The Collapse of the Soviet Union
and the Productivity of American Mathematicians

George J. Borjas and Kirk B. Doran
Harvard University and University of Notre Dame

May 2011

The Collapse of the Soviet Union and the Productivity of American Mathematicians

George J. Borjas and Kirk B. Doran

ABSTRACT

Many economists argue that human capital externalities are the source of long-term growth, and recommend an increase in the supply of highly skilled workers as a strategy to increase other workers' productivity. Using unique data on publications, citations, and affiliations of mathematicians in the United States and the former Soviet Union, we examine the impact of a large post-1992 influx of Soviet mathematicians on the productivity of their American counterparts. We find a negative productivity impact on those mathematicians whose research overlapped with that of the Soviets. We also document an increased mobility rate (to lower-quality institutions and out of active publishing) and a shift away from fields dominated by Soviet mathematicians (particularly among new entrants). The evidence suggests that gains from human capital externalities generated by new high-skill workers might only be fully captured if policy-makers adjust the amount of resources allocated to the affected fields.

The Collapse of the Soviet Union and the Productivity of American Mathematicians

George J. Borjas and Kirk B. Doran*

I. Introduction

Many economists argue that the source of long-term economic growth lies in human capital. The foundation of modern models of economic growth, in fact, is the assumption that there are strong human capital externalities (Lucas, 1988; Romer, 1986, 1990). In response to this idea, there is a consensus among policy makers that increases in the supply of highly skilled workers (through both immigration and increased training) will increase the productivity of the pre-existing high-skill workforce, and lead to a substantial increase in national wealth.

But even if the ideas of a highly qualified single worker spill over to other workers with whom they interact, the overall effect of the interaction can still be deleterious to the net productivity of other workers. In particular, in a world with constraints on the funding and dissemination of ideas (e.g., a limit on the number of faculty slots, or, more abstractly, a limit on the attention span of the potential audience), large and sudden increases in the population of producers of knowledge can result in diminishing marginal productivity for a pre-existing worker. As an example, a young academic might appreciate the hiring of a new illustrious colleague in, say, mathematics because it may improve his own ideas. At the same time, the young academic realizes that in a world with limited funding and limited research opportunities, his own services and research now become relatively less important to his own department and to the field in general.

A number of recent empirical studies have attempted to quantify the net impact of the presence of one highly skilled worker on the ideas and output of other workers. For example, Waldinger (2009) examines the productivity of the doctoral students who were

---

* Dr. Borjas is a Professor of Economics and Social Policy at the Harvard Kennedy School, Cambridge, MA, and Dr. Doran is an Assistant Professor of Economics at the University of Notre Dame, South Bend, IN. We are grateful to Patrick Ion, Graeme Fairweather, Norm Richert, and Erol Ozil from the American Mathematical Society for extensive collaboration and support in preparing the data. This project would have been impossible without the excellent programming assistance of Andrew Stellman of Stellman and Greene Consulting. The authors are grateful to the Upjohn Institute, the Kauffman Foundation, and the Sloan Foundation for their financial assistance.

left behind when superstar German scientists left Germany during the Nazi era. He finds that these students suffered in the absence of their highly skilled mentors. In more recent work, however, Waldinger (2011) finds that the coauthors left behind in Nazi Germany did not experience a loss in productivity when the superstars left. Finally, Azoulay, Zivin, and Wang (2010) document the decreased output suffered by the co-authors of superstar scientists after the superstars die. They find that the coauthors become much less productive when the superstar is no longer able to collaborate. In concluding, they note: "Although we measure the impact of losing a star collaborator, a full accounting of knowledge spillovers would require information on the benefits that accrued to the field while the star was alive. We can think of no experiment, natural or otherwise, that would encapsulate this counterfactual."

This paper attempts to measure the productivity effects of the entry of superstar scientists in a context where we can observe the counterfactual of no entry. In particular, we examine the impact of the influx of renowned Soviet mathematicians into the American mathematics community.[1] In the period between the establishment and fall of communism, Soviet mathematics developed in an insular fashion and along very different specializations than American mathematics. As a result, some U.S. subfields experienced no potential insights from Soviet mathematics after the collapse of the Soviet Union, while other fields experienced a flood of new mathematicians, theorems, and ideas.

We have constructed a data set that contains information on the authorship and affiliation of nearly every paper published in mathematics over the past 40 years. These data allow us to document the location, affiliation, and complete publication and citation records of mathematicians who were active in either the Soviet Union or the United States for the past few decades. Crucially, we have uniquely identified mathematicians even when they have the same name.

After the collapse of the Soviet Union, several hundred world-class Soviet mathematicians emigrated to other countries, with a large percentage settling in the United

---

[1] Recent applications of modern academic research papers in mathematics and related fields (such as theoretical computer science and mathematical physics) to our broader economy are so numerous and diverse that it is impossible to characterize them briefly. For a few examples, consider the Rivest-Shamir-Adleman algorithm that forms the backbone of Internet encryption, or the Reed–Solomon error correction that is used everywhere from Blu-ray Disks to deep-space communication to DSL television.

States. Even though important Soviet treatises were painfully translated into English over a process of many years and after much delay, there was almost no collaboration and little free communication between Soviet and American mathematicians during the Cold War.[2] In fact, every communication with an American mathematician was opened and read by the authorities, and special permission was required to publish outside the Soviet Union. Depending on the era in the Cold War, Soviet violators could be imprisoned or even executed (Polyak, 2001).

Our empirical analysis demonstrates unambiguously that almost all American mathematicians whose research agenda overlapped with that of the Soviets experienced a reduction in productivity after the entry of Soviet émigrés, the notable exception being the few tenured mathematicians who were placed at the most highly ranked institutions. Regardless of whether there is lower productivity arising from constraints in the dissemination of ideas or an increased productivity due to external effects, we would expect to observe behavioral reactions to these changes. Our analysis also examines the behavioral consequences of the Soviet influx. We find that American mathematicians whose research overlapped with that of the Soviets became much more likely to switch institutions; that the switch entailed a move to a much lower quality institution; that many of these American mathematicians ceased publishing relatively early in their career; that there was an increase in collaboration with the Soviet mathematicians among the institutions that hired the Soviet émigrés; and that new entrants into the mathematics community were less likely to pursue a research agenda in Soviet-dominated fields. In the aggregate, based on the pre-1990 age-output profile of American mathematicians, we can predict almost perfectly the total output of mathematicians whose work had little overlap with the Soviet research agenda, while the actual output of mathematicians working in areas similar to those of the Soviets is far below what one would have expected.

Our results are not consistent with the optimistic view emphasized by Jones and Romer (2010) that the ideas of one highly skilled worker will induce greater idea generation by other highly skilled workers. In contrast, we find that the average worker in

---

[2] Abramitzky and Sin (2010) exhaustively study the history of book translations between Western and East Bloc countries, and find very few translations from communist countries to Western ones in any fields of "exact science," including mathematics, either before or after the collapse of the Soviet Union.

the pre-existing workforce becomes less productive when new ideas *and* workers are exogenously introduced into the market. Our evidence thus suggests that it is far from inevitable that the spillovers emphasized by the economic growth literature will "win" the race between human capital externalities and the law of diminishing returns that arises from resource scarcity. In particular, human capital externalities will likely tend to win this race when capital adjusts appropriately; in some sectors important to the knowledge economy there is reason to doubt that capital adjusts quickly to the flow of new ideas; and there is a limited time horizon for capital to adjust before the new high-skill workers have retired or the pre-existing high-skill workers have been driven out of the market.

At a practical level, our evidence probably stems from constraints on the funding of research, from market failures associated with how resources are allocated in academic institutions, and from constraints in the opportunities to disseminate ideas. It seems, therefore, that the gains from the human capital externalities generated by the influx of highly skilled workers can only be fully captured if policy makers or academic institutions correspondingly adjust the amount of resources allocated to the affected fields. More generally, however, it is uncertain that even a substantial increase in resources can fully overcome the implications of the law of diminishing returns. Put simply, *real* human beings with limited time, attention span, and effort may be unable to disseminate and assimilate new knowledge in a way that reverses the crowd-out effect.

II. Historical Context

After the establishment of the Soviet Union in 1922, Soviet mathematics entered a long period of development independent from mathematics in the West. To varying degrees between 1922 and 1992, the Soviet Government instituted strict controls on which scientists could communicate with Western peers, on the parameters of scientific travel, on the acceptable outlets for publication, and on access to Western materials.[3] Just as speakers

---

[3] (Polyak, 2001) gives a firsthand account of the life of mathematicians behind the Iron Curtain. Polyak writes: "Academician N.N. Lusin was exposed to a humiliating execution after the publication of his papers in Western journals in 1936. . . When Professor Ya.Z. Tsypkin received a letter in the late 1940s from an American reader of his paper, he was summoned by the KGB and underwent a long investigation there, tottering at the edge of arrest. . . Another source of difficulties for researchers was the mania for secrecy. . . Nobody was allowed to publish any paper without special permission confirming that the publication does

of one language, when separated geographically for many generations, eventually develop separate and different dialects through natural changes over time, so Western and Eastern mathematicians, separated by Stalinist and Cold War political institutions, developed under different influences to the point of achieving very different specializations across the fields of mathematics.

Figure 1 illustrates the dramatic differences in specializations between Soviet and American mathematics. Prior to 1990, Soviet mathematicians accounted for 13 percent of all papers published in Partial Differential Equations, but these same mathematicians only accounted for two percent of the papers published in Algebraic Topology, and less than one percent of the papers published in Statistics.[4]

The primary influences behind the development of both Soviet and Western mathematics were history dependence and, to a lesser extent, state funding. In the Soviet Union, for example, the mathematical genius Andrey Kolmogorov developed important results in the area of Probability and Stochastic Processes beginning in the 1930s. In a scenario common throughout Soviet mathematical history, he then established a "school" at Moscow State University, attracting some of the best young minds over the next four decades, such as the teenage prodigy Vladimir Arnold in the 1950s. Arnold himself quickly solved Hilbert's famous "Thirteenth Problem," and initiated the field of symplectic topology. The large amount of Soviet work in these areas even in the 1980s can be seen in Figure 1 under the subject headings "Mechanics of Particles and Systems," "Global Analysis, Analysis on Manifolds," and "Ordinary Differential Equations." Since the United States did not have the unique combination of the teacher Kolmogorov and student Arnold, the amount of work in these subfields done in the United States in the 1980s was far less than

---

[sic] not contradict numerous security restrictions. All letters abroad (as well as letters from abroad) were opened and inspected. Everybody must have special permission, and a full text of the talk had to be approved if you were going to an international conference. And working in a classified institution (which was the case for many experts in mathematical programming), complicated the situation drastically. . . . The situation in the 1940s to mid-1950s was the worst. Malevolent intent by the authorities could lead a researcher to the GULAG. The period 1955–1970 was the least oppressive, it was a 'golden age' of Soviet mathematics. . . The years 1970 to 1985 were a period of stagnation in political, social, economic, and scientific life. All the troubles I have mentioned above played a more and more significant role in the development of Soviet science and thus led to its degradation."

[4] The denominator in these percentages is the number of papers published by mathematicians in American affiliations pre-1990.

would have been expected given the size and breadth of the American mathematics community.

In the United States, however, researchers like John Milnor at Princeton University and Raoul Bott at Harvard developed key ideas in the Topology of Manifolds (manifolds and cell complexes) in the 1950s. Their students and collaborators produced an enviable body of research in Geometric Topology, which, because of the lack of a similar chance originator in their country, was never replicated in the Soviet Union.[5]

Algebraic Geometry, a field of relative U.S. excellence, provides another example of how persistent history-dependence can be. The Summary Report of the Panel on Soviet Mathematics (Lefschetz, 1961) explains that: "in no part of mathematics is the Soviet Union weaker than in algebraic geometry. No significant contributions have ever . . . come from there." This was not only true in 1961, but, as Figure 1 demonstrates, it remained true in the 1980s.[6]

Finally, Soviet funding was limited (in comparison to the U.S.) in fields requiring experiments or equipment (Howe, 1990). Figure 1 also demonstrates that as late as the 1980s, this resulted in a large discrepancy between American and Soviet specialization in computer science and related fields.[7]

Figure 2 presents a different perspective on the different field specializations of Soviet and American mathematicians between 1970 and 1989. In particular, the figure illustrates the fraction of papers published by mathematicians in each country in each of the six most popular fields. It is worth noting that there is overlap in only two fields out of the top six fields. (In addition, only three fields appear in common in the Top Ten list of fields in each country).

We argue that the differing specializations of Soviet scientists as compared to American ones in the years before the Soviet collapse are largely due to history

---

[5] Also see Dubois, Rochet, and Schlenker (2011) for additional information about trends in publications and mobility of mathematicians in the past three decades.

[6] The definitions of Soviet and American mathematicians are discussed in detail below.

[7] Personal communication with Professor Lawrence Shepp, who during a long career at Bell Labs was in contact with some Soviet mathematicians as early as 1964, suggests that the fact that Soviet mathematicians avoided the most popular American field (Statistics) had little to do with funding. Specializing in statistics was politically dangerous in the USSR, as it would require a great deal of massaging of sensitive data.

dependence and state funding opportunities. Specifically, we argue that Soviet and American mathematicians did not choose their specializations in the late 1970s to mid-1980s in the belief that they would soon have an opportunity to co-author papers, compete for pages in the same journals, and apply for jobs at the same universities. In support of this claim, we cite both the restrictions on international collaboration mentioned earlier and the fact that the consensus among both Soviet and American experts almost immediately before the collapse of communism was that the political system of the existing Soviet state was not ripe for a sudden change.

Even until the late 1980s, predictions of the impending collapse of the Soviet Union were laughed at on both sides of the Iron Curtain. Walter Laqueur (1996) describes how in the Soviet Union itself, "most believed the system was so strong that it would never essentially change. Others, more optimistic, thought that change was perhaps possible over a long period – decades, or more likely, generations." In the West, Laqueur reports that Sovietologists were taken by surprise: "the U.S. government (like most others) had enormously overrated Soviet economic performance and . . . the statisticians, in the intelligence community as in academe, were in a state of disarray. According to a study published as late as 1988 by a well-known Western economist specializing in the Soviet Union, Soviet citizens enjoyed 'massive economic security' . . . the consensus was that the Soviet Union was *not* on the verge of economic bankruptcy and political disintegration."

Thus, the divergent interests and capabilities of Soviet and American mathematics which had emerged in earlier decades were not likely to have been modified in the 1980s by any serious belief that Soviet scientists would soon have the opportunity to migrate to the United States *en masse* to compete with Americans for jobs and pages in Western journals. In fact, as Figure 3 shows, the very small rate of co-authorship between people reporting Soviet research addresses and people reporting U.S. research addresses, demonstrates that there was no discernable trend before 1990 that would have engendered the hope that such collaborations would explode so suddenly in 1990 and beyond.

For these reasons, we treat the difference between the pre-1990 Soviet distribution of mathematical work across the subfields of mathematics and the pre-1990 U.S. distribution as exogenous. This assumption gives us a rare opportunity for empirical research: we know not only the distribution of actual immigrants across skill categories (i.e., fields), but we also know the pool of potential immigrants across skill categories. As Figure 1 also shows, the Soviet émigrés continued publishing more in those fields that were Soviet-dominated before the collapse of the Soviet Union.

Around 1990, as the political situation deteriorated in the Soviet Union, a large number of Soviet mathematicians began to come into regular contact with Western mathematicians through visits and immigration. According to American mathematicians who experienced this sudden increase in contact, the effect on American mathematics was immediate. In 1990, the *New York Times* reported (Kolata, 1990):

> "American scientists say they have benefited immensely from the [recent] Soviet visitors. . . 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 [Soviet mathematician] 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 . . . No one could help. But . . . Soviet scientists had done a lot of work on such problems. 'It was a whole new world I had access to,' Dr. Diaconis said. 'Together, we'll be able to solve the problem.'"

But in addition to these benefits, the American mathematical community also experienced increased competition in hiring.[8] The American Mathematical Society's 1991-1992 Academic Hiring Survey reports on the rapid incorporation of the "substantial influx of highly trained mathematicians from abroad" into the American university workforce. "Citizens of Eastern European countries and the former Soviet Union accounted for 13% of all newly-hired faculty and 15% of the tenured and tenure-eligible new hires. 71% of the Eastern European/Soviet citizens received their doctorate in Eastern Europe or the Soviet Union." The report also identifies "increased numbers of highly qualified recent U.S.

---

[8] For example, Soviet mathematician Nicolai Reshetikhin did not only help Persi Diaconis solve his puzzle, but he also was hired as a Visiting and Assistant Professor at Harvard from 1989 through 1991.

immigrants seeking employment in academia" as a leading cause of the unprecedented 12% unemployment rate of new American mathematics Ph.Ds. in 1991 (McClure, 1992). Figure 4 illustrates the trends in employment and unemployment of newly minted doctorates from North American institutions. It is evident that there was a dramatic increase in the unemployment rate (as well as a dramatic decrease in the probability of obtaining a position in research universities) at the same time that the Soviet influx was occurring.[9]

Our empirical analysis of the productivity of American mathematicians will measure the net effect of these two concurrent and competing factors: increased exposure to new ideas and increased competition for a limited number of jobs and other resources.

III. Theory

Modern theories of economic growth emphasize how human capital externalities alter the productivity of specific workers. A worker surrounded by many high-skill workers will himself become more productive by being exposed to new ideas and concepts. It would not be surprising, therefore, that the migration of world-class Soviet mathematicians to the United States following the collapse of the Soviet Union exposed American mathematicians to many new techniques and theorems, and increased the productivity of the American mathematics community.

This spillover effect, however, must coexist with the traditional laws of scarcity and diminishing returns that form the core of economic analysis. Following the excellent summary of the literature provided by Jones and Romer (2010), suppose that the production function for "mathematical knowledge" Y depends on the stock of ideas D, the stock of resources K used as inputs (e.g., faculty slots, computing resources), and the stock of mathematicians L. Capital and labor are rival goods, while ideas are nonrival. It is typically assumed that there are constant returns to K and L. Suppose that the production function can then be written as:

---

[9] The exodus of key scientific personnel from the former Soviet Union to the West led George Soros to establish a program that provided research funds to those scientists who chose to remain; see Ganguli (2010) for an analysis of the impact of this program on career choices.

(1) $$Y = D^{\phi}\left(\alpha_K K^{\delta} + \alpha_L L^{\delta}\right)^{1/\delta},$$

where ϕ gives the "externalities elasticity," the percent expansion in mathematical knowledge associated with a one percent increase in the stock of ideas.

Suppose, as is common in the literature, that the stock of ideas is proportional to the number of mathematicians. The marginal product of a mathematician is then given by:

(2) $$MP_L = AP_L\left(\phi + \alpha_L Q^{-\delta} L^{\delta}\right),$$

where $AP_L = Y/L$ and $Q = [\alpha_K K^{\delta} + \alpha_L L^{\delta}]^{1/\delta}$.

This simple model can be used to answer a crucial question: what happened to the marginal productivity of American mathematicians following the influx of Soviet immigrants who simultaneously increased the number of ideas *and* the number of workers. It is easy to show that:

(3) $$d\log MP_L = \phi\; d\log L + s_K\left(\frac{\phi + (1-\delta)s_L}{\phi + s_L}\right)(d\log K - d\log L)$$

where $s_L = \alpha L^{\delta}/Q^{\delta}$, or labor's share of output, and $s_K = 1 - s_L$. Equation (3) summarizes the opposing forces on the marginal product of mathematicians arising from the Soviet influx. On the one hand, the increase in the stock of ideas makes the pre-existing mathematicians more productive. On the other hand, the law of diminishing returns comes into play as a result of the increase in the number of mathematicians.

It is instructive to illustrate the implications of the model in terms of two alternative scenarios: the short run and the long run. By definition, resources *K* are fixed in the short run and are fully adjusted in the long run. If the academic market were competitive, additional resources would enter the market until the rate of return *r* to these resources is again equal to the world rate. The resulting change in the marginal product of mathematicians is then given by:

$$(4) \qquad d\log MP_L = \begin{cases} \left((\phi - s_k) + \dfrac{\delta s_L s_k}{\phi + s_L} \right) d\log L, & \text{if } dK = 0 \\ \\ \phi \; d\log L, & \text{if } d\log K = d\log L. \end{cases}$$

Suppose the production function is Cobb-Douglas, so that δ is zero. Equation (4) then collapses to $(\phi - s_K)\, d \log L$ in the short run. Resources are fixed and there is a race between the spillover effect and the law of diminishing returns. If the elasticity ϕ is sufficiently large, the marginal product of mathematicians rises.[10] Otherwise, the impact on the marginal product of mathematicians will be negative in the short run. Note that if the elasticity of substitution between labor and capital is less than one, the negative impact introduced by nonrival labor and resources gets larger, making it even more likely that the marginal product of mathematicians will decline.[11]

Equation (4) also shows that the marginal product of mathematicians must rise in the long run after the capital stock fully adjusts to the Soviet influx. In other words, more faculty slots are allocated to existing mathematics departments, perhaps new universities arise, and journals expand. If the institutions hiring mathematicians *act as if* they were competitive firms, the long-run adjustment in the capital stock would be given by $d \log K = d \log L$. In the words of Jones and Romer (2010): “In the long run, the benefits of a larger population which come from an increase in the stock of available ideas decisively dominates the negative effects of resource scarcity.”

However, it is unclear if the “long run” realistically captures adjustments in the market for mathematical knowledge. In an important sense, the old Keynesian quip that “in the long run we are all dead” is particularly appropriate in this context. After all, the Soviet

---

[10] The Cobb-Douglas functional form assumption provides a sense of how large ϕ must be for the spillovers to “win” the race. It is well known that the share of income accruing to capital (in the aggregate U.S. economy) is around 0.3. We do not know whether this estimate is relevant for the mathematics market, but it does suggest that the externalities elasticity must be substantial if the net effect on marginal product is to be positive.

[11] The elasticity of substitution between labor and capital is defined by $\sigma_{KL} = 1/(1 - \delta)$. It is difficult to imagine that resources and mathematicians are easily substitutable in the production of new theorems.

mathematicians, whose presence imparts new ideas that make everyone else more productive, will themselves eventually retire, and when they do presumably so will the external effects associated with their presence. Moreover, if short-run scarcity of capital causes existing mathematicians to stop publishing in the short-term after the Soviet influx, then academic markets are unlikely to make room for them to publish again in the future (academic labor markets typically have one-way doors). Thus, by the time capital adjusts, many careers of existing mathematicians are effectively over and it will be too late for the adjusted resources to help them reap the benefits of Soviet influence without incurring its costs. Finally, the long-run effect derived in equation (4) presumes a competitive academic market. Academic institutions, however, rarely behave like competitive firms. If nothing else, the tenure system imposes constraints on hiring and firing that distort the long-run allocation of resources. It would not be surprising, therefore, if an influx of world-class mathematicians reduces the productivity of the pre-existing workforce both in the short and long runs.

It is instructive to generalize this simple model to account for the fact that mathematics, as all academic disciplines, is composed of many fields, and these fields join together to form what we call "mathematical knowledge." Because the Soviet influx was particularly sizable in a relatively small number of fields, it is important to examine how the field distinction alters the lessons provided by this generic model.

The simplest way to model the existence of different fields of knowledge is to interpret *L* in equation (1) as the number of efficiency units supplied by the mathematics workforce. For simplicity, suppose there are two fields. It is convenient to use a nested CES framework to aggregate the contribution made by mathematicians in the various fields:

$$(5) \qquad L = \left(\theta_1 L_1^{\beta} + \theta_2 L_2^{\beta}\right)^{1/\beta},$$

It can be shown that the change in the marginal product of a mathematician in field *j* following a Soviet influx that potentially changed the quantities in all fields is given by:

$$(6)\quad d\log MP_{Lj} = \begin{cases} \left((\phi - s_k) + \dfrac{\delta s_L s_k}{\phi + s_L} \right) d\log L + (1-\beta)(d\log L - d\log L_j), & \text{if } dK = 0 \\ \\ \phi \, d\log L + (1-\beta)(d\log L - d\log L_j), & \text{if } d\log K = d\log L. \end{cases}$$

where $d \log L$ is the immigration-induced percent change in the total number of efficiency units, and $d \log L_j$ is the immigration-induced percent change in the number of mathematicians in field *j*.[12]

Equation (6) shows that the nested CES framework generates a redistribution effect in addition to the "race" discussed in the context of the homogeneous labor model. In particular, the marginal product of a mathematician in field *j* will now fall if the Soviets increased the supply of group *j* relative to their impact on the size of the other groups. Moreover, this impact is larger the smaller the elasticity of substitution between mathematicians in the two groups.[13] Put differently, a mathematician in group *j* could now suffer a short-run decline in his marginal product even if the external effects are relatively strong. In fact, the marginal product of the most affected mathematicians could decline even in the long run.

One problem with the nested CES model presented in equations (1) and (5) is that it assumes that the resources available to the mathematics community are commonly shared across mathematicians in all fields. This is probably not a realistic way of modeling the way that resources are allocated in academia. Consider instead the following alternative model in the context of two mathematical fields. Suppose that all mathematicians contribute to the stock of ideas that are useful in field *j* according to the Cobb-Douglas ideas production function:

---

[12] It is easy to show that the percent change in efficiency units is a weighted average of the percent change in the size of the various groups, where the weights are the output shares accruing to each of the groups.

[13] The elasticity of substitution is defined as $\sigma_{12} = 1/(1 - \beta)$.

(7) $$D_j = L_1^{\lambda_j} L_2^{1-\lambda_j}.$$

Suppose further that the field-specific production function for knowledge can be represented by a Cobb-Douglas production function:

(8) $$Y_j = D_j^{\phi_j} \left(K_j^{\alpha_j} L_j^{1-\alpha_j} \right),$$

This model allows the allocation of resources K to differ across fields. This assumption is probably more realistic in the context of academic departments, where resources and faculty slots are typically allocated to specific fields and purposes. The net impact of the Soviet influx on the marginal product of mathematicians in field 1 is given by:

(9) $$d\log MP_{L1} = \alpha_1 [d\log K_1 - d\log L_1] + \phi_1 [\lambda_1 d\log L_1 + (1-\lambda_1) d\log L_2].$$

Because the externality arises from the (weighted average of) ideas generated by mathematicians in all fields as specified in equation (7), the entry of Soviet émigrés in field 2 may have an external effect on the marginal product of mathematicians in field 1. As we will argue below, these cross-field spillovers are likely to be very small in mathematics.

One interesting implication of this model arises by noting that the total resources available to the entire mathematics field $K^* = K_1 + K_2$. Even if K^* were fixed in the short run, it may be possible to quickly reallocate resources across fields to take advantage of field-specific supply shocks. Suppose, for instance, that the Soviet influx increases the number of mathematicians in field 1 by $d \log L_1$, but leaves the number of mathematicians in field 2 unchanged. It would be optimal to move resources (e.g., faculty slots) from field 2 to field 1, even if total resources are fixed. This movement of resources would, of course, help attenuate the negative impact of the influx on the marginal product of mathematicians in group 1 while simultaneously lowering the marginal product of mathematicians in group 2.

As these models illustrate, the theoretical implications of the net productivity of a Soviet influx are ambiguous, even in the presence of human capital externalities. To our knowledge, no prior paper has directly documented whether an exogenous influx of new people and ideas into a *field of knowledge* had a net positive or negative impact on the rate of knowledge production of pre-existing workers.[14]

IV. Data and Descriptive Evidence

Data Sources:

We collected our data on Soviet and U.S. mathematics and related fields from two sources: Thomson Reuters' ISI Web of Science archive, and the American Mathematics Society (AMS) Math Reviews database. The ISI Web of Science contains research addresses and reprint addresses for many articles, allowing us to determine the geographic location and institutional affiliation of many authors at multiple points in time. The AMS has carefully researched the publication profiles of individual mathematicians in order to record unique author identifiers that precisely distinguish the papers of, for example, one John Smith from another John Smith. A staff of professional mathematicians at AMS has also recorded subject classification codes for each paper in its database: 64 subfields defined according to the 2-digit 2010 Mathematics Subject Classification (MSC2010), 104 subfields (3-digit MSC), or 610 subfields (5-digit MSC). With these classification codes we can observe how professional mathematicians have classified each paper's contributions. The empirical analysis in this paper uses the 2-digit coding throughout.

We merged the data from the ISI Web of Science with the data from the AMS paper-by-paper. This merging allows us to observe the location, field contributions, affiliations, and collaborations of uniquely identified individual mathematicians over the sample period.

---

[14] In related work, Hunt and Gauthier-Loiselle (2010) examine how the presence of immigrants in a state may affect the rate of patent creation in that state. We are interested in examining the impact of immigration into a narrowly defined scientific field on the rate of knowledge creation in that field. See also Furman, Kyle, Cockburn, and Henderson (2011) who conduct a related analysis of pharmaceutical research.

Scope and Accuracy of Data:

The ISI Web of Science is a publication database that records the titles, publication source, author names, references, and citations of millions of articles from thousands of journals worldwide. For many articles, the database records reprint addresses and research addresses for each author, as well as abstracts, key words, and funding information.

The database also contains complete citation information for each article in a primary set of 7,621 journals, selected to include all of the most important journals in each field. But articles in marginal journals also appear in the database if they either cite an article in the primary database or are cited by such an article. Thus, ISI Web of Science only misses articles that are not in important journals, are not cited by any articles in important journals, and do not themselves even cite any articles in important journals. This means the citation-adjusted weight of any missed articles must be very small, especially since ISI adds important journals to its main database by subfields, so even marginal subfields likely have the main field journals represented.

We purchased the records of all 1,179,787 articles in the primary ISI Web of Science database between 1970 and 2009 for the following categories: Mathematics, Applied Mathematics, Interdisciplinary Applications of Mathematics, Mathematical Physics, and Statistics & Probability. We also purchased all 1,921,587 articles referenced by these main articles, and all the 2,368,123 papers that cite these main articles. Not counting overlaps in these three categories, this amounted to 3,586,834 articles in Mathematics and related fields. For the reasons mentioned earlier, our database is a relatively complete record of all papers published in mathematics and related fields from 1970 until 2009. It is a useful but incomplete record of papers before 1970 (since it only includes these papers if they were cited by the post-1970 papers in the primary database).

AMS Math Reviews records the titles, publication sources, author names, references, and citations of over 2 million articles from 2,764 different journals and publication sources worldwide.[15] Importantly, the AMS Math Reviews team (located in Ann Arbor, MI)

---

[15] The 2009 AMS data does not contain the information for the complete calendar year. Hence it is not used in any of the empirical analysis reported in this paper.

painstakingly assigns the correct set of articles to each person (even in difficult cases such as identical names), tagging each author with a unique author identifier.[16]

As a result of the merging of these two data sets, we have constructed a unique data set that contains all the relevant information for every published article in mathematics and related fields for the period 1970-2008: the author(s), the affiliation, the location of the institution, the number of citations garnered by the article, and the mathematical subject field.

Merging Process:

We obtained special permission from the AMS to merge our ISI Web of Science papers by title, source, and author with the AMS database. The AMS set up their own merging site on a mirror in Germany, and we developed our own software to repeatedly query this mirror with papers from the ISI Web of Science database. We obtained 882,088 matches out of the 1,753,148 journal articles in the AMS database, or just slightly over a 50 percent match rate.[17] We checked a small random sample by hand and all matches were correct.

Summary Statistics:

Figure 5 illustrates the frequency distribution of publications and citations in the group of American mathematicians, conditional on the mathematician having at least one article published between 1978 and 2008. The sample of "predominantly" American mathematicians used in the figure consists of persons who published at least half of their

---

[16] As an example of this precision, we learned from personal communication with Professor Victor Kac of MIT that after he defected from the Soviet Union in 1977, he had to publish his work with his Soviet advisor (who remained in the USSR) under the Italian pseudonyms Gatti and Viniberghi, since otherwise the paper would need special permission from the authorities to be published abroad. Despite the difference in names between Gatti and Kac, the AMS database correctly lists the article under the unique author identifier for Victor Kac.

[17] There is a steady increase in the match rate over time, from about 40 percent for journal articles published in the late 1970s to around 75 percent for articles published after 2002. It is likely that the quality of the data improves over time, since all of the relevant information for more recent articles was electronically available at the time of publication. We are able to match over 90 percent of the papers for the most published mathematicians during the period (see the list in Table 3), suggesting that many of the missing articles are in marginal journals not covered by the ISI database. As a result, the citation-weighted count of the number of missing articles is likely to be very low.

pre-1990 work using an American affiliation. Not surprisingly, the mode number of publications for mathematicians is only one paper for the entire period. In fact, half of all mathematicians have 6 or fewer publications. At the same time, however, 4 percent of mathematicians published at least 50 papers during the period. Similarly, the mode number of citations for mathematicians who have published at least one paper is zero.[18]

Table 1 presents other summary statistics for the two samples of pre-existing American mathematicians that we will examine in the empirical analysis. In addition to the sample of predominantly American mathematicians, the table also provides statistics for "exclusively" American mathematicians, which consists of those who published all of their pre-1990 papers using an American affiliation. There are roughly 19,000 mathematicians in the first group, and 15,000 mathematicians in the second. Note that a relatively large fraction of both groups of mathematicians are relatively new entrants into the labor market as of the time of the Soviet influx: Nearly a third entered the labor market during the 1980s.

Figure 6 illustrates how the supply shock had a different quantitative impact in different fields. The identification strategy we will use in the next section requires that those fields where relatively few Soviets labored before 1990 witnessed a relatively small influx, while those fields where Soviets focused had a larger influx. In fact, the figure shows that the supply shock, as defined by the fraction of total papers published by Soviet émigrés, was very large for the fields emphasized by pre-shock Soviet mathematicians, and was very small for the fields that were dominated by American mathematicians. Our identification strategy also requires that the migration happened suddenly after the beginning of the fall of communism. This graph clearly demonstrates the sudden increase in the importance of Soviet émigrés in the American mathematics community around 1990.

V. Regression Results

The empirical analysis presented in this section will examine the extent to which any existing (or non-existing) "field overlap" between Soviet and American mathematicians altered the productivity of the latter after the émigrés arrived. In order to assess these

---

[18] The maximum number of publications by a mathematician during this period was 277 (Paul Erdös), and the maximum number of citations received was 31,962 (Edward Witten).

consequences, we first establish that the Soviet mathematicians who moved to the United States were a unique non-random sample that represented a highly positively selected group of the universe of Soviet mathematicians.

By definition, the universe of Soviet mathematicians is composed of those researchers who declared a USSR affiliation in at least one-third of their published research between 1970 and 1989. The one-third threshold is useful because a casual look at the time series of affiliations for Soviet mathematicians indicates that some of these persons would publish using a Soviet affiliation for some years, temporarily visit and publish from an East Bloc institution (for example, the University of Havana or Belgrade), and then return to their Soviet affiliation. Our data contains 2,212 unique individuals who can be identified as Soviet mathematicians by this definition.

From this universe of Soviet mathematicians, we define the émigrés as those who report an American affiliation at any point after 1992. It is not uncommon for Soviet émigrés to be affiliated with an American affiliation for a time, publish one or two papers using a foreign affiliation, and then continue publishing using their new American affiliation. By our definition, there are 272 Soviet émigrés in our data, or 12.9 percent of the pre-existing Soviet population.[19]

For each Soviet mathematician, we calculated the number of papers they published in both the 1970-1989 and the 1992-2008 periods.[20] To measure the selection that characterizes the émigrés *prior* to their migration, we use the entire sample of Soviet mathematicians to estimate a simple regression that relates two alternative dependent

---

[19] Our definition of an émigré requires the Soviet mathematician to have successfully obtained a job at an American institution and to have published from that institution. Of course, in academic job markets this number of émigrés does not reflect the actual number of people competing for the limited jobs and limited attention span of field leaders. Anecdotal evidence from the period suggests that a much larger group of Soviet mathematicians was competing for the same jobs as American mathematicians. In 1990, the New York Times reported: "Even more Soviet scholars want to come. American scientists say they are being peppered with letters and calls asking for invitations. 'It seems like I get a letter from the Soviet Union every two or three days,' said Dr. Ablowitz of the University of Colorado. In the meantime, Soviet scientists are traveling throughout the United States. 'I have run across a number of very distinguished Soviet mathematicians who have come here as visitors and spend their time going around the country and looking for a job,' Dr. Nathanson said." Therefore, in some sense, the entire Soviet mathematical community began engaging with the American mathematics community (through collaboration and competition for jobs) after the collapse of the Soviet Union, and our estimated effects result from both geographic and intellectual migration of new ideas into a pre-existing network of knowledge producers.

[20] We do not adjust for co-authorships when counting publications.

variables, the total number of papers published by a Soviet mathematician in 1970-1989 and the number of citations generated by those papers, to a dummy variable indicating if the Soviet mathematician is a future émigré. To control for the possibility that some of the mathematicians had fewer "active" years in the pre-collapse period, we include a variable indicating the year in which the mathematician published his first paper, effectively controlling for (potential) years of labor market experience.

The top row of Table 1 summarizes the evidence. It is evident that prior to their migration the future émigrés were significantly more productive both in terms of the number of papers published and in terms of the number of citations received. In particular, the future émigrés published around 3.5 more papers before 1990 than the Soviets who remained behind, and they received around 139 more citations. The data, therefore, unambiguously show that the sample of Soviet émigrés was positively selected from the universe of Soviet mathematicians.

We can examine the post-migration success of the Soviet émigrés by comparing their publication and citation record to those of America mathematicians after 1992. The sample of American mathematicians consists of persons who were predominantly affiliated with an American institution prior to 1990. For each Soviet émigré and each American mathematician, we calculated the total number of publications during the 1992-2008 period.

The bottom row of Table 1 summarizes the results from the comparison of these two groups. The results are striking. The Soviet mathematicians are substantially more productive than the Americans. During the 1992-2008 period, the typical Soviet mathematician published around 14 more papers than the typical American mathematician and those papers garnered 148 more citations. In short, the Soviet émigrés originated from the upper tail of the skill distribution in the Soviet Union and immediately moved into the upper tail of the skill distribution in the American mathematics community.

Perhaps the most telling way of summarizing the evidence is as follows: Soviet émigrés to the United States composed only 12.9 percent of the Soviet mathematics workforce, but accounted for 21.4 percent of all published papers in 1970-1989 and 38.2 percent of all citations. After migration, Soviet émigrés made up only 2.4 percent of the

American mathematics workforce, but accounted for 5.7 percent of all papers published in 1992-2008 and 4.2 percent of all citations.

The sizable quantitative difference in the publication rate of American and Soviet mathematicians is likely to have many consequences, particularly in those fields where the Soviets excelled. In the short run, for example, the total number of mathematics faculty jobs, as well as the fraction that deans and administrators allocate among the various subfields of mathematics, is likely constrained (perhaps by rent-seeking behavior within the university and hold-up problems among tenured faculty within departments) and relatively inelastic. The increasing presence of experienced and highly productive Soviet émigrés who are willing to accept junior-level jobs typically staffed by newly minted doctorates would then almost certainly have a "crowd-out" effect on the paid research jobs that American mathematicians would have otherwise filled if the Soviet Union had not collapsed and those Soviet mathematicians had remained behind the Iron Curtain. It is very likely that when fewer Americans are paid to produce research, fewer Americans will produce it, even in the context of the greatest sudden expansion in available mathematical knowledge in generations. There are also other sources of scarcity that will create crowd-out effects, such as the limited attention span of field leaders. There are insurmountable constraints on how much new knowledge a scientist can absorb, so that the location of the "marginal" article that will now get ignored shifts in the distribution of publications and reasonably good articles that would otherwise have generated some attention will remain unread and un-cited.

It is easy to demonstrate that the crowd-out effect was indeed a real phenomenon in the post-1992 American mathematics community. As we showed earlier, the field distribution of mathematicians in the Soviet Union prior to 1992 was quite different from the field distribution of American mathematicians. Our theoretical model implies that in the absence of sufficiently large externalities, the entry of an influx of Soviet mathematicians tends to hurt those American mathematicians who specialized in the Soviet fields. To test this theoretical implication, therefore, we need to calculate an index reflecting the field "overlap" between the publication record of *each* American mathematician and that of the Soviets.

We use 3 alternative indices to demonstrate the robustness of our empirical findings. The first index is simply a correlation coefficient estimated for each American mathematician in our data, estimated using publications between 1970 and 1989. Let a_{ij} be the share of those papers that mathematician i published in field j, and let s_j be the share of all Soviet papers published in field j during the same period. Our first index is simply the correlation coefficient ρ_i between these two vectors. Note that the correlation coefficient is calculated separately for each American mathematician.

A second index is designed to measure the "intensity" of the overlapping research interests. In particular, let S_i be the total number of papers written in field j in the Soviet Union between 1970 and 1989. If we calculate the dot product between the vector S and the vector of shares a_{ij} for each American mathematician, the resulting number gives the "effective" number of Soviet papers that the typical American mathematician would have written, weighted by Soviet preference. To construct an index that lies between zero and one, we divide the dot product by the number of papers published in the most populated Soviet field. Hence the "index of intensity" is given by:

$$(10) \qquad I_i = \frac{\sum_j a_{ij} S_j}{\max(S_j)}.$$

Note that this index will be equal to zero when the American mathematician publishes in fields where Soviets have never published and is equal to one when the American mathematician publishes exclusively in the field where the Soviets have done the most work.

Finally, we employ the commonly used "index of similarity" (Cutler and Glaeser, 2007) defined by:

$$(11) \qquad D_i = 1 - \frac{1}{2} \sum_j \left| a_{ij} - s_{ij} \right|.$$

The index of similarity equals one when there is a perfect overlap in the relative field distributions between Soviet and American mathematician *i*, and zero when there is total dissimilarity.[21]

Table 3 summarizes the value of the indices for the Top Ten mathematicians in our data, with the ranking being defined as the total number of publications between 1970 and 1989. Note that even among these superstars, there is a great deal of dispersion in the various indices. The correlation coefficient, for instance, ranges from zero to .7, while the index of similarity ranges from .06 to .41.

We also estimated the indices at the institution level (by aggregating all the mathematicians associated with a given institution). These institution-level indices are reported in the bottom panel of Table 3 for the Top Ten institutions, and they again show sizable dispersion in the areas of interest (relative to the Soviets) for the mathematicians in the various institutions. The correlation coefficient varies from .24 (for the University of Illinois to .67 (for the University of Minnesota).

Using these indices, it is easy to show that the Soviet mathematicians who moved to the United States joined institutions that tended to be highly ranked and had a pre-existing workforce whose research was more closely related to that of the Soviet émigrés.[22] The Soviet émigrés in our sample report a total of 136 American affiliations during the post-1992 period. We estimated regressions of whether a particular institution was “exposed” (i.e., hired a Soviet immigrant) on the various institution-level indices calculated earlier, as well as on a measure of institutional quality (given by *N*, the total number of papers published by all mathematicians at that institution between 1970 and 1989). The estimated regressions were:

$$(a) \qquad \underset{}{\mathit{Exposed}} = \underset{(.005)}{.017} + \underset{(.028)}{.164}\,\rho + \underset{(.00003)}{.001}\,N, \qquad R^2 = .369$$

---

[21] If the complement of the index of similarity (or 1 – D) were calculated for the total group of American mathematicians, its value would give the fraction of American mathematicians who must move across fields in order to ensure that American mathematicians have the same field distribution as the Soviets.

[22] There is a high degree of field overlap between the Soviet émigrés and the Soviets who remained behind. We calculated the various indices for these two groups. The correlation coefficient ρ is .758; the index of intensity is .622; and the index of similarity is .744.

$$(b) \qquad Exposed = .023 + .026\, I + .001\, N, \qquad R^2 = .359$$
$$(.007)\ (.024)\quad (.00003)$$

$$(c) \qquad Exposed = -\, .015 + .642\, D + .001\, N, \qquad R^2 = .423$$
$$(.005)\ (.041)\quad (.00004)$$

The sorting between the Soviet émigrés and American institutions was clearly not random. Instead, Soviet mathematicians tended to enter the best institutions. Mathematicians affiliated with a top-ranked institution, say MIT, published 1,675 mathematics papers during the 1970-1989 period, and that institution would have a much greater probability of landing one of the Soviet émigrés.

Moreover, holding institutional quality constant, the Soviets gravitated to those institutions that had a history of mathematical research in those fields favored by the Soviets. Each measure of the overlap index suggests that the greater the similarity between the fields that the Soviets studied and the fields that American mathematicians in that institution studied, the greater the probability that the institution hired a Soviet émigré (and this effect is strongly significant with two of the indices). For example, a shift in the correlation coefficient from zero to one increases the probability that the institution hired a Soviet mathematician by 16 percentage points.

Our theoretical model suggests the influx of Soviet mathematicians into the American market will have two effects on the pre-existing mathematics workforce: an increase in productivity through ideas spillover and a decrease in productivity through scarce rival resources. We will document the net impact summarized in equation (3) by estimating the relationship between a mathematician's marginal product and the extent to which that mathematician's pre-existing research interests overlapped with those of the Soviets. In the absence of strong external effects, the marginal product of a mathematician is more likely to fall the greater the overlap between the skill set of the mathematician and the skill set of the Soviets. Let $y_i(t)$ be a measure of the marginal product of mathematician i in year t. Consider the regression model:

$$(12) \qquad y_{it} = \phi_i + \phi_t + X_i(t)\,\gamma + \delta\, T + \theta\,(T \times Index_i) + \varepsilon_i(t),$$

where ϕ_i is a vector of individual fixed effects; ϕ_t is a vector of year fixed effects; X is a vector of standardizing variables that include the mathematician's years of work experience introduced as a quartic polynomial; T is a dummy variable indicating if the year for the particular observation is 1992 or beyond, and *Index* is one of the three alternative overlap indices defined earlier.[23] The standard errors are clustered at the individual level.

We use two alternative dependent variables in the analysis: the number of papers mathematician *i* published in a particular year, and the number of citations garnered by the papers published in that year. It is worth noting that the citations data is truncated because a paper may continue to be cited into the future (and some important papers may not have been recognized by 2008). This truncation, which is obviously serious for papers published in more recent years, implies that the citation results should be interpreted cautiously.

The identification strategy used by equation (12) can be easily described. We are examining how the research output of mathematician *i* changed after 1992, when Soviet and American mathematicians began to interact. The regression ignores the potential geographic proximity between the Soviet émigrés and the American mathematicians (e.g., they may be located in the same department or in the same city). As a result, the effect measured by the coefficient of the interaction between the overlap index and the post-1992 dummy variable, θ, is resulting from "intellectual proximity"—the possibility of substitution or complementarity in intellectual ideas.[24]

The coefficient θ captures whether the impact of the intellectual overlap between the Soviets and the specific American mathematician changed after 1992, when the Soviet ideas began to be disseminated to a wider audience in the United States and Soviet mathematicians began to publish their ideas in the same journals as American mathematicians. If the externalities arising from the entry of the highly skilled Soviets

---

[23] Of course, the regression does not include the value of the overlap index itself since the individual fixed effects subsume this variable.

[24] We also estimated models that examined if the measured impact was mitigated or exacerbated by geographic proximity. In particular, we defined a variable indicating if a particular mathematician was affiliated with an institution that eventually hired a Soviet émigré. The estimates of the parameter θ (not shown) were essentially independent of whether a particular mathematician had been exposed geographically to the Soviet influx.

specializing in a relatively small set of mathematical fields are weak or non-existent, we would expect that the American “intellectual competitors” of the Soviets would become less productive as a result of the influx.

The control group in our empirical exercise consists of American mathematicians who prior to 1990 did not specialize in the fields dominated by the Soviets. This control group not only experienced a smaller Soviet influx post-1992 (see Figure 6), but it also did not gain directly from the influx of new ideas. In particular, the exposure to Soviet ideas was unlikely to be useful to American mathematicians working in U.S. dominated fields: out of all the references cited by articles in the top 10 U.S. dominated fields, fewer than 3 percent of these references are to articles in Soviet-dominated fields.[25]

Table 4 reports the estimates of the coefficient θ obtained from a variety of different regression specifications. In addition to the two dependent variables and three measures of the overlap index, we also estimate the regression on the two alternative samples of predominantly and exclusively American mathematicians. Note that either definition of an “American” mathematician includes large numbers of foreign-born mathematicians. These foreign-born persons may have either migrated to the United States after publishing some initial work abroad, or first arrived in the United States as foreign students and stayed in the country after receiving the doctorate.[26]

Regardless of the various specification changes, the results reported in panel A of Table 4 are robust.[27] In particular, all of the overlap indices have a strong negative impact on the post-1992 productivity of American mathematicians. It is instructive to give a numerical example to emphasize that the quantitative impact on the number of publications is numerically important. In terms of the correlation coefficient, for instance, an increase from $\rho = 0$ to $\rho = 1$ (so that the pre-1990 field distribution shifts from one

---

[25] In the context of our model in equation (9), this implies that the parameter λ is close to one. In other words, the change in marginal product of the mathematicians in the treatment group can be estimated by comparing the publication records of the control and treatment groups over time.

[26] In fact, 47.5 percent of the papers published by predominantly American mathematicians between 1970 and 1989 used a foreign affiliation.

[27] In addition to the influx of Soviet researcher in the 1990s, there was also an influx of non-Soviet Eastern Bloc researchers during the early 1990s. We have re-estimated the basic models using the field distribution for all mathematicians in the East Bloc and the Soviet Union. The results are very similar to what we report in the text.

where the American mathematician's work is uncorrelated with that of the Soviets to one where the mathematician's research interests perfectly overlap) reduces the number of publications in any particular year by 0.14, or roughly 1.4 fewer published papers per decade.

It is easy to illustrate graphically the impact of the Soviet supply shock on the productivity of American mathematicians whose research agenda overlapped that of the Soviets. Consider the sample of American mathematicians who were always affiliated with an American institution prior to 1990. Suppose we classify mathematicians whose index of similarity is in the upper quartile as "highly exposed," while the group in the bottom quartile has low exposure. As shown in the top panel of Figure 7, the trends in the average number of papers published by the average mathematician in each of these two groups are revealing. Both of the groups have slight downward trends in the number of papers prior to 1990. After 1990, however, there is a precipitous decline in the publication rate of the group whose research agenda overlaps most with the Soviets.[28] The trends in the adjusted number of papers (after removing experience and year effects) also show a continuation of previous trends for the less exposed mathematicians, and a reversal of pre-1990 trends for the highly exposed mathematicians.

Table 4 also presents the estimates of the coefficient θ when the dependent variable is the number of citations (or roughly the number of publications weighted by quality). The evidence again suggests a strong decline in the number of citations written by American mathematicians with research interests that greatly overlap those of the Soviets. In other words, not only are the competing mathematicians producing fewer papers, but also the work they produce is of "lower quality" in the sense that it generates many fewer citations. A one-unit shift in the correlation coefficient would reduce the number by over 22 citations per year.

The next two panels of Table 4 address the question of whether these effects persist in the long run. In panel B, we restrict the analysis to the years 1978-1999, so that the regressions measure the short-run productivity effect, while panel C only includes the

---

[28] The data point for the year 2008 is off the linear trend. However, this likely arises because the AMS data is not complete for the year 2008, and this statistical artifact disappears after controlling for year fixed effects.

years 1978-1989 and 2000-2008, so that the productivity impact is measured roughly one to two decades after the Soviet influx occurred. The perhaps surprising finding in the table is that, if anything, the long run effects are larger than the short-run effects (although the difference is sometimes not statistically significant). As we noted earlier, it is very difficult for academics to reenter the publications market once they have taken some years off from successful active research. In academia, the short run is the long run.[29]

One potential problem with these regression results is that although they include period fixed effects to net out any trends in total output in the mathematics discipline, there may be field-specific time trends that may be spuriously correlated with the Soviet influx. To determine if the results are sensitive to the inclusion of these field-related differences, we re-estimated the regression model in a sample constructed so that a particular observation represents an author-field-year permutation. In other words, our data now consists of the publication history of each American mathematician in each of the 63 fields in mathematics for each year in our sample period, 1978-2008. Of course, many of the values in the dependent variable (e.g., the number of publications or the number of citations) in this regression will be zero simply because few mathematicians publish anything at all in a particular year, and fewer still publish a paper in more than one field.

The regression model is now given by:

$$y_{ij}(t) = \phi_i + \phi_j + \phi_t + (\phi_j \times \phi_t) + X_i(t)\,\gamma + \delta\, T_i + \theta\,(T \times Index_i) + \varepsilon_i(t), \tag{13}$$

---

[29] In the long run, we would expect resources to be moved to the affected fields. It is unclear, however, that such adjustments took place. In fact, federal obligations to universities and colleges for research in mathematics declined slightly (in real dollars) during the 1990s; see U.S. National Science Foundation (2004, Table 1a). Anecdotal evidence also suggests that it was difficult to obtain additional resources for hiring Soviet mathematicians. Personal communication with Professor Arthur Jaffe provides two such examples. Professor Jaffe was Chair of the Harvard mathematics department at the time and requested resources to hire two Soviet mathematicians classified as "Targets of Opportunity." The Dean of the College denied the request, although Harvard President Derek Bok belatedly reversed that decision. Similarly, Jaffe and Bok contacted 54 foundations requesting assistance to fund the transition of Soviet mathematicians into the American mathematics community. All of the funding requests were denied.

where $y_{ij}(t)$ is a measure of the marginal product of mathematician *i* in field *j* at time *t*; ϕ_j is a vector of field fixed effects; and $(\phi_j \times \phi_t)$ represents all possible interactions between the field and period fixed effects.

The bottom panel of Table 4 summarizes the regression results. It is evident that the coefficient θ remains negative and significant even after controlling for the field-year fixed effects, regardless of the overlap index used. The numerical magnitude of the coefficient is smaller than that of the analogous coefficient reported in Panel A, but this is simply a mechanical effect because the total impact on publications is now divided among the 63 fields that make up the mathematics discipline. Once we scale the coefficient properly, the regression coefficients reported in Panel D are, in fact, quite similar to those in Panel A. The regression results, therefore, strongly indicate that field-specific trends do not account for the post-1992 break in the evolution of publications for mathematicians whose research interests greatly overlapped with that of the Soviet émigrés.[30]

In terms of our theoretical model, the evidence shows that for the typical American mathematician the spillover effects are swamped by the diminishing returns associated with nonrival resources. Nevertheless, there may exist a select group of mathematicians for whom the externalities dominate and lead to an increase in productivity. The anecdotal evidence discussed earlier suggests that likely candidates are the mathematicians whose research strongly overlapped with the Soviet research agenda and who happened to be affiliated with the high-ranked institutions that eventually hired the Soviet émigrés.

To isolate this group, we first create a variable measuring a high degree of research overlap. The "*overlap*" dummy variable indicates if a mathematician's pre-1990 output placed him in the top decile of the distribution of the correlation coefficient ρ. We then estimated our generic model using this binary measure of research similarity. As expected, the coefficient reported in the first column of the top panel of Table 5 shows a negative impact of research similarity on the post-1992 productivity of mathematicians.

---

[30] We have also estimated regressions that control for pre-existing trends in the author-year regressions by including linear trends for the upper and bottom quartile of the overlap index. The results are similar to those reported in the text.

The next 4 columns expand the analysis by introducing a three-way interaction between the dummy variable indicating the post-1992 period (*T*), the binary overlap variable, and a dummy indicating if the mathematician's last observed pre-1990 institution was highly ranked.[31] The table shows that the three-way interaction is negative, though usually insignificant, regardless of how a "highly ranked" institution is defined.

The bottom two panels of Table 5 report a more interesting exercise. In particular, we classify mathematicians according to whether they are "tenured" or not. Panel B uses the sample of mathematicians whose first paper was published between 1980 and 1989—hence they were relatively junior faculty and many of them would not have yet attained tenure by the end of the decade. In contrast, panel C uses the sample of mathematicians whose first publication was before 1980 so that they had at least a decade of work experience prior to the Soviet influx and were likely high-level Associate Professors or Full Professors by the time of the influx.

The separation of the mathematicians into these two groups reveals that it is the older mathematicians, already tenured prior to the arrival of the Soviet émigrés, who gained from the influx. Moreover, it is the tenured mathematicians at the *Top 3 or Top 5* institutions who gained. The three-way interaction variable is positive and significant only for these groups, and there's a marked downward trend in the productivity gain as one includes more and more institutions among the "highly ranked." In contrast, the relevant coefficients for the junior faculty are consistently negative and significant, and there is a marked rise in the productivity loss as the list of highly ranked institutions becomes more selective.

Prior to the Soviet influx, there were 47 mathematicians affiliated with Top 3 and 76 mathematicians affiliated with Top 5 institutions who had a strong overlap with the Soviet research agenda. It is this select group of mathematicians who experienced an increase in productivity after 1992.[32] In contrast, the pre-existing junior faculty at those high-ranking

---

[31] The ranking is determined by the total number of papers published by the institution over the entire period.

[32] We also estimated the regression model using the number of citations as the dependent variable. Interestingly, the coefficient of the three-way interaction variable is numerically and statistically close to zero in all specifications. There's no evidence, therefore, that the "quality" of the output of this highly select group of mathematicians was improved by the externality.

institutions whose research strongly overlapped with that of the Soviets experienced a substantial decline in their marginal product.

The individual-level data can also be used to illustrate a crucial impact of the Soviet supply shock: its effect on the probability that an American mathematician remains active (i.e., publishing) in the profession. To determine the impact of the Soviet émigrés on the survival rate of American mathematicians we constructed a variable indicating whether the mathematician has "retired" by setting it equal to zero in every year prior to the last year in which we observe a publication for mathematician *i*. We then omit all post-retirement observations for the particular mathematician (in other words, we exclude the years after a particular mathematician published his last article). We used this form of survival data to estimate probit regressions where the dependent variable is the probability of retirement.

Table 6 reports the marginal effects on the probability of retirement of the interaction variable between the post-1992 period and the various overlap indices. The first row of the table reports the results for the sample of mathematicians who were predominantly affiliated with American institutions prior to 1990. We find a positive and significant impact of the overlap index on the retirement probability using both the correlation coefficient and the index of similarity. The estimate of θ when the index is measured by the correlation coefficient indicates that a mathematician whose prior research was perfectly correlated with that of the Soviets has about a one percent higher probability of retirement post-1992 than a mathematician whose work was uncorrelated. Over a decade, therefore, those mathematicians had a 10 percent higher probability of "retirement."

The remaining rows of the table show that the impact of the Soviet influx on the probability of retirement does not fall evenly on all pre-existing American mathematicians. The impact of an overlapping research agenda with the incoming Soviets is far larger in the sample of younger (as measured by the year of first publication) and likely untenured mathematicians. For mathematicians who first published a paper in the 1980s, for example, the estimate of θ when the overlap index is measured by the correlation coefficient is between 0.03 and 0.05, so that the probability of retirement in any given year after 1992 rises by 3 to 5 percentage points, or 30 to 50 percent over a decade. In other words, the

Soviet supply shock had a particular detrimental effect on the long-run career prospects of young (and untenured) mathematicians, greatly increasing their probability of retirement from the publishing business. Of course, this evidence cannot distinguish between the possibility that the mathematician indeed stopped writing publishable articles or was simply unable to find an outlet for his research in the face of increasing competition by the highly skilled Soviet mathematicians.

In fact, the Soviet influx not only affected the timing of the end of an active research career, but also the mathematician's institutional placement at the beginning. Figure 3 demonstrated a sudden decrease in the probability of newly minted Ph.D.'s obtaining a position at a research institution at the time of the Soviet influx. In order to determine if this effect was more common among mathematicians who were in direct competition with the arriving Soviet émigrés, we need a proxy for the "Soviet-ness" of the dissertation topic. We obviously cannot calculate a brand new mathematician's overlap index, so we instead use the field of the first published paper as a proxy.[33] Note that 52 percent of pre-1990 publications in the Soviet Union occurred in only eight fields.[34] We define these fields as "high exposure" fields. In contrast, only 10 percent of publications in the Soviet Union occurred in the 34 least popular fields. We call these fields "low exposure" fields. We classify the remaining 21 fields as "medium exposure" fields. We use these definitions of high-, medium-, and low-exposure fields to categorize an American mathematician's first paper. Our measure of institutional quality is based on the total number of pre-1990 papers published by the institutional affiliation that the mathematician reported in his first published paper. We then estimate regressions that relate the quality of the mathematician's initial placement to variables indicating if that first paper was in one of the dominant Soviet fields.

Table 7 reports the regressions of various measures of institutional quality on the level of exposure of the newly minted mathematician to Soviet competition. The coefficient

---

[33] We are, of course, assuming that the first published paper is related to the mathematician's doctoral dissertation.

[34] These eight fields are partial differential equations, quantum theory, global analysis and analysis of manifolds, operator theory, probability theory and stochastic processes, ordinary differential equations, statistical mechanics and structure of matter, and functions of a complex variable.

of interest, of course, is the interaction between the variable indicating if the paper is in a high-exposure field and the post-1992 dummy. This variable has a significant negative effect. In other words, mathematicians who wrote their dissertation on topics similar to those that interested the Soviets experienced a substantial decline in the quality of their first academic placement after 1992. If we use the total number of pre-1990 papers to rank institutions, a mathematician in a high-exposure field ends up in an institution that is ranked 47 slots lower in the academic pecking order after the Soviet influx began.

As we emphasized earlier, short-run effects in academia can be persistent. The placement of a newly minted mathematician into a lower-quality institution will surely have many adverse ramifications for the mathematician's future productivity (as measured by scholarly output).

VI. Behavioral Responses

The evidence strongly suggests that there should be behavioral responses in the hiring and firing decisions of American institutions as a result of the Soviet influx, as well as in the research agendas of American mathematicians. The existence of these behavioral responses has been largely ignored in the literature that attempts to document the existence of human capital externalities. We now examine a number of related variables that help to document how the sorting of American mathematicians to institutions adjusted to the Soviet influx, and how some American mathematicians began to pursue different research topics to avoid Soviet competition in the "publications market".

We begin by examining the determinants of the rate of institutional mobility of American mathematicians. In particular, we define a variable indicating if the modal institution in the post-1992 period (after the Soviet influx) was different from the last institutional affiliation observed for a particular mathematician prior to 1990. By construction, our sample consists of mathematicians whose last reported affiliation prior to 1990 was an American institution. We then estimated linear probability models (at the individual level) that examine the separation probability of a mathematician using variables that presumably control for differences in individual quality and in the monetary incentives for such mobility. In particular, we hold constant the total number of papers the mathematician published prior to 1990, a measure of institutional quality, and years of work experience.[35]

The regression includes two variables designed to capture the impact of the Soviet influx. The first is the overlap index that describes the similarity in research interests (before the supply shock) between the particular American mathematician and Soviet mathematicians. The second is a dummy variable indicating whether the institution employing the mathematician just *before* the shock was, in fact, subsequently "exposed" to Soviet émigrés (i.e., if the institution itself hired a Soviet immigrant after 1992). The impact of this variable, of course, measures if the individual mathematician was directly affected by the Soviet influx (rather than through a diffused market-wide effect).

The regression coefficients estimated in the sample of mathematicians employed by an American institution prior to the Soviet shock are reported in the top panel of Table 8.[36] The separation rate is strongly related both to geographic proximity and to intellectual proximity. In particular, the separation rate rises by about 2 percentage points if the mathematician was initially employed in an institution that eventually hired a Soviet émigré. Similarly, regardless of how the overlap index is measured, there is a strong positive correlation between the index and the separation rate. An increase in the correlation coefficient from zero to one increases the separation rate by 14 percentage points. In other words, holding direct exposure constant, those mathematicians whose pre-1990 research greatly overlapped with that of the Soviets found themselves in volatile jobs and ended up moving to other institutions at dramatically higher rates

The bottom two panels of the table replicate the analysis for the samples of "young" and "older" mathematicians. It is notable that the separation rate of younger (i.e., untenured) faculty is much more sensitive to direct exposure to the Soviet influx. In particular, the probability of moving increases by about 4 percent for a young

---

[35] The measure of institutional quality is given by the rank of the institution during the 1970-89 period, as defined by the ranking in terms of total articles published during that period. Years of work experience are defined as time elapsed since the publication of a mathematician's first article.

[36] Although we do not show the coefficients of the other variables in the regression, they mirror findings in the labor economics literature. The higher the quality of the mathematician (as measured by the number of pre-1990 publications), the lower the separation rate. This finding is analogous to the negative correlation between the probability of job separation and a worker's wage. In addition, the separation rate is lower for older mathematicians and higher for mathematicians initially employed by low-ranked institutions.

mathematician who was employed in an institution that eventually hired a Soviet émigré, and is essentially unaffected for an older mathematician.

The separation regressions, of course, raise a number of interesting questions. Most important: Where did the mathematicians who were "displaced" by the Soviet influx end up? We calculated a measure of change in institutional quality by adding the total number of papers published by mathematicians affiliated with a particular institution in the pre-1990 period. Since we can identify the last institution employing the mathematician prior to the supply shock as well as the modal institution employing him after the influx, we can construct a measure of the change in institutional quality for movers. Let N_0 be the number of papers published by the institution employing the mathematician before 1990, and N_1 be the number of papers (measured pre-1990) employing him after the influx. We define a variable giving the change in institutional quality as $\log(N_1/N_0)$.

Table 8 also reports the regression results indicating how the exposure variable and the overlap index affect the quality of the institution employing American mathematicians in the sample of movers. Note that the exposure variable is defined in terms of the institution employing the mathematician *prior* to the Soviet supply shock. The change in institutional quality is strongly and negatively related to both geographic and intellectual proximity. The impact of geographic proximity is particularly large. Conditional on moving, an American mathematician initially employed by one of the institutions that ended up hiring Soviet émigrés moved to an institution that had published 93 percent fewer papers in the pre-1990 period!

Although the numerical magnitude of this effect seems implausible, it is easy to explain. The total mathematical output produced by institutions falls precipitously once we move from the top-ranked institutions to even those ranked 100 or 200 in our index of quality. The typical career path for an academic often exhibits a downward drop in institutional quality if the individual is not granted tenure at a top-ranked institution. The Soviet mathematicians were mainly employed in high-ranked institutions. The American mathematician's move from those institutions was likely not voluntary, and inevitably resulted in the mathematician moving to a lower-tier institution where the number of total publications was substantially lower.

We can summarize the nature of the move by calculating the average rank of the institution employing the exposed American mathematician before and after the influx. We ranked all institutions according to the number of papers published in 1970-1989. The average American mathematician who moved after the institution hired a Soviet émigré was initially affiliated with an institution that ranked 33rd in our list, and that institution had published around 560 papers in the pre-influx period. After the move, the mathematician ended up in an institution ranked 395 and that institution had only published 15 papers in the pre-1990 period.[37]

Of course, although some mathematicians ended up moving to lower-quality institutions, others ended up staying at the exposed institutions. Not surprisingly, some of these mathematicians took advantage of the changed opportunity set by establishing co-authoring relationships with the Soviet émigrés. Our data allows us to construct a variable indicating if an American mathematician had ever (i.e., after 1992) co-authored a paper with any of the Soviet émigrés we identified. The mean probability of co-authorship is 5.9 percent at the exposed institutions and 3.6 percent at the unexposed institutions.

We again estimated regressions to examine the determinants of this type of co-authorship and these regressions are reported in Table 9. Not surprisingly, the mathematicians taking advantage of the Soviet skills were those who are affiliated with the exposed institutions and who worked in research fields (prior to the influx) that greatly overlapped with the Soviet research agenda. Moreover, note that the impact of direct exposure to Soviet émigrés on co-authorship rates is statistically significant only for the older, more established mathematicians.

While some American mathematicians responded to the influx by beginning a co-authoring relationship with the Soviet émigrés, others responded in an opposite direction—by switching research interests *away* from the Soviet fields. From a long-run perspective the various indices of overlap that we have used throughout the analysis are endogenous, as the mathematician changes research interests in response to supply and demand, as well as to age-related changes in mathematical skills. Up to this point, we have

---

[37] We also interacted the exposure variable with the overlap index, but these interactions, although positive, are usually insignificant.

used an exogenous measure of the overlap index by calculating it based on the publication record of the American mathematician *prior* to the Soviet influx. We now construct an analogous index based on the mathematician's publication record between 1992 and 2008. The post-influx value of the overlap index is endogenous because some mathematicians will choose different research agendas after the arrival of the Soviet émigrés.

The regressions reported in the top panel of Table 10 show the relationship between the post-1992 overlap indices and some of the variables in our analysis for the sample of pre-existing American mathematicians. Not surprisingly, research interests tend to be persistent: those mathematicians whose research interests overlapped with those of the Soviets before 1990 will have research interests that overlap even after the influx. Note, however, that the coefficient of the lagged index is uniformly less than one, ranging between .5 and .7. More interesting, direct exposure to Soviet émigrés has only a weak (negative) effect on the overlap index. The weakness of this effect is not surprising since the group of pre-existing mathematicians will inevitably have a much more difficult time retooling their skills and adjusting to the Soviet influx. As we noted earlier, it is rare for mathematicians to publish in more than three or four fields during their professional lifetime.

The bottom panel of the table looks at the larger universe of mathematicians by including both the pre-existing mathematicians, as well as the mathematicians who entered the market after 1992. We include a "new entrant" variable in this regression to show the significant difference in the research agendas chosen by the new and the older cohorts of mathematicians. To adjust for age-related differences in research choices, the dependent variable is the post-1992 overlap index for the new entrants and the pre-1990 index for the pre-existing mathematicians.[38] Hence we are effectively comparing the research choices made by mathematicians who entered the mathematics market before and after the influx at a similar stage, i.e., the beginning phase of their career. The dummy variable indicating whether the mathematician is a new entrant has a strong negative coefficient, implying that the post-influx entrants are much less likely to choose a research agenda that overlaps with

---

[38] To further ensure that the comparison of the two groups is at the same point in their research life cycle, the pre-existing mathematicians in this regression entered the market (i.e. had their first publication) between 1970 and 1989.

that of the Soviets. It is also of interest that the move away from Soviet research interests is even stronger if the new mathematician is affiliated with an institution that was directly exposed to the Soviet influx.

In sum, the adverse productivity effects that were documented in the previous section by examining the post-1992 break in the publication trends of American mathematicians led to strong behavioral reactions in that community. The Soviet influx altered the sorting between mathematicians and institutions, with the mathematicians working on Soviet-style research moving to lower quality institutions, and with new entrants in the mathematics market choosing research topics that were less likely to compete directly with the Soviet research agenda. By shying away from these research topics, the new mathematicians were also less likely to take advantage of the incoming Soviet ideas.

VII. Aggregate Effects

Moving from individual-level effects to aggregate effects, we now consider whether the total number of papers published by the cohort of American mathematicians who were active before 1990 increased or decreased when we take into account the contribution of Soviet émigrés to the United States.

Suppose we divide this pre-existing group of American mathematicians into three categories: those with a pre-1990 research agenda that was highly correlated with that of the Soviets, those with an agenda that was almost uncorrelated with that of the Soviets, and everyone else.[39] For each of these three groups of mathematicians we estimated a regression that describes the mathematician's age-product profile using *only* the 1978-1989 data. In particular, we estimated a regression of the number of papers published by a particular mathematician in a given calendar year on years of experience (introduced as a quartic polynomial) and on a vector of individual fixed effects. We used this regression to predict the post-1992 product for each pre-existing mathematician in each of the three groups. We then added the total number of predicted publications across years and across

---

[39] Specifically, we use the distribution of the correlation coefficient across mathematicians to define the first quartile as the "unexposed" group, the fourth quartile as the "exposed" group, and the middle 50 percent as the residual group.

all mathematicians in each group. Table 11 reports the predicted total number of papers and contrasts these predictions with the actual output of the pre-existing workforce.

It turns out that the predicted total number of papers over the 1992-2008 period is almost identical to what was actually observed for two of the three groups of mathematicians. Consider, for instance, the group of mathematicians in the bottom quartile of the distribution of the correlation coefficient. Based solely on the path of the pre-1990 age-product profile, we would expect this group to publish 1,341 papers over the 1992-2008 period. In fact, this group published 1,376 papers. Therefore, we tend to slightly under predict the number of publications for mathematicians whose work was basically uncorrelated with that of the Soviets. It is worth emphasizing that this prediction exercise reveals that there is little beneficial spillover for these mathematicians from being exposed to Soviet-style mathematics. Similarly, for the group whose correlation coefficient lies between the 25th and 75th percentile, we predict 2,635 papers in the post-1992 period, while they actually publish 2.620 papers. The gap between observed and actual is not statistically significant for either of these groups.

It is in the group of mathematicians who are in the top quartile of the distribution of correlation coefficients that our prediction is far off from the actual value. In particular, based on their pre-1990 age-product profile we would have expected 2,845 published papers between 1992 and 2008. In fact, this group published only 2,033 papers, a "loss" of over 800 publications.

If we add up the total number of predicted and actual papers across the three groups, our regression exercise over-predicts the actual output by about 800 papers (or 13 percent). Table 11 also shows that Soviet mathematicians contributed 343 papers to the post-1992 output, so that the contribution of Soviet mathematicians "explains" about half of the gap.

This exercise raises a number of interesting issues regarding which capital inputs to academic production are not adjusting quickly enough to make room for the externalities. One could argue, for example, that the total number of published papers in mathematics over any reasonable time period is essentially fixed. As a result, the fact that the Soviet émigrés take up scarce journal space must inevitably lead to a decline in the pages allocated to pre-existing American mathematicians. The conjecture of a "fixed pie," however, does not accurately represent the evolution of the mathematics market over the past four decades. Figure 8 shows the trend in the (log) number of mathematics papers published since 1970. Remarkably, the log number of papers has been growing almost linearly since that time, with a coefficient of .058; so an increase in the number and size of journals each year has brought about an increase in the size of the mathematics market of about 5.8 percent per year for the last four decades.

Furthermore, the cohort of mathematicians that we examine does not represent the whole American mathematics community, and could in theory have taken up a larger share of the "publications pie" if it had strongly benefitted from the insights of the Soviet newcomers. In other words, it is unlikely that the influx of Soviet mathematicians mechanically prevented some American mathematicians from publishing their work if drafts of those "missing" papers had, in fact, been circulating in the ever-widening publications market.[40]

A more likely explanation is that the papers that are "missing" from our empirical exercise are papers that would have been written, but never were. The nature of the academic market can help us understand how this situation may arise. First of all, the American Mathematical Society's Annual Surveys report surprisingly small changes from year-to-year in the total number of mathematicians working in U.S. research institutions during our sample period.[41] With the total number of research jobs unresponsive to the opportunity created by the Soviet influx, the entry of experienced Soviet mathematicians had particularly adverse effects on untenured mathematics faculty. As we have shown, some of those faculty moved to lower-ranked institutions and some of them moved out of the academic market altogether. Regardless of their final placement, many of these mathematicians found it increasingly difficult to do the type of research that leads to

---

[40] Furthermore, the fact that the total number of papers was not fixed implies that the total number of citations was not fixed either. Kaiser (2011) has extensively documented how changes in the number of people pursuing physics research are related to changes in the total amount of physics papers published. It appears that, as Orley Ashenfelter is fond of saying, every paper has a home.

[41] For instance, between the 1990-91 and 1994-95 academic years, the total employment of doctoral full-time faculty in mathematics departments of Ph.D.-granting institutions increased from 6,008 to 6,147. The number of such faculty at the 82 institutions with a Conference Board of Associated Research Councils Rank of I or II (i.e., those institutions that produce the bulk of the best research) actually decreased from 3,740 to 3,613. It is also interesting that during this time period the total employment in statistics departments remained roughly constant, changing from 1,123 to 1,099.

publishable output, either because of an increased teaching load, fewer networking possibilities, or because the responsibilities of a job as a "quant" in Wall Street limits the kind of effort required to develop publishable material. Regardless of the cause, the affected untenured mathematicians not only began to publish less, but eventually ceased publishing altogether. Put differently, the impact of the Russian influx on the productivity of these mathematicians was not a marginal impact on the number of papers per year, but instead helped to drive out some mathematicians from the publishing community.

One can argue that the substitution of the output of these "marginal" mathematicians for that of highly able Soviets was beneficial for mathematics. In this perspective, the published work of the Soviets has a higher intrinsic value for mathematics than that of the marginal mathematicians who were displaced. However, the academic market is ripe with market failures that prevent it from working this neatly. It is likely that the "marginal" mathematicians who were crowded out were untenured. Many older and tenured mathematicians who were unaffected by the Soviet influx were probably less productive than the displaced untenured mathematicians.

It is, therefore, difficult to find convincing quantitative evidence that there was an improvement in the overall "output" of either the pre-existing American workforce or of that community combined with the Soviet émigrés. Furthermore, since the new cohort of American mathematicians were less likely to produce their initial work in Soviet-style research, it is unlikely that net positive production externalities are having much of an impact on the new cohort either.

It is possible, however, that the "quality" of mathematics improved in a broader sense. For instance, experimental physicists may now be citing and using newly available Soviet work on quantum theory in their own research. We should then expect to see further ripples in the pond in the sense that the exposure to Soviet-style mathematics has affected the product and quality of work of scientists in fields other than mathematics.

We examined the extent to which scientists working outside mathematics now cite the work of Soviet mathematicians and Soviet émigrés, or of mathematics in general. This exercise revealed two key facts: First, American papers outside mathematics do cite mathematical papers. There were 256 thousand such citations in 1978, and this number rose to 1.8 million by 2007. Second, citations to Soviet mathematical research are rare throughout the entire period. These citations accounted for three-hundredths of one percent of the total in 1978 and seven-hundredths of one percent in 2007. It seems implausible, therefore, that Soviet mathematical research could have formed the basis of large improvements in American non-mathematical fields.

Finally, there are many remaining areas outside of mathematics where net positive externalities from the influx of Soviet researchers and ideas might play an important role. For instance it is possible that by helping turn some untenured American mathematicians into Wall Street "quants" who revolutionized finance with their new ideas, the Soviet influx *indirectly* produced a new wave of positive externalities (i.e., externalities directly produced by the pre-existing American mathematicians themselves). However, we are not capable of quantitatively measuring these indirect externalities because they require comparing the importance of potentially groundbreaking science that was never written with the importance of influential financial instruments whose value is still being debated. Though the realm of potential externalities in ever more distant sectors is infinite, one can characterize our findings in terms of a simple metaphor: when one drops a stone in a pond, if the first ripple and the second ripple are small, then the remaining ripples are unlikely to be larger, even though in unusual circumstances they may be.

VIII. Summary

Modern models of economic growth assume that there are strong human capital externalities that could lead to increasing returns. It is not surprising then that there is a consensus among policy makers that increases in the supply of a highly skilled workforce will increase the productivity of the pre-existing workforce, and lead to a substantial increase in national wealth.

This paper examines the impact of the influx of renowned Soviet mathematicians into the American mathematics community. In the period between the establishment and fall of communism, Soviet mathematics developed in an insular fashion and along different specializations than American mathematics. After the collapse of the Soviet Union, several hundred world-class Soviet mathematicians emigrated to other countries, with a large percentage settling in the United States. As a result, some fields in the American mathematics community experienced a flood of new mathematicians, theorems, and ideas, while other fields received few Soviet mathematicians and gained few potential insights.

We constructed a data set that contains information on the authorship, subject, and affiliation of papers published in mathematics over the past 40 years. These data allow us to document the location, affiliation, and complete publication and citation records of mathematicians who were active in either the Soviet Union or the United States for the past few decades.

Our empirical analysis shows that almost all American mathematicians whose research agenda most overlapped with that of the Soviets experienced a reduction in productivity after the entry of Soviet émigrés. Based solely on the pre-Soviet age-output profile of American mathematicians, we can predict almost perfectly the actual post-1992 output of mathematicians whose work had little overlap with the Soviet research agenda, while the actual output of mathematicians whose work should have allowed them to benefit from Soviet ideas is far below what would have expected given their pre-1990 marginal product.

There is also strong evidence of a large array of behavioral consequences among American mathematicians: those mathematicians whose research overlapped with that of the Soviets became much more likely to switch institutions; the switch entailed a move to a much lower quality institution; many of these American mathematicians ceased publishing relatively early in their career; and new entrants into the mathematics community were less likely to pursue a research agenda in Soviet-dominated fields than the previous cohort had been.

This evidence is not consistent with the conventional human capital externalities story. We do not believe this inconsistency arises because American mathematicians did not gain new ideas from the Soviet influx. Rather, we believe that this evidence suggests that there may be surprisingly resilient constraints on the number of academic jobs available and funding opportunities for a mathematician’s research agenda. The gains from knowledge spillovers might have been much greater had policy makers or academic institutions correspondingly adjusted the amount of resources allocated to the affected fields.

More generally, however, it is unclear that even a substantial increase in resources can fully overcome the implications of the law of diminishing returns. New research can only be assimilated into the existing body of knowledge if the leaders of a field allocate their limited time and effort to do so. As a result, even in a world with nonrival ideas, scarcity of resources and diminishing marginal productivity may come into play.

References

Abramitzky, Ran and Isabelle Sin, "Book Translations as Idea Flows: The Effects of the Collapse of Communism on the Diffusion of Knowledge," Stanford University Working Paper, August 2010.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang, "Superstar Extinction," *Quarterly Journal of Economics*, May 2010.

Borjas, George J. "The Labor Demand Curve *Is* Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics* 118 (November 2003): 1335-1374.

Dubois, Pierre, Jean-Charles Rochet and Jean-Marc Schlenker, "Productivity and Mobility in Academic Research: Evidence from Mathematicians", IDEI Working Paper No. 606, January 2011.

Furman, Jeff, Margaret Kyle, Iain Cockburn, and Rebecca Henderson. "Knowledge Spillovers, Geographic Location and the Productivity of Pharmaceutical Research," *Annales d'Economie et de Statistique*, forthcoming 2011.

Ganguli, Ina. "Saving Soviet Science: The Impact of Grants When Government R&D Funding Disappears," Working Paper, Harvard University, 2010.

Howe, Marvine. "For Émigré Scientists, Job Hunting is Difficult," *The New York Times*, August 12, 1990.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics*, 2(2): 31–56.

Jones, Charles I., and Paul M. Romer, "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital," *American Economic Journal: Macroeconomics*, vol. 2(1), pages 224-45, January 2010.

Kaiser, David. "Booms, Busts, and the World of Ideas," MIT Working Paper, 2011.

Laqueur, Walter. *The Dream That Failed: Reflections on the Soviet Union*. New York: Oxford University Press, 1996.

Lefschetz, Solomon. A chapter in LaSalle, J.P., chair. "*Summary Report of the Panel on Soviet Mathematics."* Research Institute for Advanced Study. Baltimore: Martin Company 1961.

Lucas, Robert E. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, Volume 22, Issue 1, July 1988, Pages 3-42.

McClure, Donald. 1992. “AMS Task Force on Employment Report to the Mathematical Community Part I: Academic Hiring Survey, 1991-1992.” *AMS Notices*, 39(4): pp. 311-316.

Polyak, B.T. “History of mathematical programming in the USSR: analyzing the phenomenon,” *Mathematical Programming* Ser. B., 2001.

Romer, Paul M. “Increasing returns and long-run growth,” *Journal of Political Economy*, 1986.

Romer, Paul M. “Endogenous Technological Change,” *Journal of Political Economy*, Vol. 98, No. 5, Part 2 (Oct., 1990), pp. S71-S102.

U.S. National Science Foundation, *Federal Funds for Research and Development: Fiscal Years 1973–2003*. Washington, DC: Division of Science Resources Statistics, July 2004.

Waldinger, Fabian. “Quality Matters: The Expulsion of Professors and the Consequences for Ph.D. Student Outcomes in Nazi Germany,” *Journal of Political Economy*, 2010.

Waldinger, Fabian. “Peer Effects in Science Evidence from the Dismissal of Scientists in Nazi Germany,” Working Paper, University of Warwick, March 2011.

Figure 1. Ratio of Soviet papers to American papers across subfields of mathematics, 1970-1989

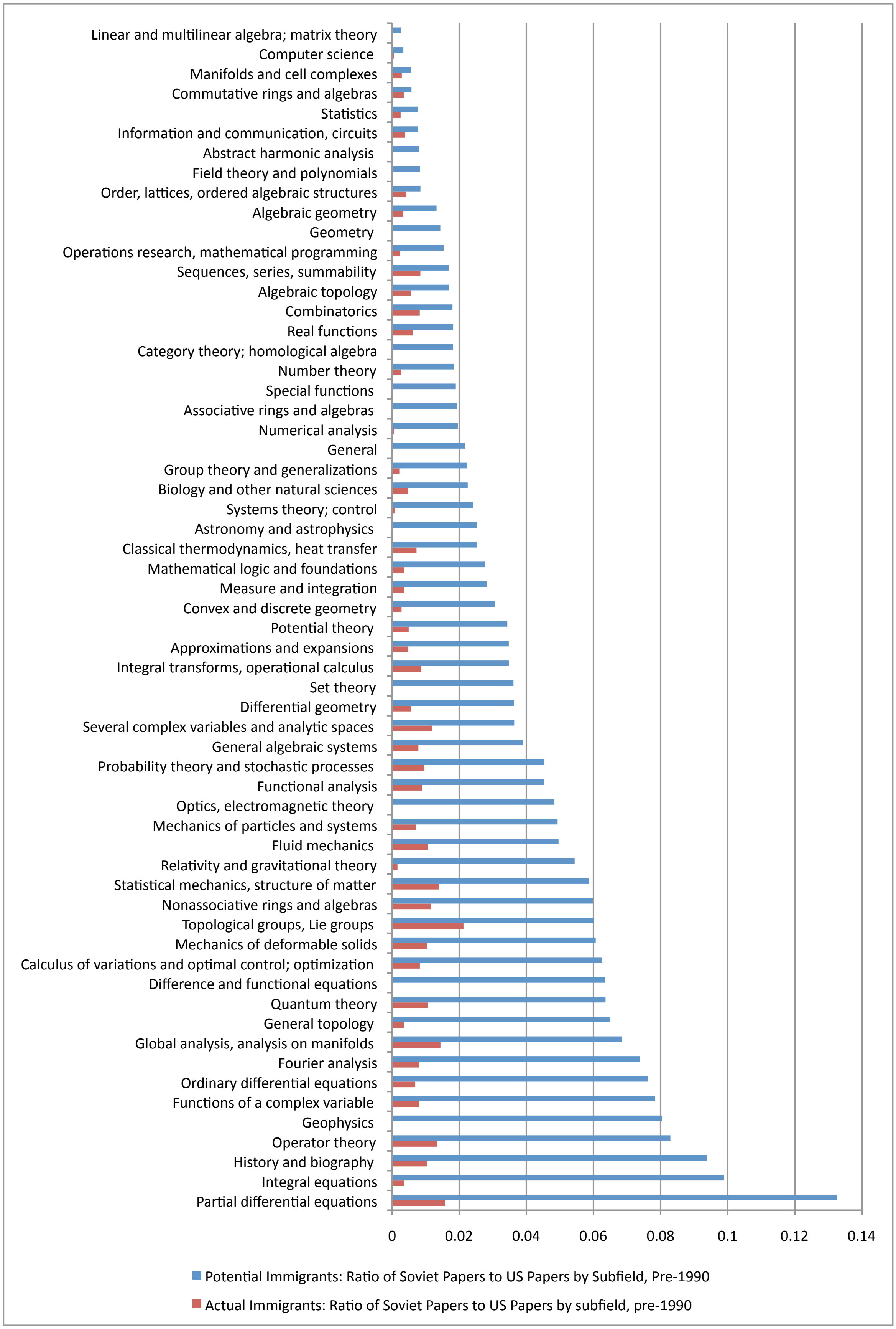
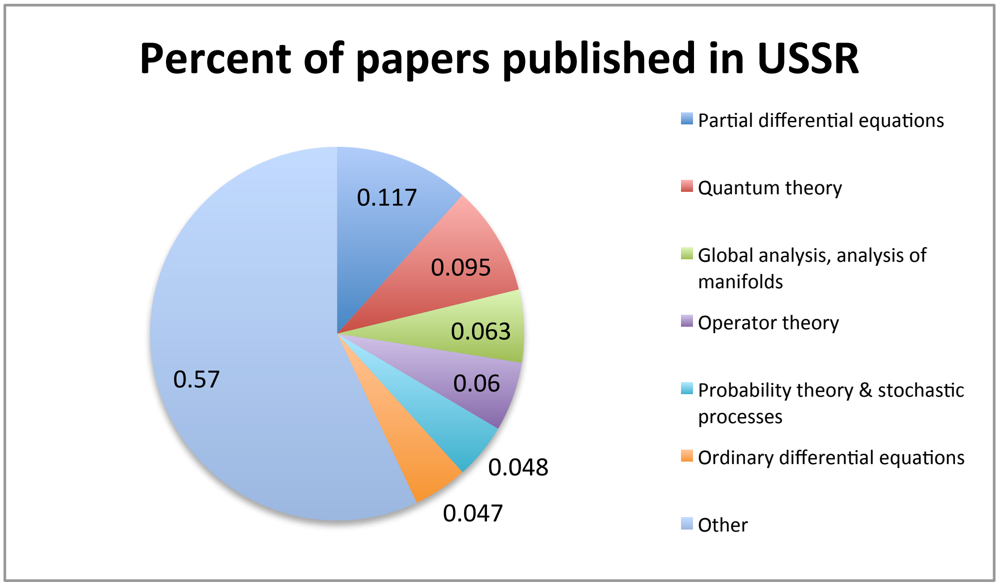
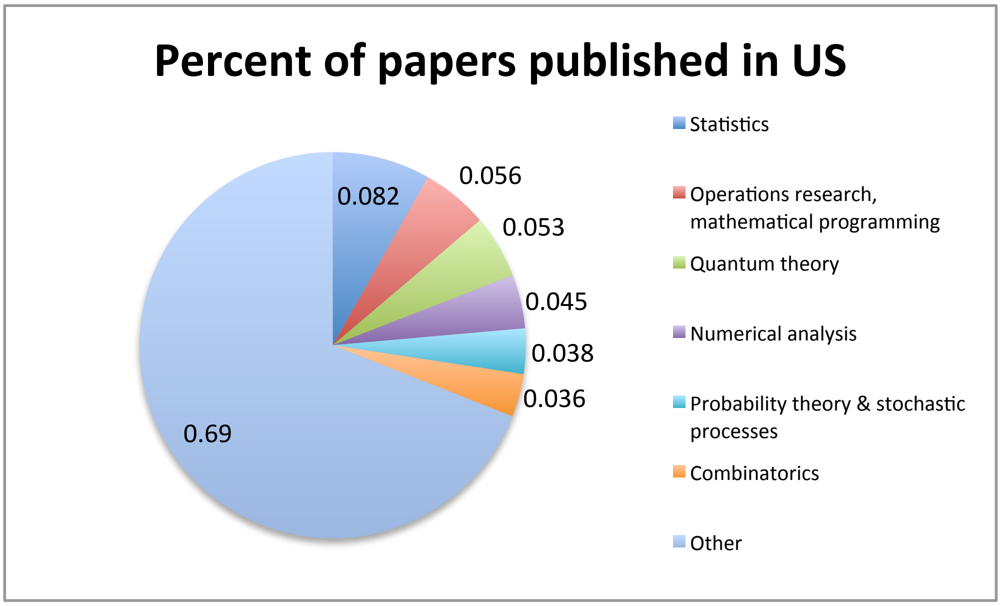


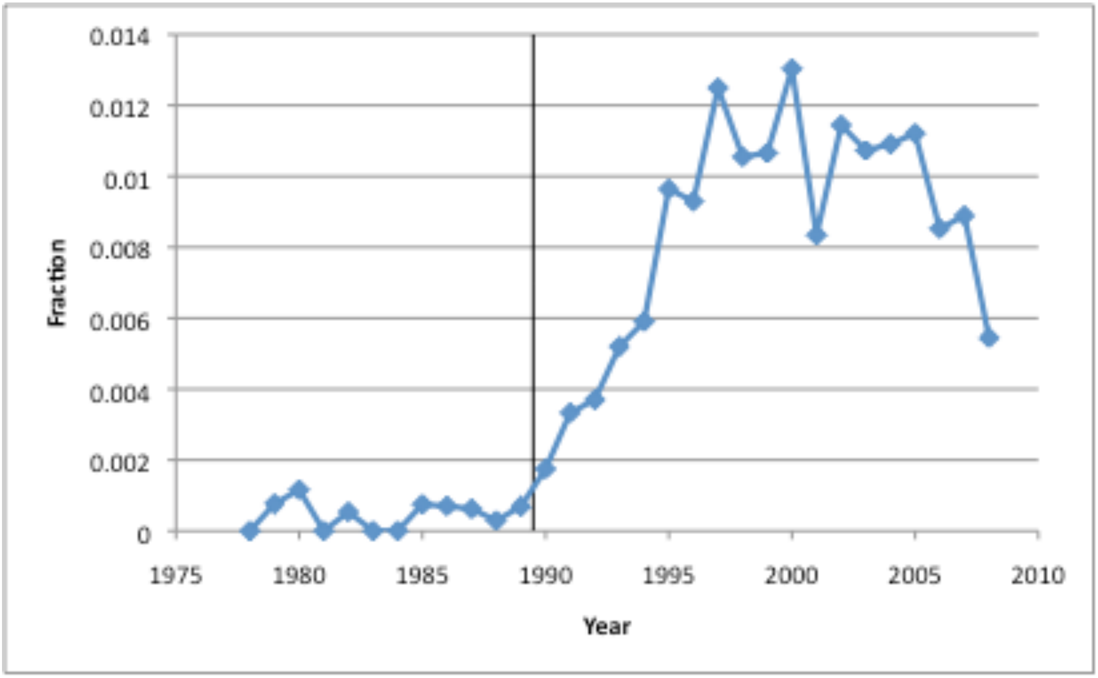
Figure 2. Distributions of pre-1990 publications in mathematics in the USSR and US in top six fields





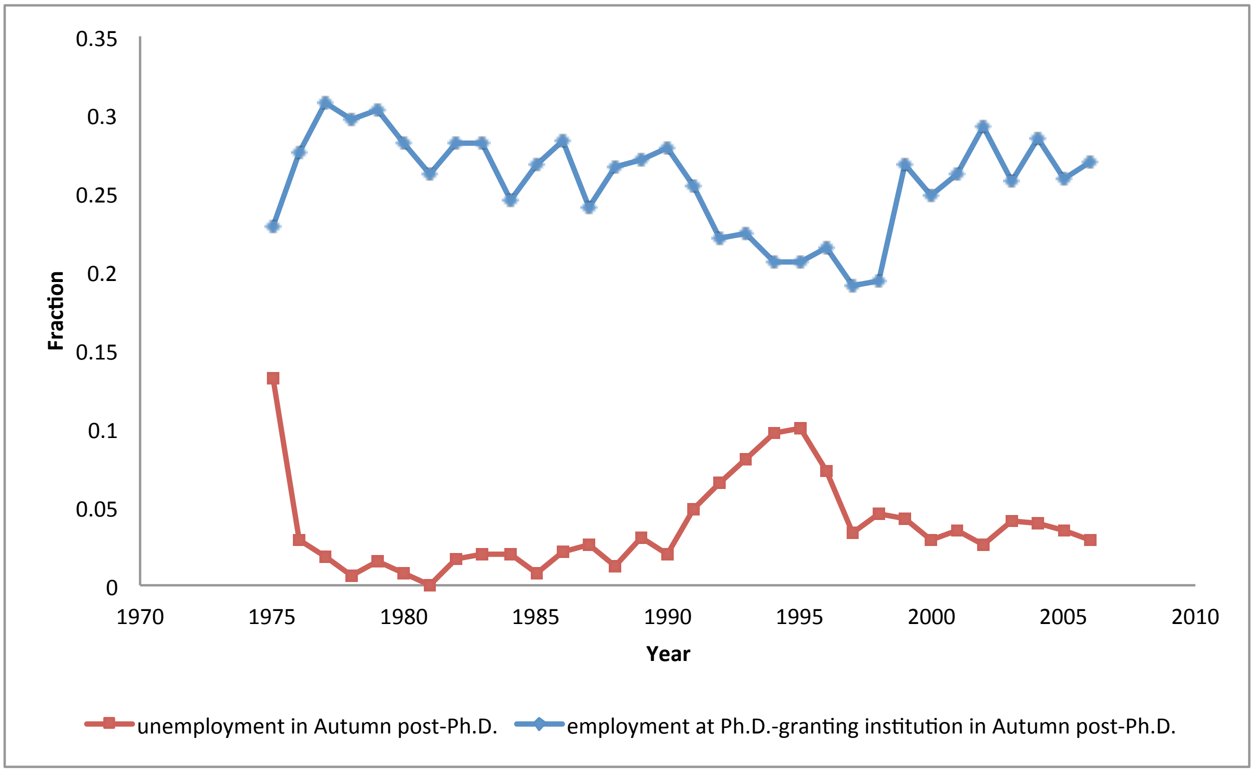
Notes: The data consist of mathematicians who published at least 1 paper between 1970 and 1989. The sample of American mathematicians includes persons who were predominantly affiliated with an American institution before 1990.

Figure 3. Trends in coauthorship rates between Soviet and American mathematicians



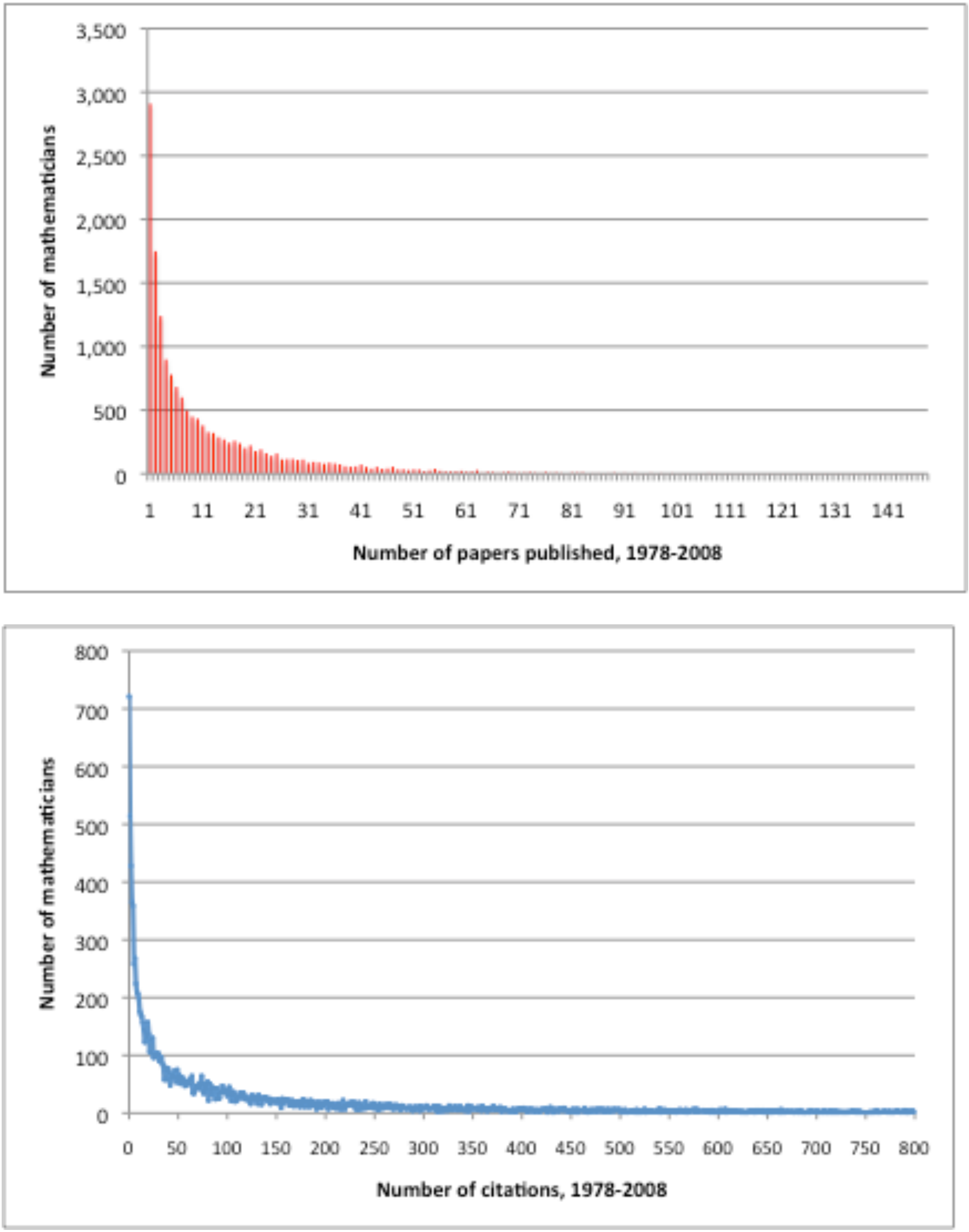
Notes: The denominator of this fraction is the number of papers published each year where at least one author reports an American affiliation. The numerator is the number of such papers in which one other author also reports a Soviet affiliation.

Figure 4. Employment trends for new doctorates in mathematics granted by North American institutions



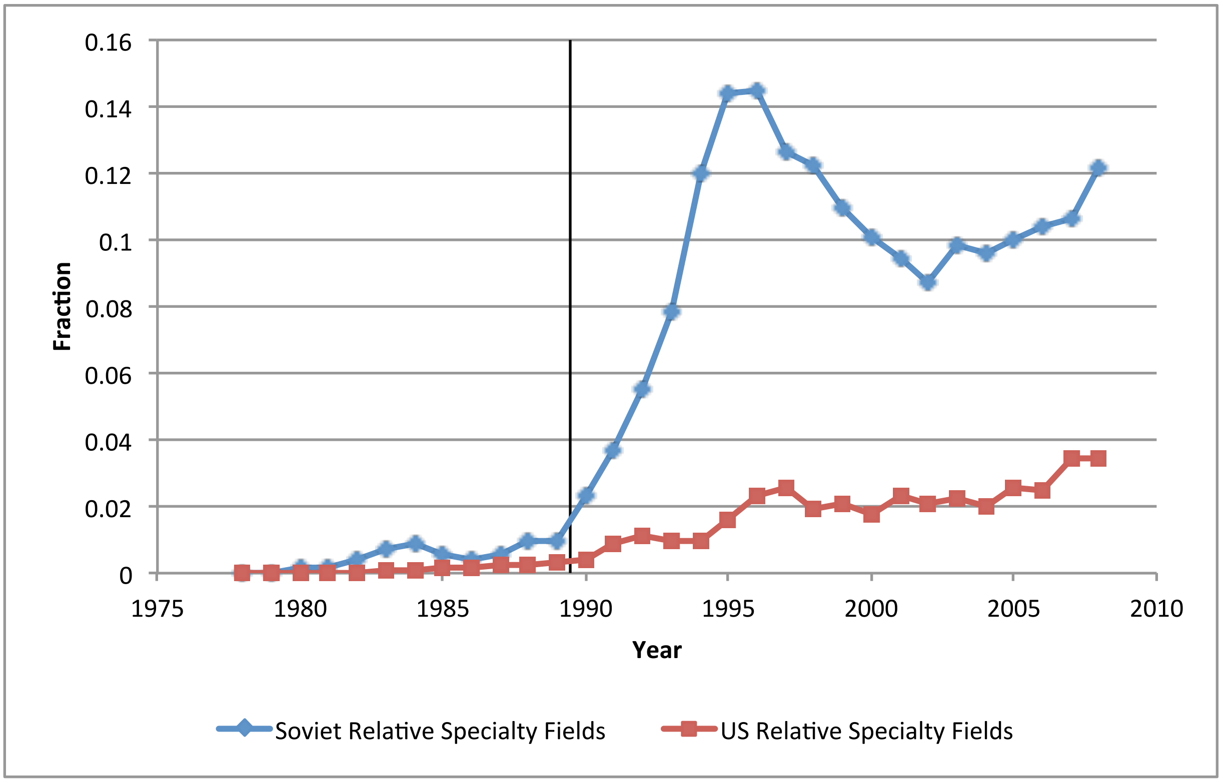
Source: Data compiled by the authors from annual reports produced by the American Mathematical Society.

Figure 5. Frequency of papers and citations among American mathematicians



Notes: The sample consists of American mathematicians who published at least 1 paper between 1978 and 2008, and who were predominantly affiliated with an American institution before 1990.

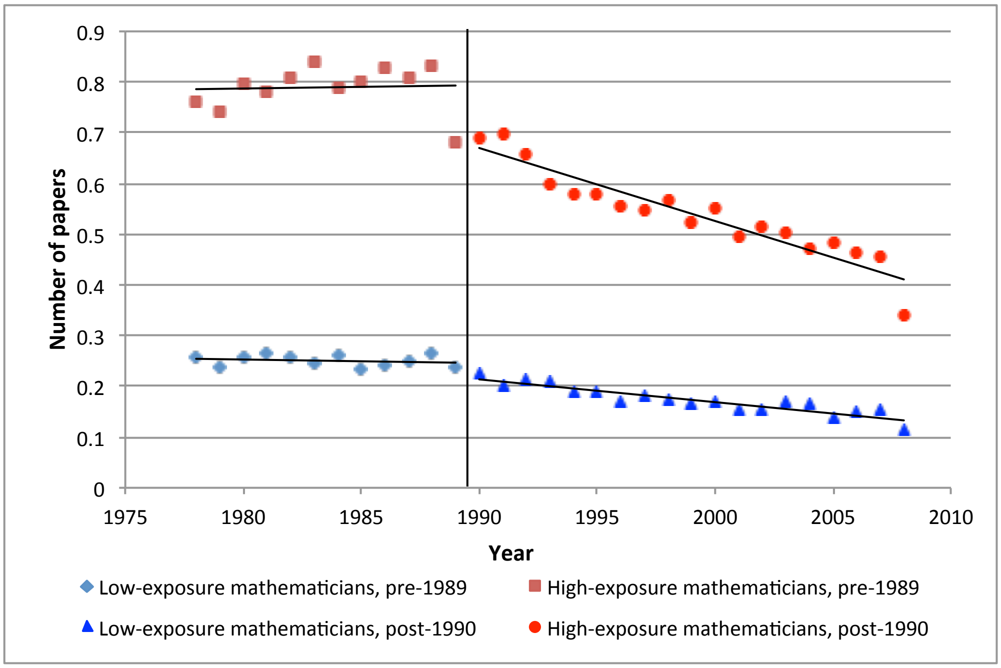
Figure 6. Fraction of publications published by Soviet émigrés



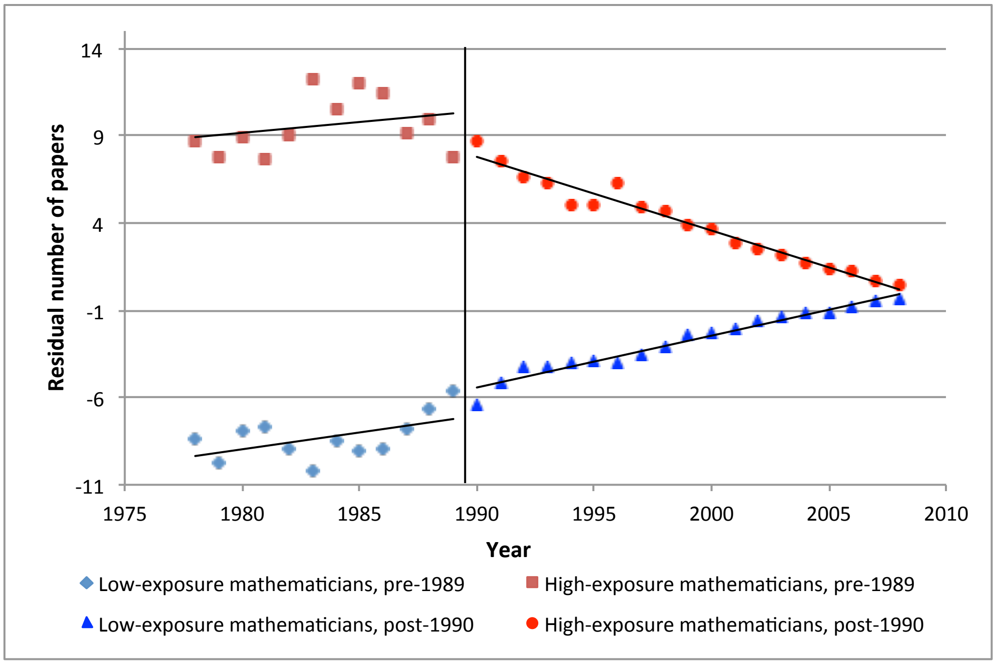
Notes: Soviet relative specialty fields are defined as the 10 fields at the bottom of Figure 1. The U.S. relative specialty fields are defined as the 10 fields at the top of Figure 1. The numerator of the fraction illustrated in this graph is the number of papers published by a Soviet émigré in the specialty fields. The denominator is this quantity plus the number of papers published by an American author in the specialty fields, where an American mathematician is defined to consist of persons who were active before 1990 and who were exclusively affiliated with an American institution in that period.

Figure 7. Impact of index of similarity on output of American mathematicians

A. Annual number of papers per mathematician

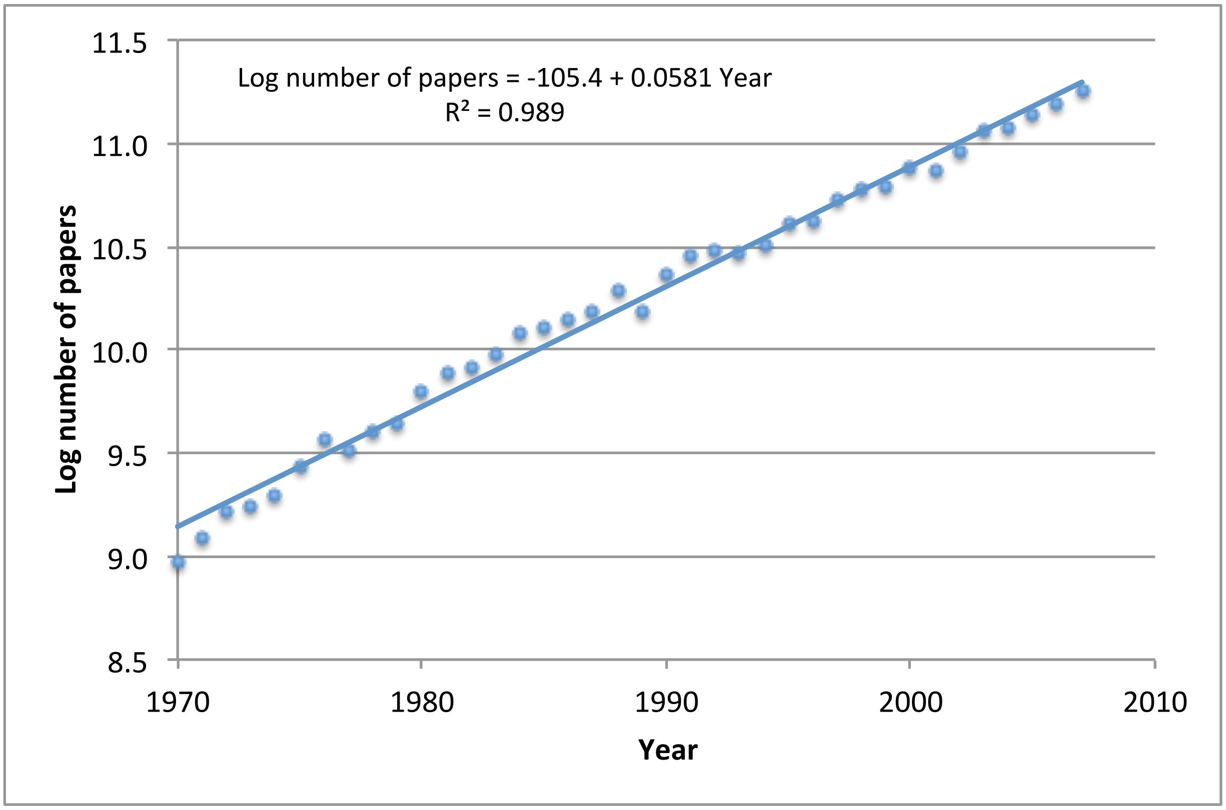


B. Annual number of papers, removing experience and year fixed effects



Notes: The sample consists of mathematicians who were predominantly affiliated with American institutions prior to 1990. The residual number of papers in panel B are calculated from a regression that includes a quartic in years of experience and year fixed effects.

Figure 8. Trend in the number of papers published in mathematics, 1970-2007



Notes: The figure illustrates the trend in the total number of papers published in mathematics.

Table 1. Summary Statistics

Variable:	Predominantly American	Exclusively American
Number of pre-existing mathematicians	19,186	15,812
Total number of papers	266,812	128,571
Total number of citations	4,979,009	2,217,465
Median papers per mathematician	7.0	5.0
Mean papers per mathematician	13.7	10.5
Median citations per mathematician	64.0	45.0
Mean citations per mathematician	257.3	194.4
Percent publishing in four or fewer fields	81.1	85.6
Median year of first publication	1974	1975
Percent who first published after 1980	28.8	33.5

Notes: The sample of "predominantly American" mathematicians consists of persons who published at least half of their pre-1990 papers using a U.S. affiliation; the sample of "exclusively American' mathematicians consists of persons who published all of their pre-1990 papers using an American affiliation.

Table 2. The productivity of Soviet émigrés (relative to Soviet stayers and American mathematicians)

	Number of papers		Number of citations	
	(1)	(2)	(1)	(2)
Pre-1990 comparison of Soviet émigrés with Soviets who remained	3.570 (.347)	3.579 (.312)	139.252 (16.117)	139.363 (16.029)
Post-1992 comparison of Soviet émigrés with American mathematicians	14.180 (.635)	14.247 (.632)	148.522 (14.364)	148.687 (14.366)
Includes year of first publication	No	Yes	No	Yes

Notes: Standard errors are reported in parentheses. There are 2,112 observations in the regressions reported in row 1, and 11,190 observations in the regressions reported in row 2. The baseline sample of Soviet mathematicians in the first row includes all mathematicians who published at least a third of their papers using a Soviet affiliation prior to 1990. The baseline sample of American mathematicians in the second row includes only those mathematicians who were predominantly affiliated with American institutions prior to 1990, and had an American affiliation just prior to the Soviet influx.

Table 3. Top ten mathematicians and institutions with the highest number of pre-1990 publications

	Number of papers	Correlation coefficient	Index of intensity	Index of similarity
Mathematician:				
Paul Erdös	242	-0.003	0.133	0.120
Barry Simon	164	0.700	0.683	0.414
Avner Friedman	162	0.588	0.628	0.295
Peter C. Fishburn	132	0.035	0.173	0.109
Michel Talagrand	127	0.174	0.263	0.191
Ciprian Foias	125	0.473	0.543	0.328
Leonard Carlitz	113	-0.013	0.123	0.056
Joel L. Lebowitz	110	0.135	0.249	0.197
Jean Bourgain	110	0.184	0.380	0.143
Jürg Fröhlich	103	0.469	0.602	0.202
Institution:				
University of California, Berkeley	2152	0.424	0.245	0.423
University of Wisconsin	2021	0.334	0.234	0.422
University of Illinois	1921	0.238	0.197	0.486
Massachusetts Institute of Technology	1675	0.482	0.310	0.461
University of Maryland	1347	0.580	0.324	0.369
Stanford University	1338	0.585	0.341	0.439
Princeton University	1334	0.274	0.224	0.515
University of Minnesota	1225	0.672	0.325	0.336
University of Michigan	1200	0.535	0.282	0.406
University of California, Los Angeles	1085	0.466	0.239	0.407

Notes: All of the calculations are based on the publications data available at the individual or institution level between 1970 and 1989. The bottom panel calculates the various overlap indices by aggregating across all individuals affiliated with a particular institution at the time of publication.

Table 4. Impact of Soviet supply shock at the individual level

	Mathematicians predominantly in U.S.		Mathematicians always in U.S.	
	Number of papers	Number of citations	Number of papers	Number of Citations
A. Author-year regressions				
Correlation coefficient	-.136	-22.460	-.094	-16.983
	(.024)	(1.893)	(.023)	(2.134)
Index of intensity	-.067	-15.475	-.035	-10.931
	(.018)	(1.474)	(.017)	(1.560)
Index of similarity	-1.201	-86.586	-1.075	-72.167
	(.078)	(5.693)	(.083)	(7.253)
B. Author-year regressions, short run				
Correlation coefficient	-.100	-16.329	-.085	-12.080
	(.025)	(1.783)	(.024)	(1.950)
Index of intensity	-.047	-11.276	-.036	-7.782
	(.019)	(1.434)	(.017)	(1.506)
Index of similarity	-.953	-61.569	-.880	-49.994
	(.080)	(5.153)	(.084)	(6.152)
C. Author-year regressions, long run				
Correlation coefficient	-.164	-27.699	-.098	-21.218
	(.029)	(2.271)	(.029)	(2.640)
Index of intensity	-.084	-19.088	-.032	-13.716
	(.022)	(1.745)	(.021)	(1.878)
Index of similarity	-1.396	-107.975	-1.224	-91.048
	(.094)	(6.885)	(.103)	(9.226)
D. Author-field-year regressions				
Correlation coefficient	-.0022	-.3868	-.0019	-.3773
	(.0006)	(.0542)	(.0006)	(.0695)
Index of intensity	-.0011	-.2750	-.0006	-.2510
	(.0005)	(.0431)	(.0004)	(.0518)
Index of similarity	-.0205	-1.4480	-.0188	-1.4405
	(.0019)	(.1650)	(.0022)	(.2370)

Notes: Standard errors are reported in parentheses and are clustered at the individual level. The table reports the coefficient of the interaction between the overlap index and a post-1992 dummy variable. Each observation in panels A, B, and C represents a unique author-year permutation; each observation in panel D represents a unique author-field-year permutation. The regressions in panel B only include observations for the years 1978-1999, while the regressions in panel C only include observations for the years 1978-1989 and 2000-2008, inclusive. The regressions in panel A have 549,202 observations in the sample of predominantly American mathematicians and 448,076 observations in the sample of exclusively American mathematicians; the respective statistics for Panel B are 377,236 and 307,348 observations; for panel C are 395,714 and 322,580 observations, and for panel D are 13,713,988 and 10,717,513 observations. The sample in panel D represents a 50 percent random sample of the population. All regressions include a dummy variable indicating if the year of the observation is after 1992, the mathematician's years of experience (as a quartic polynomial), a vector of year fixed effects, and a vector of individual fixed effects. The regressions in panel D also include all interactions between the year fixed effects and the field fixed effects.

Table 5. Impact of Soviet supply shock at the individual level, by quality of institution and age of mathematicians (Dependent variable = Number of papers)

Sample/Variable	All institutions	Top 3 institutions	Top 5 institutions	Top 10 institutions	Top 25 institutions
A. All mathematicians					
$T \times$ Overlap	-.032	-.031	-.028	-.023	-.016
	(.014)	(.015)	(.015)	(.015)	(.017)
$T \times$ Overlap $\times$ High rank	---	-.027	-.039	-.048	-.043
		(.051)	(.042)	(.039)	(.030)
B. First publication after 1980					
$T \times$ Overlap	-.064	-.036	-.035	-.034	.017
	(.026)	(.027)	(.028)	(.030)	(.037)
$T \times$ Overlap $\times$ High rank	---	-.288	-.203	-.139	-.183
		(.087)	(.070)	(.060)	(.051)
C. First publication before 1980					
$T \times$ Overlap	-.020	-.031	-.028	-.021	-.032
	(.017)	(.017)	(.018)	(.017)	(.019)
$T \times$ Overlap $\times$ High rank	---	.154	.072	.005	.035
		(.058)	(.051)	(.052)	(.038)

Notes: Standard errors are reported in parentheses and are clustered at the individual level. The sample of American mathematicians consists of persons who were predominantly affiliated with American institutions prior to 1990, and whose last reported affiliation prior to the Soviet supply shock was an American institution. There are 508,786 observations in panel A, 345,116 observations in panel B, and 163,670 observations in panel C. The variable T is a dummy variable indicating if the observation is observed between 1992 and 2008; "overlap is a dummy variable indicating if the mathematician's correlation coefficient is in the top decile of the distribution; and "high rank" is a dummy variable indicating if the mathematician's affiliation was in the Top 3, Top 5, Top 10, or Top 25 institutions. All of the regressions include a dummy variable indicating if the year of the observation is after 1992, the mathematician's years of experience (as a quartic polynomial), a vector of year fixed effects, and a vector of individual fixed effects. The regressions also include all possible two-way interactions between the post-1992 dummy variable, the overlap indicator, and the variable indicating whether the institution is highly ranked.

Table 6. Impact of Soviet supply shock on the probability that American mathematicians "retire" from publishing

	Measure of overlap		
Sample	Correlation coefficient	Index of intensity	Index of similarity
All pre-existing mathematicians	.009 (.005)	.001 (.004)	.155 (.015)
First publication in 1985-1989	.051 (.016)	.031 (.012)	.357 (.065)
First publication in 1980-1984	.028 (.010)	.012 (.009)	.275 (.034)
First publication in 1970-1979	-.007 (.007)	-.007 (.007)	.059 (.020)
First publication before 1970	-.010 (.008)	-.015 (.008)	.057 (.021)

Notes: Standard errors are reported in parentheses and are clustered at the individual level. The regression uses the sample of mathematicians who were predominantly affiliated with an American institution between 1978 and 1989. The regression model is a probit, where each individual contributes one observation until the last year in which we observe the individual publishing a paper. The dependent variable is set to zero in every year except for the last year in which a publication is observed. The coefficient reported in the table is that of the interaction variable between the respective overlap index and a dummy variable indicating if the author-year observation is observed after 1992. The sample consists of American mathematicians who were predominantly affiliated with American institutions between 1978 and 1989. The sample sizes (by row) are: 262,937 observations, 33, 557 observations, 54,037 observations, 107,622 observations, and 65,721 observations. All of the regressions include a dummy variable indicating if the year of the observation is after 1992, the mathematician's years of experience (and its square), and a vector of year fixed effects.

Table 7. Impact of Soviet supply shock on the institutional quality of a mathematician's first affiliation

	Dependent variable		
Independent variable	Number of papers, 1970-89	Log number of papers, 1970-89	Institution rank, 1970-89
High-exposure field × post-1992 dummy	-67.953	-.320	47.144
	(22.900)	(.068)	(.014)
Medium-exposure field × post-1992 dummy	-11.166	-.042	3.831
	(18.299)	(.055)	(.066)

Notes: Standard errors are reported in parentheses. Each observation in the regression represents a mathematician observed during the year of his first publication using an American affiliation. The alternative dependent variables measure various aspects of the mathematician's first affiliation, including the total number of papers published by persons affiliated with that institution between 1970 and 1989, and the institutional rank as defined by the total number of papers published in that period (with rank number 1 representing the most productive institution). The sample has 30,238 observations. The regressions are restricted to mathematicians affiliated with an institution that published at least one paper in the 1970-89 period. All of the regressions include a dummy variables indicating if the first paper was in a field with high- or medium- Soviet exposure, as well as the dummy variable indicating if the year of the mathematician's first publication is after 1992.

Table 8. Impact of Soviet supply shock on institutional mobility of mathematicians

Sample/ Independent variable	Measure of overlap/Dependent variable					
	Correlation coefficient		Index of intensity		Index of similarity	
	Moved	Change in quality	Moved	Change in quality	Moved	Change in quality
A. All mathematicians						
Exposed	.023	-2.627	.023	-2.628	.024	-2.620
	(.013)	(.080)	(.013)	(.080)	(.013)	(.080)
Index	.141	-.649	.114	-.521	.297	-2.330
	(.029)	(.205)	(.025)	(.178)	(.081)	(.571)
B. First publication after 1980						
Exposed	.046	-2.580	.047	-2.582	.045	-2.569
	(.017)	(.113)	(.017)	(.113)	(.017)	(.114)
Index	.110	-.325	.085	-.209	.378	-1.843
	(.037)	(.290)	(.030)	(.232)	(.124)	(.953)
C. First publication before 1980						
Exposed	.004	-2.685	.004	-2.684	.005	-2.681
	(.020)	(.112)	(.020)	(.112)	(.020)	(.112)
Index	.165	-.945	.147	-.943	.252	-2.515
	(.044)	(.289)	(.042)	(.275)	(.108)	(.716)

Notes: Standard errors are reported in parentheses. The sample consists of mathematicians whose last observed affiliation before 1990 was in the United States. The dependent variable in the "moved" regressions is a dummy variable set to unity if the mathematician's last observed affiliation before 1990 is different from the modal affiliation after 1992, while the dependent variable in the change in quality regressions is the log difference in the total number of papers published between 1970 and 1989 by the modal institution employing the mathematician after 1992 relative to the last affiliation prior to 1990. The "change in quality" regressions are estimated only in the sample of mathematicians who switched institutions. The "moved" regressions have 6,803 observations in panel A, 3,345 observations in panel B, and 3,458 observations in panel C. The "change in quality" regressions have 4,064 observations in panel A, 2,125 observations in panel B, and 1,939 observations in panel C. All regressions hold constant the total number of papers the mathematician published before 1990, and the mathematician's years of work experience. The "moved" also hold constant the quality of the pre-1990 institution as defined by the institution's rank in the total number of papers published in that period.

Table 9. Determinants of probability of coauthoring with Soviet émigrés

	Measure of overlap		
Sample/ Independent variable	Correlation coefficient	Index of intensity	Index of similarity
A. All mathematicians			
Exposed	.022	.022	.021
	(.007)	(.007)	(.007)
Index	.103	.082	.321
	(.016)	(.014)	(.043)
B. First publication after 1980			
Exposed	.014	.015	.014
	(.010)	(.010)	(.010)
Index	.097	.081	.259
	(.022)	(.018)	(.071)
C. First publication before 1980			
Exposed	.026	.026	.026
	(.010)	(.010)	(.010)
Index	.109	.085	.341
	(.023)	(.022)	(.056)

Notes: Standard errors are reported in parentheses. The dependent variable is a dummy variable indicating if the mathematician coauthored with a Soviet émigré at any point after 1990. The sample of young mathematicians consists of persons whose first publication appeared on or after 1980, while the sample of older mathematicians consists of persons whose first publication appeared prior to 1980. The regressions in panel A have 6232 observations; the regressions in panel B have 2820 observations; and the regressions in panel C have 3412 observations. All regressions hold constant the total number of papers the mathematician published before 1992, the mathematician's years of work experience, and the quality of the pre-1990 institution as defined by the institution's rank in the total number of papers published in that period.

Table 10. Impact of Soviet supply shock on choice of research topics (Dependent variable = Value of overlap index)

	Measure of overlap		
Sample/Variables	Correlation coefficient	Index of intensity	Index of similarity
A. Sample of pre-existing mathematicians			
Index, 1970-89	.711	.706	.618
	(.008)	(.008)	(.010)
Exposed	.005	.004	.004
	(.003)	(.004)	(.002)
B. Pooled sample of new entrants and pre-existing mathematicians			
New	-.043	-.035	-.015
	(.002)	(.002)	(.001)
Exposed	.015	.013	.004
	(.002)	(.002)	(.001)

Notes: Standard errors are reported in parentheses. In panel A, the dependent variable is the value of the index for the individual mathematician in the 1992-2008 period. The sample is restricted to persons who have always been affiliated with an American institution. In panel B, the dependent variable is the value of the overlap index in the 1970-1989 period for the pre-existing mathematicians or the value of the overlap index in 1992-2008 for the new entrants. The pre-existing mathematicians in panel B entered the market between 1970 and 1989. The regressions in panel A have 10,007 observations and the regressions in panel B have 43,010 observations. The regressions in panel A hold constant the log number of papers the mathematician published before 1992, the mathematician's years of work experience, and the quality of the pre-1990 institution as defined by the institution's rank in the total number of papers published in that period. The regressions in panel B hold constant the total number of papers the institution's rank as measured by the total number of papers published in 1970-1989, where the institution is the modal institution pre-1990 for the pre-existing mathematicians and the modal post-1992 institution for the new entrants.

Table 11. Predicting annual output of pre-existing mathematicians

Group:	**Predicted number of papers, 1992-2008**	**Actual number of papers, 1992-2008**	**Difference**
Distribution of correlation coefficient			
Bottom 25 percent	1,333.7	1,376.4	-42.7
			(54.6)
Middle 50 percent	2,644.6	2,620.2	24.4
			(74.8)
Top 25 percent	2,985.3	2,033.8	951.5
			(66.4)
All American mathematicians	6,963.6	6,029.5	934.1
			(114.0)
Soviet émigrés	---	343.0	
Total	6,963.6	6,372.5	591.1
			(114.0)

Notes: The standard errors in parentheses give the forecast error of the predicted number of papers. The sample of American mathematicians consists of persons who were predominantly affiliated with an American institution prior to 1990. The predicted number of papers is based on a regression estimated separately in each of the three samples of pre-existing American mathematicians ranked according to their placement in the distribution of the correlation coefficient. The unit of observation in this regression is the author-year, and the regressors include a quartic in experience and individual-level fixed effects. This regression only includes observations between 1970 and 1989. We then use this regression to predict each mathematician’s output in each year between 1992 and 2008. The statistics reported in the table are the sums of these predictions across the relevant group of mathematicians and all years.