Furthermore, they present citation-based evidence that the relative impact of team versus individual output is increasing over time, even after controlling for self-citations. Moreover, these authors report that despite the common characterization of mathematics as a less collaborative field, team size has increased significantly: "Surprisingly, even mathematics, long thought the domain of the loner scientist and least dependent of the hard sciences on lab scale and capital-intensive equipment, showed a marked increase in the fraction of work done in teams, from 19 percent to 57 percent, with mean team size rising from 1.22 to 1.84."

Scholars have advanced a number of hypotheses to explain this trend. Hesse et al. (1993) and others emphasize the role of reduced communication costs due to advances in communication technology (Agrawal and Goldfarb 2008; Kim, Morse, and Zingales 2009) or reductions in the cost of travel. Stephan (2012) discusses several more alternatives. For example, increasing capital intensity in many fields, such as the role of particle accelerators in physics, may increase the returns to collaboration due to the indivisibilities of research equipment. Changing norms may mean that contributors, who in the past may have been listed in the acknowledgements, are increasingly likely to be included as coauthors, especially in lab-based sciences. Academics also may find increasing returns to mitigating publication risk by diversifying their research portfolios as publication requirements for promotion and tenure rise. Jones (2009) emphasizes the "knowledge burden" hypothesis in which successive generations of innovators face an increasing education burden due to the advancing knowledge frontier. This advancing frontier, he posits, requires innovators to specialize more and thus necessitates working more collaboratively, which alters the organization of innovative activity towards teamwork. Jones provides descriptive statistics consistent with this theory. For example, he shows that over time: (i) the number of co-authors on academic publications increases, (ii) Nobel laureates are older when they perform their great achievement, (iii) the number of co-inventors per patent increases, (iv) the age at first innovation increases, and (v) the probability of switching fields decreases. However, these statistics are also consistent with some of the other explanations.

Because different theories have different implications for policy, it is useful to report empirical evidence that identifies one particular causal effect. For example, Jones (2008) presents a model in which the knowledge burden leads to a poverty trap. Jones argues that as the knowledge frontier shifts outward, individuals compensate by specializing, and thus the returns to collaborating increase. However, in economies where the market for complementary skills is thin, individuals are less likely to invest in the human capital necessary to reach the frontier. This results in an increasingly thin market for specialized skills and thus further lowers the returns to human capital acquisition (creating a trap). Therefore, one policy prescription is to subsidize skills development in a concentrated area (e.g., infectious diseases) in order to address the complementary skills shortage for a finite period of time until the private returns to acquiring specialized skills are sufficient for the labor market to sustain the cycle without further intervention. This policy initiative is not appropriate in the absence of this causal effect if knowledge accumulation does not increase the returns to collaboration and the observed rise in team size is driven by other factors in the economy, such as rising capital costs and/or falling communication costs.

Our main result based on the difference-in-differences estimation cannot be explained by most of the aforementioned theories, other than knowledge burden. Most of the other explanations, such as changing authorship norms, falling communication costs, and increasing capital costs associated with research equipment and other indivisibilities, occurred gradually over time and would not produce a sharp increase in team size immediately after 1990 with no pre-trend. Furthermore, almost all of the alternate explanations would predict any increases in team size to apply similarly across all subfields of mathematics, not disproportionately in the particular subfields that we classify as Soviet-rich. To be clear, we do not rule out the possibility that the aforementioned mechanisms affect team size. It is possible that they all do. After all, we observe increases in team size across all subfields of math both before and after 1990. However, we isolate the effect of a specific mechanism by focusing on the disproportionate increase in team size of Soviet-rich subfields that occurred at a specific point in time when Soviet knowledge was suddenly made available to the worldwide community of scholars.

There is one alternative explanation that we believe warrants further explanation and that we are unable to completely reject: a shock to the labor market that resulted from Soviet mathematicians suddenly being able to collaborate and travel at precisely the same time as the shock to the knowledge frontier. Borjas and Doran (2012, forthcoming, 2013) emphasize the labor market impact of increased competition from Soviet scholars. They report that the influx of Soviet mathematicians to American and European universities resulted in increased competition for jobs and journal slots, and so it is plausible that this may have caused an increase in collaboration, particularly in Soviet-rich subfields.

We believe our results are more likely driven by an outward shift in the knowledge frontier than changes in the labor market for four reasons. First, our main result is robust to dropping all papers with Soviet-named authors. Second, and perhaps most importantly, we show the same pattern persists in Japan, a country that experienced very little immigration of Soviet mathematicians. More generally, we provide additional evidence from a set of 15 countries and report that the estimated Soviet-rich effect is not strongly correlated with the number or fraction of local Soviet mathematicians. Third, we show that papers in Soviet-rich subfields disproportionately cited Soviet papers after the fall of the Soviet Union, suggesting that non-Soviet scholars did indeed draw upon the insights of Soviet mathematicians. Fourth, we report evidence of a disproportionate increase in specialization in Soviet-rich fields, consistent with the mechanism described in Jones (2009).

We structure the remainder of the paper as follows. In Section I, we provide historical context for our instrument, explaining how knowledge was developed in the Soviet Union and yet kept secret from Western mathematicians, creating the conditions for the 1990 shock to the frontier. In Section II, we describe our differences-in-differences empirical strategy by comparing the propensity to collaborate in Soviet-rich versus Soviet-poor subfields before and after the knowledge shock. In Section III, we describe the mathematics publication data we use to construct our sample as well as the method we employ for classifying subfields as Soviet-rich or Soviet-poor. We present our results in Section IV and our conclusions in Section V.

I. Historical Context

The interpretation of our results as evidence for the knowledge burden hypothesis relies on the assertion that the collapse of the Soviet Union around 1990 caused an outward shift in the knowledge frontier in mathematics and that it did so more for some subfields than others. We rely on three observations to substantiate this assertion: (i) the Soviet Union's effect on the knowledge frontier in mathematics was significant, (ii) the Soviet Union's effect on the knowledge frontier was greater in some subfields than others, and (iii) the knowledge produced in the Soviet Union was kept secret from the outside world such that its effect on the frontier came reasonably suddenly in the years following 1990. We offer historical context for each of these three points below.

First, the Soviet Union's contribution to knowledge in the field of mathematics was meaningful and significant. The Soviet Union was and Russia continues to be a world-renowned center of scientific research, with mathematics holding a prominent position. Lauren Graham, a historian of Soviet science and technology, states: "Of all fields of knowledge, it was mathematics to which Russia and the Soviet Union made the greatest contributions. The Soviet Union became a world power in mathematics" (Graham and Dezhina 2008). Graham attributes the Soviet Union's strength in scholarly research in mathematics to the fact that it attracted great minds; mathematics was uniquely detached from politics, conferred status and prestige, and offered financial rewards superior to many other occupations.

Second, the Soviet Union's contribution to knowledge was significantly greater in some subfields of mathematics than others. Borjas and Doran (2012) show this empirically by comparing across subfields the fraction of Soviet-to-American papers published during the period 1984–1989. Graham (1994) notes that although Soviet mathematicians were strong across the entire spectrum of theoretical and applied mathematics, they seemed to have made the greatest advancements, relative to the rest of the world, in pure theory. One explanation for this is politics. Soviet policies were strict about secrecy and focused on maintaining control over technological developments. It was easier for Soviet mathematicians to build on their progress in theoretical mathematics than in areas where technology implementation was more immediate. Many advances in applied mathematics were stalled for political reasons, with exceptions linked to government interests such as the space program (Graham 1994). Differences in subfields were further amplified due to path dependency: Subfields that attracted bright minds early on were more likely to subsequently attract more bright minds due to mentorship opportunities (Borjas and Doran 2012). The importance of mentorship is well known in science (Merton 1973) and was likely particularly salient in this setting due to restrictions on travel and access to foreign journals. For example, the success of Moscow mathematics can be traced back to Ergorov and his student N. N. Luzin (Tikhomirov 2007). Luzin, whose famous work was mainly focused on the theory of functions, a subfield of theoretical mathematics, mentored subsequent generations of eminent Soviet mathematicians. On the other hand, little outstanding mentorship was available to practitioners in some other subfields of theoretical mathematics, like algebraic geometry (Borjas and Doran 2012).

Third, the knowledge produced in the Soviet Union was kept secret from the outside world such that its effect on the frontier came reasonably suddenly in the years following 1990. Soviet researchers were prevented from publishing their findings, traveling to conferences, communicating or collaborating with non-Soviets, and even accessing non-Soviet references. 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). Furthermore, Soviet advancements in mathematics remained relatively unknown in the United States until the collapse of the Soviet Union mainly because the USSR government kept much of Soviet science secret (Graham and Dezhina 2008). In addition, what escaped the secrecy filter was subject to the natural barrier imposed by the Russian language. Graham and Dezhina (2008) note: "the Russian language was known by few researchers outside the Soviet Union, and consequently the achievements of Soviet researchers were more frequently overlooked than those presented in more accessible languages." Borjas and Doran (2012) provide extensive evidence that Soviet knowledge in mathematics was not known in the West, although translations of some Soviet scientific journals were available before the collapse of the Soviet Union.

The limited diffusion of Soviet mathematics into the West is evident in the aftermath of the collapse of the Soviet Union. Starting in 1990, Soviet discoveries began to spread through the West and were considered new and important. Communication and travel restrictions were lifted, publications were translated and indexed, and ideas and knowledge began to flow out from the Soviet Union into the broader research community. The following quote by Harvard mathematician Persi Diaconis (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:

> *"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 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.*

To be clear, for our interpretation of the results to apply, we do not require there to have been no information leaking out of the Soviet Union prior to 1990. Instead, we require the knowledge available to non-Soviet researchers to have increased after 1990 in Soviet-rich relative to Soviet-poor fields.

In sum, the fall of the Iron Curtain provides a plausible natural experiment differentially affecting the knowledge frontier across subfields of theoretical mathematics. This historical event was exogenous to the mathematics research community and set free a large pool of accumulated knowledge. Furthermore, Borjas and Doran (2012), who pioneered the use of this event as an instrument for causal identification in the setting of mathematics, present comprehensive evidence indicating that the timing of the collapse took the global mathematics community by surprise; even in the late 1980s, both the Western mathematical community and Soviet scholars were quite certain that Soviet mathematics would remain secluded for the foreseeable future.

II. Estimation Strategy

We employ a difference-in-differences estimation strategy in which we compare collaboration rates in subfields where the Soviets were particularly strong (*treated*) with other subfields (*control*), both before and after the fall of the Iron Curtain (1990). In other words, we examine the difference between treated and control subfields in two periods, before and after the treatment. Thus, we distinguish between the rise in team size that is directly attributable to changes specific to Soviet-rich fields from the underlying differences between treated and control subfields as well as the underlying changes in collaboration patterns in theoretical mathematics over time.

The objective of our empirical analysis is to estimate the effect of the collapse of the Soviet Union on collaboration in Soviet-rich fields relative to other fields. We measure collaboration as a count of the number of unique authors on a publication. Thus, we estimate the following linear regression model, using the academic paper as our unit of analysis:

$$(1) \quad TeamSize_{it} = \beta(SovietRich_i \times AfterIronCurtain_t) + \mathbf{Subfield}_i + \gamma_t + \epsilon_{it}.$$

$TeamSize_{it}$ is the logged count of authors for each academic paper i published in year t. $SovietRich_i$ is an indicator variable equal to 1 if academic paper i belongs to the treated group, and 0 otherwise. $AfterIronCurtain_t$ is an indicator variable equal to 1 if academic paper i is published after 1990, and 0 otherwise. This applies to academic papers in both treated and control subfields. We include subfield and time fixed effects. Hence, the main effects $SovietRich_i$ and $AfterIronCurtain_t$ drop out of the estimating equation.

We are primarily interested in the estimated coefficient on the interaction between $SovietRich_i$ and $AfterIronCurtain_t$, which equals 1 for publications in treated subfields that were published after the knowledge shock and 0 for all others. We interpret a positive estimated value of this coefficient as implying that the average team size of Soviet-rich subfields increased disproportionately, relative to Soviet-poor subfields, after the collapse of the Soviet Union. After establishing this relationship, we provide evidence that we interpret as being consistent with a mechanism driven by an outward shift in the knowledge frontier.

III. Data

We follow three main steps to collect and prepare our data. First, we extract publication data, then we rank subfields in mathematics with respect to the relative contribution by Soviets, and finally we process the data for analysis.

A. *Data Collection*

We collect data on every publication in theoretical mathematics published during the 41-year period 1970–2010. This represents 20 years of data both before and after the collapse of the Soviet Union in 1990. We follow Borjas and Doran's (2012) interpretation of historical events that identifies 1990 as the year when academic seclusion was significantly lessened. We recognize that the political and social turmoil preceding and following the fall of the Iron Curtain spanned a period of roughly three years, between 1989 and 1991. Our results are robust to choosing 1989 or 1991 as the cutoff rather than 1990.

We collect these data from the American Mathematical Society (AMS). The Mathematical Reviews (MR) division of AMS maintains a comprehensive bibliographic database of worldwide academic publications in mathematics. The MR database includes all mathematics-related journal publications 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 (Appendix Figure 1).

B. *Classification*

Our empirical strategy relies on exploiting variation across subfields in the degree to which the knowledge frontier shifted outward. Specifically, we distinguish between subfields of theoretical mathematics where the Soviets were particularly strong in the years prior to the collapse versus subfields where they were less strong. Borjas and Doran (2012) created this identification tool; we directly employ their insight on how to classify these data to distinguish between Soviet-rich and Soviet-poor.

We rely on the careful and exhaustive work of the MR division, which classifies each paper in mathematics using Mathematics Subject Classification (MSC) codes. The MSC schema is used internationally 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. The theoretical mathematics group is comprised of a total of 40 active primary MSC codes (14 algebra, 19 analysis, 7 geometry). We drop the 6 subfields that did not exist throughout the 40-year duration of our study period as well as 1 subfield for which we are not able to obtain the full data, leaving us with 33 subfields within theoretical mathematics. Next, we adopt the Borjas and Doran (2012) ranking of the remaining 33 subfields, which is based on the degree to which Soviets contributed to a particular subfield. They construct their rank by calculating the ratio of Soviet-to-American publications in the subfield over the period 1984–1989 and define a publication as Soviet if at least one author has a Soviet institutional affiliation. They similarly define American publications. We list the 33 subfields and their rank in Table 1.

Although we use the Borjas and Doran (2012) ranking throughout the paper, we also show in Appendix Table 1 that the main results are robust to an alternate measure. While broadly similar, this alternative measure differs on three dimensions. First, this measure defines a publication as Soviet based on author name data rather than author affiliation data.[2] Second, this measure compares Soviet publication output relative to the rest of the world rather than relative to US-only publication output. Third, this measure uses ratios based on data from 1970 to 1989 rather than from 1984 to 1989. In the end, the rankings are reasonably similar, with a Spearman Rank Correlation Coefficient of 0.84, and the qualitative results are unchanged.

C. *Data Processing*

In our main specifications, we drop all Soviet publications from the sample. We define Soviet publications as those with at least one Soviet author. We do this to avoid potential confounding effects. After 1990, not only was Soviet knowledge set free to contribute to global advancements in mathematics, but collaboration restrictions were also lifted for Soviet mathematicians. By excluding publications with at least one Soviet author, we partially account for the possibility of increased co-authorship rates due to removing the constraint previously preventing collaboration with Soviets.[3]

After dropping Soviet publications, our sample includes 563,462 publications spanning the 41-year period. We focus on a comparison between the 3 top (Soviet-rich) and bottom (Soviet-poor) ranked subfields, which represents 133,497 publications, as our main specification; however, we show the results are robust to alternative definitions of Soviet-rich: (i) top 3 ranking subfields relative to all others, (ii) top 5 ranking subfields relative to all others, (iii) top 10 ranking subfields relative to all others, and (iv) a continuous measure that relies on variation within the 33 ranked subfields of theoretical mathematics. We provide descriptive statistics in Table 2.

IV. Results

A. *Main Result: Disproportionate Team Size Increase in Soviet-Rich Subfields after* 1990

We report the estimated coefficients of equation (1) in Table 3. We present our main specification in column 1. The key result is the estimated coefficient on the

[2] We identify Soviet last names based on conversations with experts and documented rules regarding Soviet surname endings. We then test and calibrate our algorithm by manually looking up and verify if academics identified as having Soviet last names were indeed Soviets.

[3] Our results remain robust when adding Soviet authors back to the sample (Appendix Table 2).

Table 1—Subfield Rank Based on Proportion of Soviet Publications (1984–1989)

Subfield rank as per Borjas and Doran (2012)	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	Geometry	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 adapt this ranking from Borjas and Doran (2012) and base it on the ratio of the number of Soviet versus American papers published in the particular subfield between 1984 and 1989. We define papers as Soviet if at least one author has a Soviet institutional affiliation. We similarly define American papers. Thus, the most Soviet-rich subfield is integral equations, which is the field with the highest fraction of papers with Soviet-based authors. The most Soviet-poor subfield is commutative rings.

interaction term $SovietRich_i \times AfterIronCurtain_t$, which is positive and statistically significant. This result implies that the difference in average team size between papers in Soviet-rich versus Soviet-poor subfields is greater after the shock than before.

We do not present estimates of the main effects of $SovietRich_i$ or $AfterIronCurtain_t$ because we drop these terms from the estimating equation due to the year and subfield fixed effects. Also, we cluster our standard errors by subfield. We cluster to address the possibility that shocks experienced in each of the 33 subfields may be correlated, both within subfield and over time (Bertrand, Duflo, and Mullainathan 2004; Donald and Lang 2007).

This main result is robust to various definitions of Soviet-rich. In the first column, we define Soviet-rich and Soviet-poor as the three top and three bottom subfields, respectively. In the next three columns, we define Soviet-rich as the top three, five,

Table 2—Descriptive Statistics

	Mean	Standard deviation	Minimum	Maximum	Number of observations
All observations					
log of author count	0.342	0.425	0	2.89	563,462
AfterIronCurtain	0.625	0.484	0	1	563,462
SovietRich (top 3)	0.169	0.374	0	1	563,462
AfterIronCurtain × SovietRich (top 3)	0.121	0.326	0	1	563,462
SovietRich (top 5)	0.183	0.386	0	1	563,462
AfterIronCurtain × SovietRich (top 5)	0.128	0.334	0	1	563,462
SovietRich (top 10)	0.287	0.452	0	1	563,462
AfterIronCurtain × SovietRich (top 10)	0.194	0.395	0	1	563,462
Top and bottom three fields only					
log of author count	0.362	0.428	0	2.19	133,467
AfterIronCurtain	0.678	0.467	0	1	133,467
SovietRich	0.711	0.453	0	1	133,467
AfterIronCurtain × SovietRich	0.511	0.500	0	1	133,467
Count of Soviet references by name	0.811	1.528	0	12	1,012
Percent of Soviet references by name	0.046	0.084	0	0.6	1,012
Count of Soviet references by journal	0.295	0.881	0	12	1,012
Percent of Soviet references by journal	0.015	0.047	0	0.5	1,012
Count of subfields (all authors)	2.039	1.705	1	47	315,161
log of count of subfields (all authors)	0.490	0.613	0	3.850	315,161
Count of subfields (junior authors)	1.931	1.471	1	26	41,282
log of count of subfields (junior authors)	0.463	0.580	0	3.258	41,282

Table 3—Teams in Soviet-Rich Subfields Exhibit a Disproportionate Increase in Size after 1990

	Dependent variable: log of author count				
	Top and bottom three subfields	Soviet-rich defined as top three of subfields	Soviet-rich defined as top five of subfields	Soviet-rich defined as top ten of subfields	Continuous
AfterIronCurtain × SovietRich	0.0780***	0.0489***	0.0394***	0.0389**	0.0011
	(0.0117)	(0.0106)	(0.0138)	(0.0152)	(0.0011)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.113	0.106	0.106	0.106	0.106
Observations	133,497	563,462	563,462	563,462	563,462

Notes: The unit of analysis is the publication. All models are OLS with robust standard errors, clustered by subfield. The results remain robust to Poisson estimation (Appendix Table 3).

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

and ten subfields, respectively, and Soviet-poor as all remaining subfields. The point estimates decrease in magnitude as we broaden the definition of Soviet-rich, which is not surprising since the difference between Soviet-rich and Soviet-poor is less stark, but the point estimates remain positive and significant. In the last column, we employ a continuous rank measure of Soviet-rich/poor, assigning the most Soviet-rich field a rank of 33 and the least a rank of 1. The coefficient remains positive, though significance is lost in the continuous specification.

Next, we examine the timing of this effect to demonstrate that there is no underlying trend toward increased collaboration in Soviet-rich relative to Soviet-poor fields

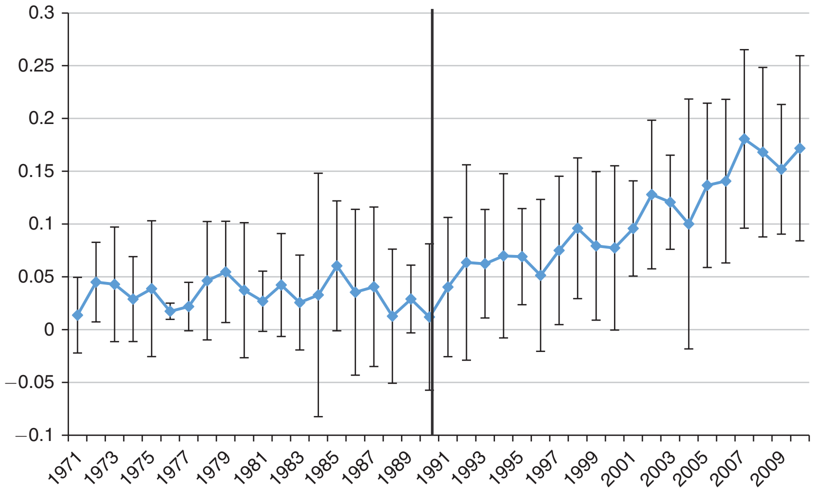

Figure 3. Plot of Estimated Coefficients on Interaction between Soviet-Rich and Year (DV = Team size)

Notes: We base this figure on 40 years of publication data for the 3 top and 3 bottom ranked subfields of theoretical mathematics. Each point on the graph represents the coefficient value on the covariate *Year* × *SovietRich* and thus describes the relative difference in collaboration rates between Soviet-rich and Soviet-poor fields in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 1970. We also present these results in table form in Appendix Table 4.

before 1990. For example, perhaps the network of scholars in Soviet-rich fields was better positioned to leverage the diffusion of electronic communication technology that led to increased scientific collaboration starting in the 1980s (Agrawal and Goldfarb 2008). In this case, one might worry that the effect of lowered collaboration costs, although spread out over many years and during a slightly earlier period than the 1989–1991 events in the Soviet Union, could explain the result.

To check for such a possibility, we examine the timing of the relationship between the collapse of the Soviet Union and changes in the relative team size in Soviet-rich versus Soviet-poor subfields. Specifically, we run a similar regression to the one shown in Table 3, column 1; however, we replace the single interaction $SovietRich_i \times AfterIronCurtain_t$ with a sequence of dummy variables representing each year interacted with *SovietRich*.

We present the results in Figure 3. Each point represents the coefficient value on the covariate *Year* × *SovietRich* and thus describes the relative difference in collaboration rates between Soviet-rich and Soviet-poor fields in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 1970. We also present these results in table form in Appendix Table 4. The most notable insight from Figure 3 is that the difference between team sizes in Soviet-rich and Soviet-poor fields was stable between 1971 and 1990. Then, starting in 1990, the difference in average team size began to increase, as evidenced by the higher coefficients. The difference in team size becomes statistically significant after about 8 years and then continues to increase for the 12 remaining years in the sample.

B. *Knowledge Burden versus Competition in Labor Markets*

The results we present in the prior section are intriguing because they are consistent with the knowledge burden explanation but not with many of the other explanations for the observed rise in team size. First, most of the other explanations, such as changing authorship norms, falling communication costs, and increasing capital costs associated with research equipment and other indivisibilities, occurred gradually over time and would not produce a sharp increase in team size immediately after 1990 with no pre-trend. Instead, we expect these mechanisms to result in increasing team size throughout the 1980s. Indeed, we do observe an increase in team size during the 1980s. However, this occurs across all subfields. Yet we do not observe a disproportionate increase in the size of teams in Soviet-rich relative to Soviet-poor subfields until 1990. This underscores the second point. Almost all of the alternate explanations would predict any increases in team size to apply similarly across all subfields of mathematics, not so disproportionately in the particular subfields that we classify as Soviet-rich.

To reiterate, we do not rule out the possibility that any of the aforementioned mechanisms affect team size. As described above, it is possible that they all do. Of the various mechanisms, there is one alternative that we believe deserves further examination as to whether it drives the main results: a shock to the labor market that results from Soviet mathematicians being suddenly able to collaborate and travel at precisely the same time as the shock to the knowledge frontier. Borjas and Doran (2012, 2013a, 2013b) emphasize the labor market impact of increased competition from Soviet scholars. They report that the influx of Soviet mathematicians to American and European universities resulted in increased competition for jobs and journal slots, and so it is plausible that this may have caused an increase in collaboration, particularly in Soviet-rich subfields.

Thus, we take three additional steps to seek further evidence of whether the knowledge burden is at least partially responsible for the rise in team size. First, recognizing that Soviet migration varied across countries, we examine whether the collaboration effect varied by immigration level. Second, we examine whether individuals in Soviet-rich subfields increased their level of specialization as the knowledge burden thesis predicts. Finally, we search for direct evidence of a shock to knowledge flows by examining the pattern of citations to Soviet papers.

We thus turn to the second step, where we look across countries to examine whether the increase in collaboration in Soviet-rich fields relates to the degree of labor market competition by country. We focus first on Japan, a country with no documented evidence of Soviet immigration in mathematics (Dubois, Rochet, and Schlenker 2014) and which consistently ranks in the top ten mathematics research (Dubois, Rochet, and Schlenker 2014). We show a disproportionate team size increase in Japanese journal publications in Soviet-rich relative to Soviet-poor subfields after 1990. We present the raw yearly change in mean team size for both Soviet-rich and Soviet-poor subfields in Figure 4. In Table 4, we use our main difference-in-differences specification (1) to estimate the effect of the knowledge shock on team size in Soviet-rich versus Soviet-poor fields in the Japanese setting by restricting the data to the subset of publications from Japanese journals. The

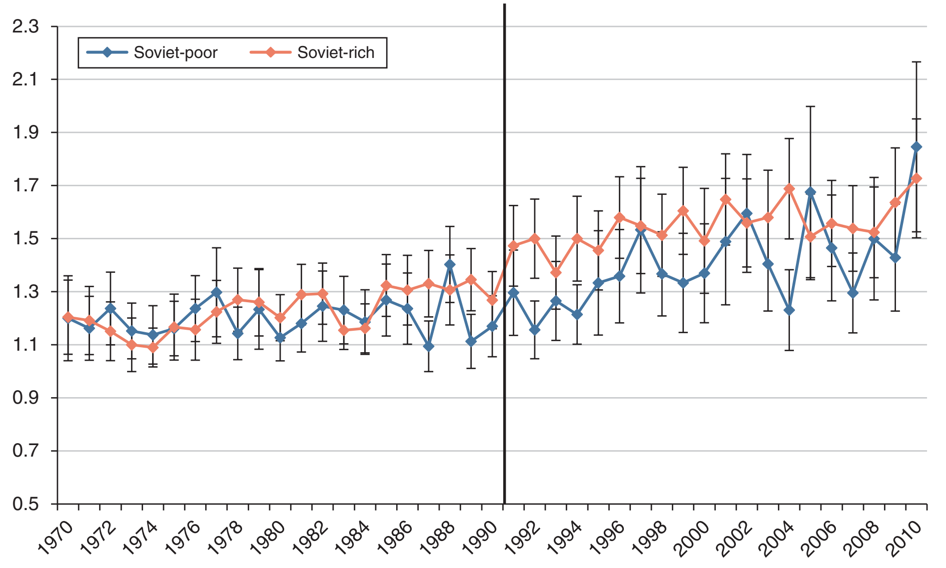

Figure 4. Disproportionate Increase in Yearly Mean Team Size on Japanese Publications in Soviet-Rich Subfields after 1990

Notes: We base this figure on 40 years of publication data in Japanese journals for the 3 top and 3 bottom ranked subfields of theoretical mathematics. We plot the yearly average team size for Soviet-rich and Soviet-poor subfields, before and after 1990.

Table 4—Teams in Soviet-Rich Subfields Exhibit a Disproportionate Increase in Team Size after 1990 (*Top and bottom three subfields from Japanese journals*)

	Dependent variable: log of author count		
	All Japanese journals	Ranked Japanese journals	Unranked Japanese journals
AfterIronCurtain × SovietRich	0.0692***	0.0678***	0.0956***
	(0.0173)	(0.0203)	(0.0206)
Year fixed effects	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes
R^2	0.076	0.093	0.101
Observations	5,096	3,859	1,237

Notes: The unit of analysis is the publication. All models are OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

estimated coefficient in column 1 indicates a clear increase in collaboration after 1990 in Soviet-rich subfields.

We recognize that the labor market for mathematicians is to some extent global. For example, scholars based outside of Japan may be interested in publishing in high-profile Japanese mathematics journals. As such, we distinguish between those Japanese journals that are included in Thomson Reuters' impact factor ranking and those that are not. We assume that Japanese journals with impact factors ranked by

Thomson Reuters are significantly more likely to attract interest from international scholars than unranked journals and, thus, be more susceptible to influence from shocks to global competition for journal space. We interpret the disproportionate team size increase in Japanese journal publications, both ranked (column 2) and unranked (column 3), in Soviet-rich relative to Soviet-poor subfields after 1990 as more consistent with the knowledge burden hypothesis than the labor market explanation that is predicated on the influx of Soviet scholars in some countries, such as the United States, but negligible in Japan.

In Figure 5, we account for the possibility that an underlying trend toward increased collaboration drives these results by showing the regression coefficients of a modified specification that replaces the single interaction $SovietRich_i \times AfterIronCurtain_t$ with a sequence of dummy variables representing each year interacted with *SovietRich*. While noisy due to the relatively small sample in just one country, we see no clear pre-trend, and the average values post-1990 are higher than the values pre-1990.

Next, we extend the analysis to journals published by country in all countries in our data for which we had at least 500 observations. In Figure 6, we present the estimated coefficients of our main specification across these 15 countries. The vertical axis shows the estimated coefficient by country of collaboration rates on the interaction of Soviet-rich and after 1990. The countries are ordered by the fraction of scholars with Soviet names who published in the journals.[4] The results suggest that there is no systematic relationship between the estimated increase in collaboration rates and competition from Soviet scholars in the journals. The correlation coefficient between the coefficient on the increase in collaboration rates and the fraction of papers published by Soviet authors in that country's journals is -0.0369.

Broadly, we think this lack of evidence of differences across countries suggests that the collapse of the Soviet Union had a similar impact across countries, regardless of the degree of direct labor market competition from Soviet scholars. We interpret this as more consistent with the knowledge burden hypothesis than with increasing competition under the assumption that, while competition might affect all countries, it is likely to have a larger effect in countries with higher levels of local Soviet competition.

C. *Researchers in Soviet-Rich Subfields Disproportionately Increase Citations to Soviet Prior Art*

Next, we combine results from Borjas and Doran (2012) with evidence from our data to document that the collapse of the Soviet Union does seem to have generated a knowledge shock in Soviet-rich fields. This is a necessary (but insufficient) condition for our interpretation that the increase in collaboration in Soviet-rich

[4] This is a measure of competition from Soviets for journal space rather than a direct measure of labor market competition. Unfortunately, we do not have information across many countries on the rate of immigration of Soviet mathematicians (except broad statements such as the rate was low in Japan and high in Germany). The coefficient is positive in most countries (except Spain, Belgium, and Singapore), though only significant in those countries with relatively precise estimates (Japan, the United States, and Germany). Still, we view this as a good proxy for labor market competition.

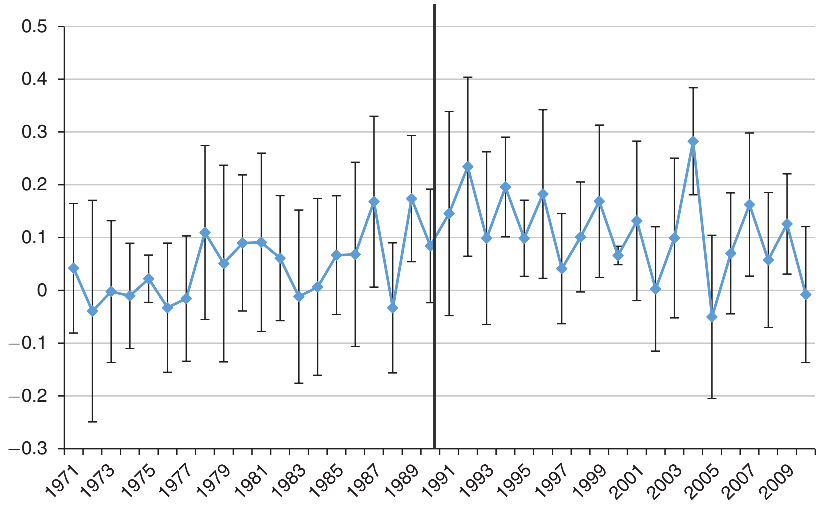

Figure 5. Plot of Estimated Coefficients on Interaction between Soviet-Rich and Year (DV = Team size) for Japanese Journals

Notes: We base this figure on 40 years of publication data in Japanese journals for the 3 top and 3 bottom ranked subfields of theoretical mathematics. Each point on the graph represents the coefficient value on the covariate *Year* × *SovietRich* and thus describes the relative difference in collaboration rates between Soviet-rich and Soviet-poor fields in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 1970. We also present these results in table form in Appendix Table 5.

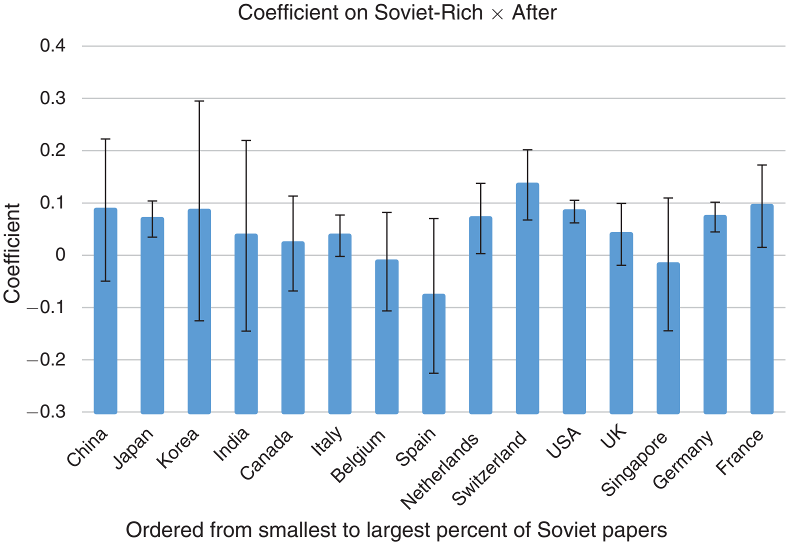

Figure 6. Plot of Estimated Coefficients on Interaction between Soviet-Rich and Year (DV = Team size) for Various Countries (*Identified by journal affiliation*) Ordered by Estimated Soviet Competition (*Identified by the fraction of publications with Soviet authors relative to the total number of publications per country*)

Notes: We base this figure on 40 years of publication data in select country journals for the 3 top and 3 bottom ranked subfields of theoretical mathematics. Each column represents the estimated coefficient of our main difference-in-differences specification (1). The bars surrounding each point represent the 95 percent confidence interval. We also present these results in table form in Appendix Tables 6 and 7.

subfields after the collapse of the Soviet Union was driven by an increased burden of knowledge.

Borjas and Doran (2012, 1156, figure IV) show a sharp increase in citations to Soviet papers by American mathematicians. We verify this trend in our data. In particular, we use backward citation data for papers from the top 3 and bottom 3 subfields that were published in one of the top 30 journals of mathematics (as measured by Thomson Reuters' impact factor). We further restrict the data to a window of four years before and after the collapse of the Soviet Union (1988–1993) for tractability (this data collection process is manual). We extract 1,217 publications that meet these criteria and are authored by non-Soviet scholars. We find full text information on 1,012 of these papers (using Thompson Reuters' Web of Knowledge database) for which we extract the list of references. We count references to Soviet pre-1990 publications and calculate the percentage of Soviet references relative to the total number of references. We define a citation (to a pre-1990 publication) as Soviet if it was published in a Soviet journal. We use these data to estimate a difference-in-differences linear regression, similar to the one estimated in Section IVA:

$$(2) \quad SovietArt_{it} = \beta(SovietRich_i \times AfterIronCurtain_t) + \mathbf{Subfield}_i + \gamma_t + \epsilon_{it}.$$

We report our estimated coefficients in Table 5. The results show that citations to pre-1990 papers increased disproportionately in Soviet-rich fields after 1990 (measured by count of citations in the first column and fraction of citations in the second column). Furthermore, this result is stronger for papers written by larger teams (third and fourth columns). Thus, the combined evidence of Borjas and Doran (2012, figure IV) and our Table 5 suggests that Soviet-rich subfields appear to have experienced a knowledge shock after the fall of the Soviet Union.

D. *Evidence of Specialization*

Next, we provide further evidence consistent with our interpretation of the results reflecting the knowledge burden thesis. In particular, we document a disproportionate increase in researcher specialization (or, conversely, a decrease in diversification) in Soviet-rich subfields relative to Soviet-poor subfields after the fall of the Iron Curtain.

For this subsection, we switch the unit of analysis to the author-year and examine the degree of diversification for authors who published in a given year. We measure diversification using an author-level count of the number of subclassification areas (as defined by the AMS in their MSC schema) that were used in the authors' publications over the previous five years. Each of the 33 subfields in our data has a large number of subclassifications. Ideally we would know the contribution of each scholar to the publication. In the absence of that information, we note that the use of subclassification codes from papers may underestimate the overall increase in specialization.

If the shock to the knowledge frontier did indeed lead to increased specialization, then we expect to observe a disproportionate decline in the number of

Table 5—Teams in Soviet-Rich Subfields Exhibit a Disproportionate Increase in Propensity to Cite Soviet Pre-1990 Art after 1990. The Disproportionate Increase in References to Soviet Prior Art Is Driven by Larger-Sized Teams in Soviet-Rich Subfields

	Dependent variable: References to Soviet art			
	Count of Soviet references (defined by Soviet journal)	Percentage of Soviet references (defined by Soviet journal)	Count of Soviet references (defined by Soviet journal)	Percentage of Soviet references (defined by Soviet journal)
AfterIronCurtain × SovietRich × LogAuthorCount			0.5628***	0.0167***
			(0.1078)	(0.0029)
AfterIronCurtain × SovietRich	0.4020***	0.0172***	0.1416	0.0097**
	(0.0423)	(0.0036)	(0.0832)	(0.0037)
AfterIronCurtain × LogAuthorCount			0.0558**	0.0033
			(0.0197)	(0.0024)
SovietRich × LogAuthorCount			−0.2472**	−0.0122***
			(0.0922)	(0.0017)
LogAuthorCount			0.0412	0.0002
			(0.0787)	(0.0007)
Year fixed effects	Yes	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes	Yes
R^2	0.117	0.180	0.128	0.084
Observations	1,012	1,012	1,012	1,012

Notes: The unit of analysis is the publication. The sample includes the top/bottom three subfields drawn from the top 30 journals four years before and after the collapse of the Soviet Union. All models are OLS with robust standard errors, clustered by subfield.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

subclassifications for authors in Soviet-rich relative to Soviet-poor fields after 1990. We examine all authors, observing authors multiple times if they published more than once. Specifically, for author *a* in year *t*, we estimate:

$$(3) \qquad DegreeOfDiversification_{at} = \alpha(SovietRich_a \times AfterIronCurtain_t) + \mathbf{Subfield}_a + \gamma_t + \epsilon_{ia}.$$

We present results of two measures of specialization: a five-year rolling count of unique subclassifications codes per author per year and the same count normalized by the number of publications per author per year.

In Table 6, we show a disproportionate decrease in diversification for authors who published in Soviet-rich fields after the fall of the Soviet Union. Or, stated differently, we find evidence of a disproportionate increase in specialization for authors who published their work in Soviet-rich subfields. Furthermore, the decrease in diversification happened gradually over time in line with the knowledge burden mechanism (Jones 2009). Specifically, in Figure 7, we plot the yearly estimated difference between the weighted count of subclassification codes of authors publishing in Soviet -rich and Soviet-poor subfields. We observe that the decrease in diversification did not happen sharply after 1990, but rather gradually over time.

Table 6—Authors in Soviet-Rich Subfields Exhibit a Disproportionate Increase in Specialization

	Top and bottom three subfields only			
	Dependent variable is count of subfields	Dependent variable is count of subfields normalized by publication count	Dependent variable is count of subfields	Dependent variable is count of subfields normalized by publication count
AfterIronCurtain × SovietRich	−0.0979*** (0.0359)	−0.0068^ (0.0045)	−0.0412 (0.0428)	−0.0173*** (0.0052)
SovietRich	−0.3708*** (0.0276)	0.0378*** (0.0037)	N/A	N/A
Year fixed effects	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	Yes	Yes
R^2	0.024	0.008	0.001	0.002
Observations	315,161	315,161	315,161	315,161

Notes: Author specialization (or negative diversification) is measured by a count of distinct codes published over the last five years. We interpret fewer distinct codes as more specialized. Thus, a negative estimated coefficient implies that the covariate is positively correlated with specialization. All models are OLS with robust standard errors, clustered by author.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
^Significant at the 15 percent level.

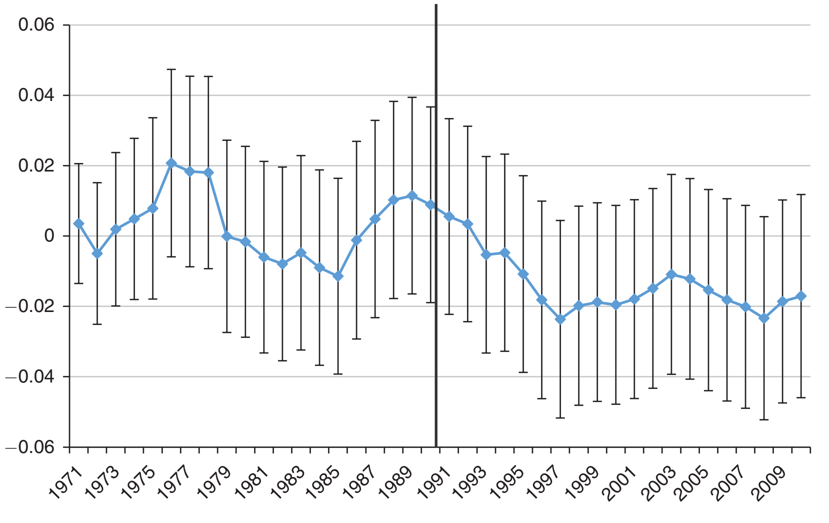

Figure 7. Plot of Estimated Rate of Specialization (*count of subfields normalized by count of publications*)

Notes: We base this figure on 40 years of publication data for authors publishing in the 3 top and 3 bottom ranked subfields of theoretical mathematics. Each point on the graph represents the coefficient value on the covariate *Year* × *SovietRich* and thus describes the relative difference in specialization rates between authors publishing in Soviet-rich and Soviet-poor fields in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 1970. We also present these results in table form in Appendix Table 8.

V. Discussion and Conclusion

We report evidence that the collapse of the Soviet Union led to a disproportionate increase in collaboration among non-Soviet mathematicians in the subfields that the Soviets were strongest. This occurred across many countries and is not correlated to the degree of local labor market competition by Soviet scholars for publication in local journals. We also report evidence of a disproportionate increase in citations to Soviet prior art as well as in specialization. We interpret this as consistent with the knowledge frontier explanation but not the other explanations for the widely documented increase in team size. In other words, although this evidence is not intended to (and does not) rule out the possibility that other explanations also played a role in the across-the-board increase in team size in science, we interpret it to provide support for the knowledge frontier mechanism as an explanation for at least some of the observed increase in scientific collaboration over time.

At the same time, it is important to clarify the limitations to our interpretation of these results. First, our results are based on the impact of a sudden outward shift in the knowledge frontier. An underlying assumption of our interpretation is that the team size response to a shock is similar to that for a gradual outward shift in the knowledge frontier. However, that may not be the case. Researchers may be able to absorb gradual increases in the knowledge frontier in a manner that does not generate as high returns to collaboration as those resulting from a sudden shock that may be more costly for researchers to internalize. Thus, our empirical results may not measure the impact of a gradual shift in the knowledge frontier.

Second, other factors resulting from the collapse of the Soviet Union surely also played a role on the field of mathematics. Borjas and Doran (2012, forthcoming, 2013) emphasize the labor market impact of increased competition from Soviet scholars. This increased competition may have also driven an increase in collaboration if, for example, returns to collaboration increased due to reasons such as risk mitigation (diversification of research projects). While we view our results as more consistent with the knowledge burden hypothesis than other explanations, we do not reject that changing labor markets also played a role.

Third, we focus on one particular field: mathematics. Adams et al. (2005) show that mathematics is somewhat of an outlier in team size relative to other disciplines. Mathematics has relatively small teams. In the first year of their study, 1981, mathematics publications had the fewest number of authors (of 12 fields). Furthermore, mathematics had the lowest annual growth rate in team size from 1981 to 1990 and the second lowest from 1990 to 1999. In contrast, physics and astronomy had the highest growth rates, which was likely at least partly driven by the increasing role of capital-intensive equipment (e.g., particle accelerators) in those fields. Therefore, our results may overestimate the effect, in terms of the *percentage* increase in team size, in fields where capital equipment plays a more central role.

In terms of the policy implications of these findings, Jones (2009, 2010, 2008) presents a variety of interventions that are potentially welfare-enhancing in the presence of the knowledge frontier effect. These include subsidies and rewards to incentivize entry into research careers, team-based evaluation of grant applications, and national or regional subsidies and specialization to prevent poverty traps due to underinvestment in human capital from coordination failures arising from thin markets for complementary skills. Although our study offers no means by which to comment on the suitability of these interventions to particular policy settings, the evidence we present here does suggest that the knowledge frontier phenomenon is worthy of further research and policy attention.

Appendix

Table A1—Robustness to Alternative Ranking Measures

	Dependent variable: log of author count				
	Top and bottom three subfields	Soviet-rich defined as top three of subfields	Soviet-rich defined as top five of subfields	Soviet-rich defined as top ten of subfields	Continuous
AfterIronCurtain	0.0800***	0.0475***	0.0404***	0.0364**	0.0012
× SovietRich	(0.0126)	(0.0107)	(0.0150)	(0.0157)	(0.0010)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.112	0.107	0.107	0.106	0.106
Observations	133,497	563,462	563,462	563,462	563,462

Notes: In this table, we use our own ranking measure rather than Borjas and Doran (2012); our measure uses worldwide publications from 1970–1989 rather than US publications from 1984–1989 and relies on identifying Soviet publications by ethnicity of names rather than by affiliation data. All models are OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A2—Robustness to Including Soviet Papers

	Dependent variable: log of author count				
	Top and bottom three subfields	Soviet-rich defined as top three of subfields	Soviet-rich defined as top five of subfields	Soviet-rich defined as top ten of subfields	Continuous
AfterIronCurtain	0.0696***	0.0424***	0.0346**	0.0283*	0.0010
× SovietRich	(0.0110)	(0.0108)	(0.0136)	(0.0154)	(0.0011)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.107	0.103	0.102	0.102	0.102
Observations	169,305	689,793	689,793	689,793	689,793

Notes: In this table, we present estimates of the main effect when including Soviet authors back into the sample. The unit of analysis is the publication. All models are OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A3—Robustness to Poisson Estimation

	Dependent variable: log of author count				
	Top and bottom three subfields	Soviet-rich defined as top three of subfields	Soviet-rich defined as top five of subfields	Soviet-rich defined as top ten of subfields	Continuous
AfterIronCurtain	0.0916***	0.1010***	0.0887***	0.0841***	0.0027***
× SovietRich	(0.0268)	(0.0155)	(0.0150)	(0.0125)	(0.0006)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes	Yes	Yes
LL	−90,211.22	−370,710.15	−370,713.82	−370,708.67	−370,720.84
Observations	133,498	563,462	563,462	563,462	563,462

Notes: In this table, we present estimates of the main effect using a Poisson estimation. The unit of analysis is the publication.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A4—Coefficient Estimates Used to Plot Figure 3

Dependent variable: log of author count per publication per year Top and bottom three subfields only			
SovietRich × 1971	0.0137	SovietRich × 1991	0.0403
	(0.0139)		(0.0257)
SovietRich × 1972	0.0452*	SovietRich × 1992	0.0635
	(0.0147)		(0.0360)
SovietRich × 1973	0.0428*	SovietRich × 1993	0.0624**
	(0.0212)		(0.0200)
SovietRich × 1974	0.0290	SovietRich × 1994	0.0700*
	(0.0156)		(0.0303)
SovietRich × 1975	0.0389	SovietRich × 1995	0.0691**
	(0.0250)		(0.0178)
SovietRich × 1976	0.0174***	SovietRich × 1996	0.0513
	(0.0030)		(0.0280)
SovietRich × 1977	0.0218*	SovietRich × 1997	0.0750**
	(0.0089)		(0.0273)
SovietRich × 1978	0.0465*	SovietRich × 1998	0.0960*
	(0.0218)		(0.0259)
SovietRich × 1979	0.0550**	SovietRich × 1999	0.0792**
	(0.0187)		(0.0274)
SovietRich × 1980	0.0373	SovietRich × 2000	0.0773*
	(0.0249)		(0.0303)
SovietRich × 1981	0.0269*	SovietRich × 2001	0.0958***
	(0.0111)		(0.0175)
SovietRich × 1982	0.0422*	SovietRich × 2002	0.1279***
	(0.0190)		(0.0274)
SovietRich × 1983	0.0256	SovietRich × 2003	0.1206***
	(0.0175)		(0.0173)
SovietRich × 1984	0.0327	SovietRich × 2004	0.1000*
	(0.0449)		(0.0461)
SovietRich × 1985	0.0604*	SovietRich × 2005	0.1376***
	(0.0239)		(0.0303)
SovietRich × 1986	0.0353	SovietRich × 2006	0.1406***
	(0.0306)		(0.0302)
SovietRich × 1987	0.0406	SovietRich × 2007	0.1806***
	(0.0294)		(0.0329)
SovietRich × 1988	0.0127	SovietRich × 2008	0.1680***
	(0.0247)		(0.0312)
SovietRich × 1989	0.0289*	SovietRich × 2009	0.1518***
	(0.0125)		(0.0239)
SovietRich × 1990	0.0119	SovietRich × 2010	0.1717***
	(0.0270)		(0.0341)
Year fixed effects	Yes	Subfield fixed effects	Yes
R^2	0.114	Observations	133,497

Note: Estimated with OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A5—Coefficient Estimates Used to Plot Figure 5—All Japanese Journals

Dependent variable: log of author count per publication per year Top and bottom three subfields only			
SovietRich × 1971	0.0417 (0.0626)	SovietRich × 1991	0.1454 (0.0987)
SovietRich × 1972	−0.0394 (0.1071)	SovietRich × 1992	0.2342*** (0.0865)
SovietRich × 1973	−0.0024 (0.0685)	SovietRich × 1993	0.0987 (0.0835)
SovietRich × 1974	−0.0106 (0.0509)	SovietRich × 1994	0.1957*** (0.0482)
SovietRich × 1975	0.0219 (0.0229)	SovietRich × 1995	0.0986*** (0.0368)
SovietRich × 1976	−0.0330 (0.0625)	SovietRich × 1996	0.1824*** (0.0815)
SovietRich × 1977	−0.0157 (0.0606)	SovietRich × 1997	0.0410 (0.0532)
SovietRich × 1978	0.1095 (0.0842)	SovietRich × 1998	0.1010* (0.0532)
SovietRich × 1979	0.0507 (0.0951)	SovietRich × 1999	0.1688** (0.0737)
SovietRich × 1980	0.0897 (0.0658)	SovietRich × 2000	0.0660*** (0.0089)
SovietRich × 1981	0.0909 (0.0862)	SovietRich × 2001	0.1316* (0.0771)
SovietRich × 1982	0.0609 (0.0604)	SovietRich × 2002	0.0026 (0.0601)
SovietRich × 1983	−0.0120 (0.0837)	SovietRich × 2003	0.0990 (0.0772)
SovietRich × 1984	0.0065 (0.0854)	SovietRich × 2004	0.2824*** (0.0518)
SovietRich × 1985	0.0665 (0.0574)	SovietRich × 2005	−0.0504 (0.0789)
SovietRich × 1986	0.0680 (0.0891)	SovietRich × 2006	0.0699 (0.0585)
SovietRich × 1987	0.1679** (0.0826)	SovietRich × 2007	0.1625** (0.0692)
SovietRich × 1988	−0.0333 (0.0629)	SovietRich × 2008	0.0574 (0.0653)
SovietRich × 1989	0.1737*** (0.0610)	SovietRich × 2009	0.1257*** (0.0484)
SovietRich × 1990	0.0841 (0.0549)	SovietRich × 2010	−0.0082 (0.0657)
Year fixed effects	Yes	Subfield fixed effects	Yes
R^2	0.114	Observations	133,497

Note: Estimated with OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A6—Coefficient Estimates Used to Plot Figure 6—Part 1

	Dependent variable: log of author count							
	China	Japan	Korea	India	Canada	Italy	Belgium	Spain
AfterIronCurtain × SovietRich	0.0864*** (0.0680)	0.0692*** (0.0173)	0.0848 (0.1051)	0.0372 (0.0912)	0.0224 (0.0454)	0.0373*** (0.0108)	−0.0123 (0.0471)	−0.0778 (0.0740)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.158	0.076	0.087	0.090	0.111	0.076	0.249	0.115
Observations	11,838	5,096	931	2,695	1,250	5,261	523	672

Notes: The unit of analysis is the publication. All models are OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A7—Coefficient Estimates Used to Plot Figure 6—Part 2

	Dependent variable: log of author count						
	Netherlands	Switzerland	US	UK	Singapore	Germany	France
AfterIronCurtain × SovietRich	0.0703** (0.0336)	0.1360*** (0.0324)	0.08176*** (0.0115)	0.0399 (0.0296)	−0.0175 (0.0635)	0.0730*** (0.0142)	0.0938** (0.0394)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subfield fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.097	0.170	0.101	0.097	0.113	0.131	0.081
Observations	17,730	4,611	32,080	13,395	2,243	9,644	2,933

Notes: The unit of analysis is the publication. All models are OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A8—Coefficient Estimates Used to Plot Figure 7

Dependent variable: Count of subfields weighted by publication count Top and bottom three subfields only			
SovietRich × 1971	0.0035 (0.0087)	SovietRich × 1991	0.0055 (0.0142)
SovietRich × 1972	−0.0050 (0.0103)	SovietRich × 1992	0.0034 (0.0142)
SovietRich × 1973	0.00019 (0.0111)	SovietRich × 1993	−0.0053 (0.0142)
SovietRich × 1974	0.0048 (0.0117)	SovietRich × 1994	−0.0048 (0.0143)
SovietRich × 1975	0.0078 (0.0131)	SovietRich × 1995	−0.0108 (0.0143)
SovietRich × 1976	0.0207 (0.0136)	SovietRich × 1996	−0.0182 (0.0143)
SovietRich × 1977	0.0183 (0.0138)	SovietRich × 1997	−0.0237* (0.0143)
SovietRich × 1978	0.0180 (0.0139)	SovietRich × 1998	−0.0198 (0.0144)
SovietRich × 1979	−0.0001 (0.0140)	SovietRich × 1999	−0.0188 (0.0144)
SovietRich × 1980	−0.0016 (0.0138)	SovietRich × 2000	−0.0196 (0.0144)
SovietRich × 1981	−0.0060 (0.0138)	SovietRich × 2001	−0.0180 (0.0144)
SovietRich × 1982	−0.0079 (0.0140)	SovietRich × 2002	−0.0149 (0.0145)
SovietRich × 1983	−0.0048 (0.0140)	SovietRich × 2003	−0.0109 (0.0145)
SovietRich × 1984	−0.0090 (0.0142)	SovietRich × 2004	−0.0122 (0.0145)
SovietRich × 1985	−0.0114 (0.0142)	SovietRich × 2005	−0.0154 (0.0146)
SovietRich × 1986	−0.0012 (0.0143)	SovietRich × 2006	−0.0182 (0.0147)
SovietRich × 1987	0.0048 (0.0142)	SovietRich × 2007	−0.0201 (0.0147)
SovietRich × 1988	0.0102 (0.0143)	SovietRich × 2008	−0.0233 (0.0147)
SovietRich × 1989	0.0115 (0.0143)	SovietRich × 2009	−0.0186 (0.0147)
SovietRich × 1990	0.0089 (0.0142)	SovietRich × 2010	−0.0171 (0.0147)
Year fixed effects	Yes	Subfield fixed effects	Yes
R^2	0.114	Observations	133,497

Note: Estimated with OLS with robust standard errors, clustered by subfield.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table A9—Robustness to Authors in Soviet-Rich Subfields Exhibit a Disproportionate Increase in Specialization

	Top and bottom three subfields only			
	Dependent variable is count of subfields	Dependent variable is count of subfields weighted by publication count	Dependent variable is count of subfields	Dependent variable is count of subfields weighted by publication count
AfterIronCurtain × SovietRich	−0.1595***	−0.0006	−0.1888***	−0.0043
	(0.0334)	(0.0045)	(0.0429)	(0.0053)
SovietRich	−0.2006***	0.0159***	N/A	N/A
	(0.0253)	(0.0036)		
Age	0.1145***	−0.0116***	0.1153***	−0.0104***
	(0.0026)	(0.0003)	(0.0420)	(0.0003)
Age^2	−0.0030***	0.0003***	−0.0026***	0.0002
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Year fixed effects	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	Yes	Yes
R^2	0.098	0.056	0.078	0.047
Observations	315,161	315,161	315,161	315,161

Notes: Author specialization (or negative generalization) is measured by a count of distinct codes published over the last five years. We interpret fewer distinct codes as more specialized. Thus, a negative estimated coefficient implies that the covariate is positively correlated with specialization. All models are OLS with robust standard errors, clustered by author.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

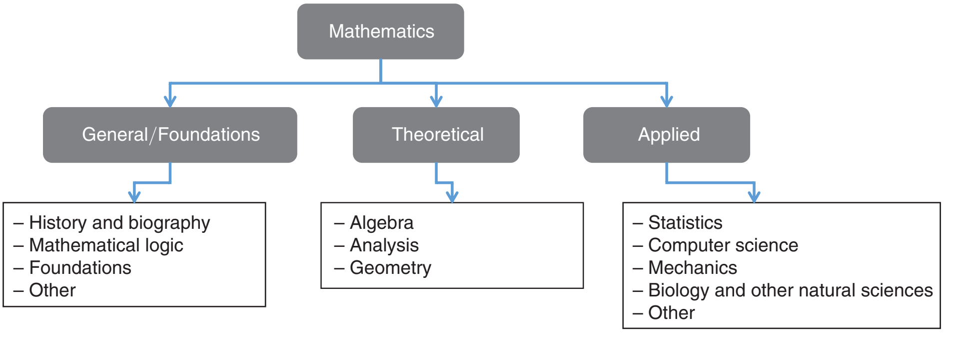

Figure A1. Mathematics Taxonomy

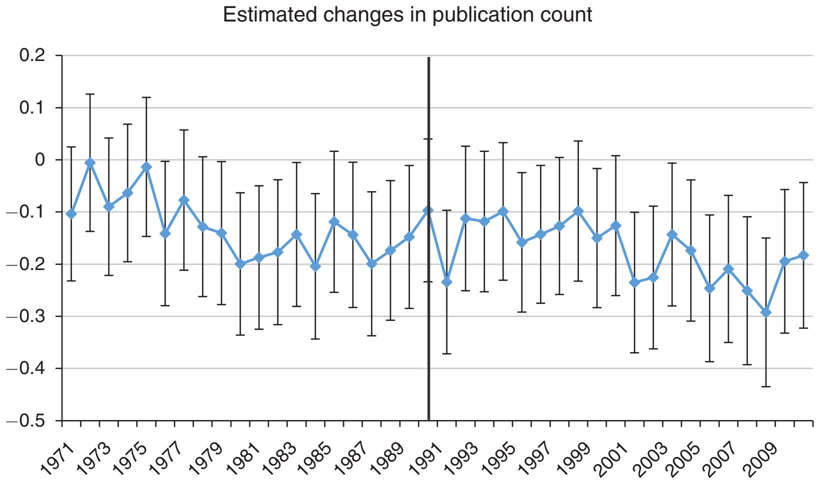

Figure A2. Estimated Difference in Yearly Publication Count for Authors in Top and Bottom Three Subfields

Notes: We base this figure on 40 years of publication data for authors who published in the 3 top and 3 bottom ranked subfields of theoretical mathematics. We plot the estimated difference in publication counts for Soviet-rich and Soviet-poor authors, before and after 1990.

REFERENCES

Acemoglu, Daron, Tarek A. Hassan, and James A. Robinson. 2011. "Social Structure and Development: A Legacy of the Holocaust in Russia." *Quarterly Journal of Economics* 126 (2): 895–946.

Adams, James D., Grant C. Black, J. Roger Clemmons, and Paula E. Stephan. 2005. "Scientific Teams and Institutional Collaborations: Evidence from U.S. Universities, 1981–1999." *Research Policy* 34 (3): 259–85.

Agrawal, Ajay, and Avi Goldfarb. 2008. "Restructuring Research: Communication Costs and the Democratization of University Innovation." *American Economic Review* 98 (4): 1578–90.

Agrawal, Ajay, Avi Goldfarb, and Florenta Teodoridis. 2016. "Understanding the Changing Structure of Scientific Inquiry: Dataset." *American Economic Journal: Applied Economics*. http://dx/doi.org/10.1257/app.20140135.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Difference-In-Differences Estimates?" *Quarterly Journal of Economics* 119 (1): 249–75.

Bikard, Michael, Fiona Murray, and Joshua Gans. Forthcoming. "Exploring Trade-offs in the Organization of Scientific Work: Collaboration and Scientific Reward." *Management Science*.

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. 2013. "Which Peers Matter? The Relative Impacts of Collaborators, Colleagues and Competitors." University of Notre Dame Working Paper 021.

Borjas, George J., and Kirk B. Doran. Forthcoming. "Cognitive Mobility: Native Responses to Supply Shocks in the Space of Ideas." *Journal of Labor Economics*.

Dubois, Pierre, Jean-Charles Rochet, and Jean-Marc Schlenker. 2014. "Productivity and mobility in academic research: Evidence from mathematicians." *Scientometrics* 98 (3): 1669–1701.

Fons-Rosen, Christian. 2012. "Knowledge Flows Through FDI: The Case of Privatisations in Central and Eastern Europe." http://www.econ.upf.edu/~fonsrosen/images/CFR_JMP_20121030_all.pdf.

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.

Grossman, Jerrold W. 2002. "The Evolution of the Mathematical Research Collaboration Graph." *Congressus Numerantium* 158: 201–12.

Hesse, Bradford W., Lee S. Sproull, Sara B. Kiesler, and John P. Walsh. 1993. "Returns to Science: Computer Networks in Oceanography." *Communications of the ACM* 36 (8): 90–101.

Jones, Benjamin F. 2008. "The Knowledge Trap: Human Capital and Development, Reconsidered." National Bureau of Economic Research (NBER) Working Paper 14138.

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.

Jones, Benjamin F., and Benjamin A. Olken. 2005. "Do Leaders Matter? National Leadership and Growth Since World War II." *Quarterly Journal of Economics* 120 (3): 835–64.

Kim, Han E., Adair Morse, and Luigi Zingales. 2009. "Are Elite Universities Losing Their Competitive Edge?" *Journal of Financial Economics* 93 (3): 353–81.

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

Mowery, David C., Richard R. Nelson, Bhaven N. Sampat, and Arvids A. Ziedonis. 2004. *Ivory Tower and Industrial Innovation: University-Industry Technology Transfer before and after the Bayh-Dole Act*. Stanford: Stanford University Press.

Newman, Mark E. J. 2004. "Coauthorship Networks and Patterns of Scientific Collaboration." *Proceedings of the National Academy of Science* (*PNAS*) 101 (1): 5200–5205.

Oettl, Alex. 2012. "Reconceptualizing Stars: Scientist Helpfulness and Peer Performance." *Management Science* 58 (6): 1122–40.

Polyak, B. T. 2002. "History of mathematical programming in the USSR: Analyzing the phenomenon." *Mathematical Programming* 91 (3): 401–16.

Rosen, Sherwin. 1981. "The Economics of Superstars." *American Economic Review* 71 (5): 845–58.

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.

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

Sonnenwald, Diane H. 2007. "Scientific Collaboration." *Annual Review of Information Science and Technology* 41 (1): 643–81.

Stephan, Paula. 2012. *How Economics Shapes Science*. Cambridge: Harvard University Press.

Stuen, Eric T., Ahmed Mushfiq Mobarak, and Keith E. Maskus. 2012. "Skilled Immigration and Innovation: Evidence from Enrollment Fluctuations in US Doctoral Programmes." *Economic Journal* 112 (565): 1143–76.

Waldinger, Fabian. 2010. "Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany." *Journal of Political Economy* 118 (4): 787–831.

Waldinger, Fabian. 2012. "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany." *Review of Economic Studies* 79 (2): 838–61.

Walsh, John P., and Todd Bayma. 1996. "Computer Networks and Scientific Work." *Social Studies of Science* 26 (3): 661–703.

Wuchy, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. "The Increasing Dominance of Teams in Production of Knowledge." *Science* 316 (5827): 1036–39.

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