Stockholm Institute of Transition Economics

WORKING PAPER

November 2014

No. 30

Immigration & Ideas: What Did Russian Scientists 'Bring' to the US?

Ina Ganguli

STOCKHOLM SCHOOL OF ECONOMICS

STOCKHOLM INSTITUTE OF TRANSITION ECONOMICS

Stockholm Institute of Transition Economics (SITE) · Stockholm School of Economics · Box 6501 · SE-113 83 Stockholm · Sweden

Immigration & Ideas:
What Did Russian Scientists 'Bring' to the US?

Ina Ganguli*

November 1, 2014

Abstract

This paper examines how high-skilled immigrants contribute to knowledge diffusion using a rich dataset of Russian scientists and US citations to Soviet-era publications. Analysis of a panel of US cities and scientific fields shows that citations to Soviet-era work increased significantly with the arrival of immigrants. A difference-in-differences analysis with matched paper-pairs also shows that after Russian scientists moved to the US, citations to their Soviet-era papers increased relative to control papers. Both strategies reveal scientific field-specific effects. Ideas in high-impact papers and papers previously accessible to US scientists were the most likely to "spill over" to natives.

*SITE, Stockholm School of Economics (e-mail: ina.ganguli@hhs.se). I received helpful input from Patrick Gaulé, Richard Freeman, Bill Kerr, Sarah Turner, and participants of the NBER High-Skill Immigration Conference and the Stockholm Institute for Transition Economics (SITE) Network Conference. I appreciate data assistance from Raviv Murciano-Goroff and Sifan Zhou. I gratefully acknowledge support from the Harvard Labor & Worklife Program and from the Eurasia Program of the Social Science Research Council with funds provided by the US State Department under the Program for Research and Training on Eastern Europe and the Independent States of the Former Soviet Union (Title VIII).

1 Introduction

Many scientists and engineers today leave their home countries, particularly from the developing world, to go abroad to pursue further training or to continue their careers (Weinberg, 2011; Hunter, Oswald and Charlton, 2009). Evidence shows that such immigrants are key contributors to U.S. innovation along a number of dimensions, usually measured by patents and publications (e.g. Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010). In addition to immigrants' direct contributions to research and innovation through their own papers and patents, the knowledge embodied in these workers has the potential to further benefit their new home country. Economic theory suggests that the human capital or ideas of these highly-skilled workers can "spill over" to others and increase their productivity; thus, having more highly-skilled workers in a country should increase the productivity of the existing ones.

In this paper, I draw upon the end of the Soviet Union and the resulting influx of immigrant scientists to the U.S. to provide the first estimates of the extent which immigrants bring new ideas across national borders that are the basis for knowledge spillovers. During Soviet times, the USSR was relatively "closed" to contact with researchers outside of the Eastern bloc. When the Soviet Union ceased to exist in 1991, there were new opportunities for Russian scientists to emigrate, travel and communicate with foreign researchers. Estimates from the 2000 Census suggest that close to 10,000 Russian scientists and engineers across many science and technology fields immigrated to the United States in the 1990s.[1] For my empirical strategy, I compile a rich panel

[1] Author's counts from the 2000 Public Use Microdata Sample (PUMS) of the Decennial Census (5% sample) using sample weights to be representative of the total population. Included are individuals who immigrated to the US after 1990, were born in the USSR/Russia, were at least 35 at the time of immigration, and reported an occupation of Mathematician, Engineer or Scientist.

dataset of the location of Russian scientists and US citations to Soviet-era publications. First, I exploit variation in the number of scientist emigres across US cities and scientific fields and citations to Soviet papers originating in those cities; and second I use a difference-in-differences approach by comparing the number of citations to Soviet-era papers published by scientists who migrated to the US and by those who did not.

Several recent studies have measured knowledge spillovers by examining the impact of the arrival or departure of highly skilled individuals on the patenting and publishing activity of others. The findings have been mixed, with some evidence pointing to positive spillovers (Moser, Voena and Waldinger, 2014; Waldinger, 2010; Hunt and Gauthier-Loiselle, 2010), negligible impacts (Waldinger, 2012; Kerr and Lincoln, 2010), or negative productivity impacts (Borjas and Doran, 2012).[2] Given the important role of human capital externalities as presumed inputs to innovation for economic growth (Romer, 1990) and policy discussions concerning the types of immigrants that should be allowed to enter the US, these mixed findings suggest a need for a greater understanding of the extent to which immigrants contribute (or do no contribute) to the flow of new ideas into the country that presupposes knowledge spillovers.

This paper builds on important literatures examining the geographic localization of knowledge flows and the role of immigrants in knowledge transfer. These studies, rather than estimating the productivity gains arising from knowledge spillovers, seek to measure knowledge flows themselves using the 'paper trail' left by patents and publications through their forward citations (Jaffe, Trajtenberg and Henderson, 1993).

[2] Also drawing on the Soviet collapse, Borjas and Doran (2012) use the influx of Soviet mathematicians to the US after the end of the USSR to estimate the impact on the productivity of native US mathematicians. They find lower publication rates for natives in subfields with greater overlap with Soviet mathematicians, which can be explained by increased competition in the mathematics labor and publication markets.

The importance of geographic proximity for the flow and production of ideas has been central to models of knowledge diffusion and industry agglomeration. The oft-cited passage in Marshall perhaps describes the mechanism best: "..so great are the advantages which people following the same skilled trade get from near neighbourhood to one another... if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas" (Marshall, 1920). Economic models based on this idea posit that individuals accumulate human capital by interacting and learning from one another through close geographic proximity (Glaeser, 1999) as well as through social proximity (e.g. Agrawal, Cockburn and McHale, 2006).

In particular, face-to-face interactions facilitated by geographic proximity are thought to be particularly important for transferring 'tacit' knowledge. Since codified knowledge (i.e. knowledge recorded in articles) is based on a larger stock of knowledge, much of which is tacit and can only be transmitted in person, knowledge embodied in individuals might be especially important for the production of new ideas (Breschi and Lissoni, 2009). It is through this mechanism - the colocation of immigrants with natives over a geographic unit that facilitates face-to-face interactions - that immigrants presumably contribute to the flow of new ideas into a destination country, which can then serve as inputs into natives' individual knowledge production functions.

Several studies have used the mobility of inventors and patent or article citations to document localized knowledge spillovers and the important role of people in transmitting ideas. These studies show that mobile scientists and engineers do influence knowledge flows and have provided significant understanding of the conditions under which these knowledge flows flourish, such as through social networks (Breschi and Lissoni, 2009) and star scientists (Azoulay, Zivin and Sampat, 2012; Zucker and Darby, 2006) and across patent technology classes (Agrawal, Cockburn and McHale, 2006).

While these studies have exploited the mobility of individuals within the US, there is also evidence that individuals can contribute to knowledge transfer across national borders. Kerr (2008) and Agrawal et al. (2011) use patent citations and the ethnicity of inventors to show that ethnic scientific communities play an important role in diffusing knowledge to immigrants' home countries.[3] Moreover, Agrawal et al. (2011) shows that knowledge transferred back to the home country is especially valuable as an input for high-impact inventions produced there. Although these studies suggest that we should expect immigrants to facilitate knowledge flows to natives in the destination country, the existing body of research lacks empirical evidence on whether and under what conditions this occurs.

There are several challenges in estimating the causal impact of immigration on knowledge flows to natives. The natural experiment provided by the Soviet collapse and the rich data used in this analysis allow for an arguably "clean" causal estimate for a few reasons. First, the influx of Soviet scientists to the US provides a large number of immigrants whose location and 'ideas' can be traced over time and space using a 'paper trail' of paper-to-paper citations. Second, the influx was spread over many locations in the US and across many scientific fields, providing useful geographic and field variation that can be exploited to address endogeneity concerns and provide insight into the conditions under which cross-border knowledge flows can flourish. Third, the opportunity for tacit

[3] Kogut and Macpherson (2011) examine the diffusion of economic policy ideas, using membership data from the American Economic Association, to show how the mobility of US-trained economists abroad impacted the adoption of economic policies by their destination countries.

knowledge to be shared with US scientists increased tremendously after 1991, with this tacit knowledge representing the large body of Soviet-era knowledge that, while codified and accessible before and after the USSR, was relatively unknown to US scientists (Borjas and Doran, 2012).

I use two empirical strategies drawn from the labor and innovation literatures to estimate the causal impact of immigration on the diffusion of knowledge. First, I use an annual city-field panel dataset of the number of migrants arriving to US cities across scientific fields after the end of the USSR and the number of citations to Soviet-era articles originating in US cities in that period. This type of spatial analysis has been used to study the impacts of immigration on various economic outcomes across local labor markets, and in this case exploits variation across cities, fields and time in the number of immigrant scientists arriving to the US from Russia to identify the impact of immigration on the flow of ideas. Since the location choices of migrants are likely to be endogenous, I instrument for the post-Soviet distribution of Russian scientists across US cities with the 1990 distribution of all Soviet-born immigrants. An advantage of this spatial approach is that the outcome measure includes citations to all Soviet-era papers, rather than only the papers published by the immigrant themselves, so it suffers less from concerns about superficial citing behavior, e.g. friends citing their friends.

Recent innovation studies have examined the impact of individual inventor and scientist mobility within the US on citation patterns using comprehensive panel datasets of papers and patents (e.g. Azoulay, Zivin and Sampat, 2012; Singh and Agrawal, 2011). Following these approaches, I next use matching methods to create a product-level dataset of paper pairs of migrants and non-US migrants.[4] To address endogeneity concerns, a difference-in-differences strategy allows me to compare the forward citations of Soviet articles "treated" by migration of the author to the US to the citations of similar control papers not treated. The matching procedure matches on several characteristics constant over time, but also on citation trends before migration. Thus, I can compare US citations to very similar pre-1990 Soviet papers authored by migrants and non-migrants in both the pre- and post-move periods. If immigrants 'bring' their ideas with them, U.S. authors should be citing the migrants' papers more than the similar papers of Soviet scientists who did not move.

For both empirical approaches, I draw upon a unique panel dataset of Russian scientists, their publications and locations, and citing publications across many fields of science. I identify and match Russian scientists to their publications, citations and affiliations during the pre- and post-Soviet periods using the Thomson Reuters ISI Web of Science database and use a unique name-matching technique to account for the transliteration of Russian names. I can observe the publishing activity of Russian scientists before and after the end of the USSR and I observe which scientists immigrated to the US.

Both approaches confirm that immigrants do contribute to cross-border knowledge flows. The city-field-level analysis shows that citations to Soviet-era publications increased significantly in cities and fields in years when more migrants arrived there. I also find that after a Russian scientist moved to the US, citations to his or her papers published during Soviet times increased relative to similar control papers authored by

[4] Note that the non-migrant sample includes migrants to other countries as well as those who stayed in Russia.

non-migrants. The increase is small, but significant and only occurs for a few years. However, given that the Soviet papers are relatively dated and that the estimated coefficients are at the paper-level, even this small increase suggests that immigrants do contribute significantly to the flow of ideas across countries.

Both approaches also show there are differences in the impact of migration across fields, which suggests that there may be some types of knowledge or conditions that differ by field that make ideas more likely to be diffused by migrants. Russian migrants in the Life Sciences and Physics tended to diffuse ideas in the US more than migrants in Mathematics. The differences are smaller, but persist after accounting for changes in the total number of native papers that could possibility cite Soviet-era work in cities by field. I also find that the effect is mainly driven by citations to high impact papers, as measured by citations accrued before the Soviet collapse. Moreover, codified knowledge that was already accessible to US scientists (through translated and international journals) was also more like to be transmitted to natives, suggesting that immigration may be especially important in facilitating the transfer of tacit knowledge through face-to-face interactions.

This paper makes several contributions to important areas of inquiry in the economics of immigration and innovation literatures. First, it provides new evidence that ideas are indeed embodied in immigrants, and they are an important channel through which knowledge is diffused to natives. Importantly, the evidence shows that the extent to which knowledge is diffused through immigrants depends on factors such as the scientific field and quality of the idea, and the transmission of tacit knowledge can be especially important through colocation with immigrants. The multi-level approach shows that the findings are robust to the method and unit of analysis. This study also contributes to a small group of papers have been able to track the individual mobility of high-skilled workers across countries to see how they influence knowledge production (Kogut and Macpherson, 2011; Kahn and MacGarvie, 2008). Finally, the findings provide new evidence concerning the impacts of the breakup of the USSR, perhaps one of the most important events of the 20th century.

The paper proceeds as follows. I provide some background about the historical context and the mechanism of immigration in Section 2. In Section 3 I describe the construction of the datasets. The empirical strategy and results follow in Sections 4 and 5 respectively, and Section 6 concludes.

2 Soviet Science and Mechanisms of Immigration

While the USSR had a large scientific community, it was very "closed" to contact with researchers outside of the Eastern bloc. Scientists were rarely able to travel outside of the USSR and were not allowed to emigrate due to emigration and exit restrictions[5], although there was substantial contact with scientists in other communist countries. Scientists had access to the numerous Soviet journals as well as to journals from Eastern bloc countries. In addition, they were able to order reprints of articles from foreign journals, but these had to be requested via Moscow. Scientists in this analysis would have mainly worked in the Academy of Sciences, where basic research was conducted and findings were published in academic journals.

With the end of the USSR in 1991, scientists gained many freedoms, including greater mobility and contact with the western world. Several factors appear to have

[5]The Soviet government allowed Jews to emigrate, with the numbers increasing in the 1970s and 1980s. The US government granted these individuals refugee status, but in 1989 a ceiling of 50,000 was set for Soviet refugees.

influenced the emigration decisions of scientists. First, there was legislation on both the Soviet/Russian and US-sides that acted as mechanisms of immigration. The Soviet "Law on the Procedures of Exit from the USSR and Entry to the USSR for Citizens of the USSR" established the right of Soviet citizens to emigrate and to return to the USSR. The law passed the Supreme Soviet in May 1991, but only entered into force on January 1, 1993 (Moody, 1996). The 1991 "Law on the Employment of the Population of the Russian Federation" also gave Russian citizens the right to work abroad (Ivakhnyuk, 2009).

On the US-side, the Soviet Scientists Immigration Act of 1992 (SSIA, 1992) facilitated the immigration of scientists to the US. The SSIA authorized "the admission to the United States of certain scientists of the independent states of the former Soviet Union and the Baltic states as employment-based immigrants under the Immigration and Nationality Act." It allotted 750 immigrant visas to eligible scientists, which were individuals with expertise in nuclear, chemical, biological or other high technology fields or worked on defense projects in these areas. Importantly, it did not require qualifying scientists to have an offer of employment in the United States. Many of the 750 visas would have likely gone to scientists who had worked in military facilities, where publication of scientific research was not common.[6]

A key factor influencing emigration decisions was the economic crisis that accompanied the end of the USSR, which led to poor living and working conditions. As Ganguli (2011) describes, there were dramatic drops in funding for science and the wages

[6]The SSIA expired after 4 years, on October 24, 1996, but then it was reinstated on September 30, 2002 for 4 years, and the limit on the number of visas was increased to 950. An interim rule from October 19, 1995 required SSIA applicants "to submit corroborative evidence of claimed expertise including the official labor book, any significant awards or publications and other comparable evidence or an explanation of why such evidence cannot be obtained." Thus, in the final year, it was presumably more difficult to obtain a visa.

of scientists. While many scientists chose to move abroad to the United States, Israel or Europe to continue their careers, others remained at home and sought opportunities to continue their research, despite the economic instability. Some, meanwhile, left science completely and pursued other career options.

Estimates of the number of scientists who left to go abroad after the end of the USSR vary so widely that they are almost meaningless. Using data on Soviet scientists who had published in top journals during Soviet times, Ganguli (2012) shows that migrants tended to be younger, were more likely to have coauthored with a non-USSR scientist in the past, were more likely to come from Moscow, and were more likely to come from the upper deciles of the productivity distribution.

One factor facilitating immigration and collaborations between Russian and Western researchers were the numerous foreign grant programs aimed at fostering international collaboration and exchange programs for Russian scientists. In the 1990s, the US Government spent $350 million per year on average to support science and technology cooperation with Russia, and $200 million per year specifically on joint research projects (Wagner et al., 2002). There was particular concern among western countries about the outflow of knowledge former Soviet scientists possessed about nuclear, chemical, and biological weapons. Worry that rogue nations or terrorist groups would recruit scientists to build weapons of mass destruction, or that they would sell their knowledge without leaving, was a key motivation for the creation of many western assistance programs.

During Soviet times, scientists in the US appear to have had relatively rarely accessed Soviet publications. As Graham and Dezhina (2008) note, not many people outside the USSR knew Russian. Locke (1956) describes a case when American mathematicians didn't discover an important paper published in 1950 in a Soviet Academy of Sciences journal for 5 years, even though this journal was available in the US, "simply because most U.S. scientists and engineers cannot read Russian." There were efforts, however, to make Soviet research available to US researchers. Beginning in the 1950s, there were private translations of several Soviet journals, as well as large-scale programs by the NSF and the NIH, which both provided selective and "cover-to-cover" translations (Garfield, 1972). There was variation in the number of journals translated by field, and the number translated increased through the 1980s. Berry (1988) notes that the greatest number of translated journals were in Physics and Mathematics, so that almost all the main Soviet journals in these fields were translated into English by the 1980s. While the English editions of journals usually appeared 6-12 months after the Soviet publication, the Institute for Scientific Information (ISI) and other indexing/abstract services indexed Soviet journals in a timely manner (Garfield, 1972).

After the end of the USSR, there was greater availability of Soviet publications in the US, but information took time to flow freely, and by the time publications were readily available on the web, many Soviet publications were already quite old. The primary channels through which Soviet research was communicated to US researchers was most likely though face-to-face contact at conferences, during research exchanges, and through immigrants.

3 Data

To proceed with the analysis, I constructed a dataset that links the Soviet-era publications of Russian scientists to the timing and location of their post-Soviet migration behavior, as well as to the papers published in the US (not by them) that subsequently cited their Soviet-era publications.

I first identified a sample of scientists who were "doing science" in Russia around the time of the Soviet collapse. I use publication data from the Thomson Reuters ISI Web of Science (ISI)[7] to create a sample of Russian scientists who were actively doing scientific research near the end of the Soviet Union in 1991. I identified the top Soviet and Russian language journals in the ISI database and extracted all authors publishing in these journals between 1986-1994 (see Appendix Table B5 for the list of Russian/Soviet journals used to create the sample). The ISI database includes over 100 top journals of the former USSR and Russian language journals. It includes journal backfiles to 1900, however journals entered the database at different times. By the 1970s, most of the Russian journals in the database in later years had entered the database.

I next identified the subset of authors who had an address that included a city in the former Russian Republic of the Soviet Union. I dropped any individuals with a foreign address before 1990. I further restricted to authors who "stayed in science" after 1991 and I could identify their location, meaning they published at least one article after the end of the USSR through 2008. For each scientist I have a record of their publications from the year they first enter the ISI database through 2008. For each paper, I know all the basic information, such as the year it was published, the journal, author and corresponding author addresses, the number of coauthors, each author's position, subject categories of the journal, and number of pages. Online Appendix A provides a full description of the preparation of the publication data, including information on transliteration and name

[7]Web of Science ® prepared by THOMSON REUTERS ®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA:

matching and assigning scientific fields to individuals.

3.1 Migrants

To determine who migrated to the US, where and in which year, I use information from the author addresses on each publication. From the addresses I can identify when someone first published in the US and in which city. If I observe a Russian scientist in my sample who is publishing in a US city after 1991, I define them as a "migrant" and I define the year of migration as the year they first publish in the US. While many Russian scientists may have moved later within the US (or back to Russia), I restrict the analysis to the first time I observe a scientist arriving in the US. My sample is limited to those who stayed in science because I can only observe a location for them, but this is also a reasonable definition given the presumed mechanism by which knowledge is transferred to natives. It is unlikely that the immigrants who left science would be transmitting scientific knowledge to natives, as the likelihood that they would be in close geographic proximity and sharing ideas with these individuals would be low.

Given that the date of the first publication in the US is a very noisy measure of the year of migration, a more accurate assessment of the migration year would be preferred using information from CVs or other data not culled from publications. Unfortunately, the lack of CVs and other information for the scientists in my sample (via websites, etc.) makes it difficult to determine a more exact move year.

The full sample of migrants I observe moving to the US between 1992-2002 is 809. For the paper-level analysis, I restrict this sample to migrants who are ever cited by a US paper, which reduces the number of migrants to 535. Figure 1 shows the total number of migrants I observe by year. It shows that there is variation across time, but that there were inflows of migrants throughout the 10-year period. The peak is in 1995 and then again an increase in 1999 (likely due to the 1998 Russian economic crisis). Figure 2 shows the distribution of migrants by the main scientific fields. Physics (39%) and Life Sciences (30%) make up the largest share of migrants in the sample.[8]

To benchmark my sample of Soviet scientist migrants with other counts, I used the 2000 Public Use Microdata Sample (PUMS) of the Decennial Census (5% sample). This comparison is discussed in Online Appendix A. I also use information on the distribution of all Soviet immigrants in the US in 1990 across US cities for an instrumental variables approach in the city-field level analysis described below. To create this variable, I use the 1990 PUMS (State 5% sample). I match each US city in my sample to a Metropolitan Statistical Area (MSA)/Primary Metropolitan Statistical Area (PMSA) in the PUMS. Then I calculate the number of 15-65 year olds reporting a birthplace in the USSR in that MSA/PMSA.

3.2 US Citations to Soviet-era Publications

For the citation data, I matched all articles published between 1980-1990 by the Russian scientists in the sample to the papers that subsequently cited them in the US. After making this link, I calculate the number of US citations to Soviet papers by year, US city, and scientific field of the Russian scientist. Note that I exclude self-cites and citing papers that include a Russian coauthor (to prevent including cites from papers by

[8] My sample includes fewer migrants in mathematics than Borjas and Doran (2012) (38 vs. 336). While their sample includes any individual who published in any math journal in the MathSciNet database, I take a sample of individual Soviet authors and assign each a field based on the subject area of the majority of their publications (as described in Online Appendix A). For example, my sample will assign an individual who published in mainly physics journals with a few math publications (which is not uncommon) as a physicist, while she will be a mathematician in the Borjas and Doran sample. In fact, many of the Physics migrants in my sample appear in the MathSciNet database used by Borjas and Doran (2012). Of a random sample of 10 of my Physics migrants, 6 appeared in the MathSciNet database, and the subject areas associated with their top publications as assigned by the AMS are indeed physics-related (e.g. optics, fluid mechanics, relativity and gravitational theory).

the immigrants themselves or with Russian collaborators).

I limit the Soviet articles to those published from 1980-1990 to exploit the fact that during Soviet times, scientists in the US rarely accessed Soviet publications, despite thir availability as discussed in the previous section. While there was even greater access to Soviet research after 1991, it was still limited and Soviet-era research had already started to be outdated, if only due to the lengthening lag compared to the speed as which science advances. Since the articles of migrants and non-migrants published during this time were thus all subject to the same level of isolationism, restricting the analysis to subsequent citations to these papers lessens the worry that other factors correlated with migration are driving the results rather than the mobility itself.[9] Moreover, there was a large economic collapse that accompanied the end of the USSR, and many scientists exited the science sector (Ganguli, 2011). Restricting the publications to the pre-transition period further prevents confounding the impact of migration from other factors like the economic conditions that may have differentially impacted on the quality of migrants' vs. non-migrants' research.

However, a trade-off to using articles produced only from 1980-1990 is that the number of citations to these articles originating in the US is low, even in the post-Soviet period, partly due to to the growing lag from publication mentioned before. Thus, the variation in the number of citations originating in the US is much lower than if I had included citations to more contemporary articles of migrants and non-migrants in the analysis.

In Figure 3, I show the distribution of the year of publication of all Soviet-era

[9]While there was a shift towards greater openness after 1985 during perestroika, interviews with scientists suggest there were no dramatic changes in the availability of Soviet/US publications.

papers in my sample. There is a slight increase in articles published from 1988-1990, but the articles are otherwise evenly distributed across the years. Figure 4 presents the total number of total US citations to Soviet articles published 1980-1990 from 1992-2002. The number of citations is declining until 1997, then peaks, and then begins to decrease. In Figure 5 I show the total citations from 1992-2002 by the originating city (for the top cities only). The full sample includes 179 total US cities (MSA/PMSA). The top cities are not a surprise, mainly large cities with large universities. However, there is a great deal of variation across the cities and also reflect areas that specialize in specific fields (e.g. Los Alamos for Physics, Bethesda for Life Sciences).

Descriptive statistics for the full sample of Soviet-era articles (1980-1990) published by migrants and non-migrants are presented in Table 1. There are 19,752 articles published by non-migrants compared to 2,570 articles by migrants. Covariates at both the article-level and scientist-level are included. Clearly, migrants look very different from non-migrants. Their articles are slightly more recent on average and have fewer coauthors. Migrants are also more productive during the Soviet period. There are many differences across the fields, with migrants more likely to be in the Life Sciences, Mathematics and Physics than non-migrants. These differences across fields also likely contribute to the productivity differences. Finally, migrants are more likely to come from Moscow than non-migrants.

4 Empirical Approaches

4.1 City-Field-Level Panel

The presumed mechanism driving knowledge flows between immigrants and natives are interactions and communications that are facilitated over a geographic unit, or a location. As Feldman (2000) describes, there is no consensus for the correct unit of analysis for a location in studies of localized knowledge flows, but many studies have settled on the concept of cities (MSAs). Pairing this conceptualization with the definition of local labor markets as an MSA in many immigration studies, I begin with an analysis of the contribution of immigrants to the flow of ideas into a location using an annual city-field level panel on the number of migrants arriving to US cities by field after the end of the USSR and the number of citations originating there in that period. In the spirit of spatial correlation studies estimating the economic impacts of immigration on local labor markets (e.g. Card, 2001), this approach exploits variation across cities, fields and time in the number of immigrant scientists arriving to the US from Russia to identify the impact of immigration on the flow of ideas. I estimate the following regression:

$$y_{ijt} = \beta_1 Migrants_{ijt} + \theta_i + \rho_j + \lambda_t + \psi_{jt} + \varepsilon_{ijt} \tag{1}$$

where y_{ijt} are US citations to Soviet papers published by authors in US city i in field j in year t, $Migrants_{ijt}$ is the number (inflow) of Russian scientist migrants, θ_i are city fixed effects, ρ_j are field fixed effects, λ_t are year fixed effects and ψ_{jt} represents field x year fixed effects. The field and year interactions control for changes in the fields over time, and robust standard errors are clustered at the city-field level. The coefficient on *Migrants*, β_1, provides an estimate of the citations to Soviet-era papers in a given field in a particular US city and year due to the presence of additional Russian migrants.

An advantage of this estimation strategy is that the outcome measure of citations includes all Soviet papers in my sample (described in the Data section), rather than only the papers published by the immigrant themselves. Typically, studies that estimate the impact of individual mobility on the diffusion of ideas limit the analysis to papers or patents published by the focal (moving) individual. While the subsequent product-level analysis is similarly limited, it is also useful to see how the greater stock of knowledge embodied in an individual may be diffused after a move. Since Soviet scientists presumably had knowledge about not only his/her own research, but also those of their colleagues from the USSR, this approach can more fully capture the potential spillovers. A drawback, however, is that if a Soviet scientist immigrated to the US but did not publish any further scientific articles, they are not included in $Migrants$. If those who left science are indeed still contributing to the transmission of knowledge to US natives, then I would still be measuring their contribution to knowledge flows (y_{ijt}) on the left-hand side, but there would be measurement error in the $Migrants$ variable, leading to the problem of "errors in variables" on the right-hand side.

A benefit to this approach is it does not suffer as much from the concern noted in Jaffe, Trajtenberg and Henderson (1993) about whether paper citations represent true knowledge spillovers if scientists cite a friend or colleague "just to be nice" because it is not costly to do so. Note that this measure also excludes self-cites and US citing papers including a Russian collaborator. Finally, in this analysis, the number of immigrants arriving in the US is not limited to only those scientists who were cited at some point by US authors (during the pre- or post-Soviet period), which is the case for the following paper-pair level analysis.

The dependent variable in this regression contains many zero values since often no one in a given field, city and year has published a paper citing a Soviet article. This is not surprising given that papers from the Soviet period would be cited less in the US in general due to the lack of communication during the Soviet period, and then further disaggregating them by city and field leads to many empty cells. There are also many zeroes for the *Migrants* variable, since there are often no inflows of migrants to many cities in certain years.[10]

Spatial correlation studies of the economic impacts of immigration, suffer from two key empirical challenges (see discussion in Borjas, 2003). First, since immigrants do not choose their destinations randomly, there are endogeneity concerns if unobserved factors lead immigrants to enter labor markets correlated with the outcome of interest. In this case, the coefficient on *Migrants* would be biased upwards if immigrants are more likely to move to areas where Soviet papers would be cited more for these other unobserved reasons. Second, it is possible that immigration leads to displacement of natives due increased labor market competition. If natives move to other cities, then potential knowledge flows would not be realized and there would be no significant effect of *Migrants*.

I use two strategies to address these concerns with the city-field-level analysis. To address the endogeneity of the distribution of Russian scientist migrants across US cities after the end of the USSR, I exploit the notion that one factor in the location decisions of Soviet scientists arriving in the US was existing Russian immigrant enclaves. While scientists' professional networks would have naturally also played an important role, I argue that the broader immigrant network, including non-scientists, likely also facilitated location decisions. As described earlier, the economic conditions in Russia in the 1990s were dire, and qualitative evidence suggests that in many parts of Russia, subsistence,

[10] Of the 9,845 observations in the sample, 93% (9,237 observations) are zeroes. Therefore, I have also estimated negative binomial models of specification (1) given that this is count data and there are many zeros, and the results are similar.

rather than science, was at the forefront of scientists minds when deciding to emigrate (see discussion in Ganguli, 2011). Russian scientists immigrating to the US typically were credit constrained, and often did not have offers of employment. This was also facilitated by the Soviet Scientists Immigration Act, which did not require scientists to have offers of employment in order to immigrate. Most Russian scientists were unlikely to speak English fluently, and would have been unfamiliar with life and practices in the US.

Thus, the broader immigrant network, and not only the professional network, would be expected to be important as a means to enter the US and adjust. The professional network of Russian scientists in the US might even be expected to be less of a factor for location decisions, given the high competition among Russian scientists to get academic jobs at this time. A quote from a Russian emigre in a *Science* article from 1990 describes the situation facing emigres at the time: "'The person who arrives to this country from the Soviet Union is like a child,' says emigre Evgeny Chudnovsky, physicist at CUNY's Lehman College. 'He doesn't understand what's going on. He is surprised by all the competition.'" (Holden, 1990). The article also describes how subsistence was initially the primary concern for many scientists arriving in the US: "Many have been trying to support their families with menial jobs - taxi driving, dog walking, doorkeeping..." In other words while they might have preferred to go to destinations where "intellectual action" was happening, it is likely they went where they could initially assimilate.

Using the idea of the supply-push instrument described in Card (2001), I thus decompose the actual inflow of Russian scientists in a given scientific field moving to a US city into an exogenous "supply-push" component and a residual component, where the supply-push component is based on total inflows of Russian scientists in a field and the fraction of Russian immigrants living in the city before the end of the USSR in 1990. In this case, I instrument for the post-Soviet distribution of Russian scientists ($Migrants_{ijt}$ in specification (1)) across cities with the initial distribution of all Soviet-born immigrants.[11] The key exclusion restriction is that while existing immigrant networks play an important role in the location choices of immigrants, which is captured by the 1990 distribution of Soviet-born immigrants in US cities, the determinants of the 1990 distribution of all Soviet migrants is not likely to be correlated with the post-1990 citation behavior to Soviet-era scientific articles.

The instrument essentially weights the inflows of Russian scientist migrants across US cities using the 1990 distribution of Soviet-born immigrants, so that the instrument for the number of Russian scientists migrating to US city i in field j in year t is:

$$\frac{AllSovietMigrants_{i1990}}{AllSovietMigrants_{1990}} \times ScientistMigrants_{jt} \tag{2}$$

where $AllSovietMigrants_{i1990}/AllSovietMigrants_{1990}$ is the share of all Soviet-born immigrants in the 1990 Census in city i and $ScientistMigrants_{jt}$ is the total number of Russian scientists in field j arriving in the US in year t.

To construct the $AllSovietMigrants_{i1990}/AllSovietMigrants_{1990}$ component of the instrument, I used the 1990 PUMS (State 5% sample) and matched each US city in my sample to a Metropolitan Statistical Area (MSA)/Primary Metropolitan Statistical Area (PMSA) in the PUMS. Then I calculated the number of 15-65 year olds reporting a birthplace in the USSR in that MSA/PMSA. Table B2 in the Appendix shows the first stage estimates and shows that the instrument is a significant predictor of the number

[11] Similar approaches have been used in many recent studies of the impacts of immigration (e.g. Hunt and Gauthier-Loiselle, 2010; Cortes, 2008).

of Russian scientist migrants. The R^2 values are low, but an F-Statistic of 32.8 (Column 1 of Table B2) suggests a strong first stage. However, for most cities, the shares of $AllSovietMigrants_{i1990}/AllSovietMigrants_{1990}$ are small and are driven by a handful of people in all but the largest cities.[12] Out of the 179 cities, 43 had zero Soviet immigrants in the 1990 Census. New York (28.6%) and Los Angeles (18.6%) by far had the largest shares of Soviet immigrants in 1990, followed by Chicago (5.1%). As a robustness check, I additionally estimate all results excluding New York and Los Angeles. Appendix Table B1 shows the 10 cities with the larges shares of Soviet Migrants in 1990.

The second issue, the displacement of natives, is an important factor to consider in this setting, especially given evidence in Borjas and Doran (2012) that the arrival of Soviet mathematicians after the end of the USSR led to lower publishing rates among US mathematicians, in part due to increased labor market competition. If native mathematicians published less when there was increased migration, then the potential for knowledge flows, or citations to Soviet work, could decrease relative to areas with fewer migrants. More generally, since some cities and fields would have had different levels and time trends in the number of native authors and native papers written than others, it is important to account for this in the estimation. I create a measure of the set of native papers that could potentially cite Soviet-era papers, which is defined as the total number of papers written by natives in a city, field and year and include this variable as a control in the regressions. For each US city, I count the number of papers by field after excluding papers that appear in my sample of papers of Soviet emigres. Field is determined using the Web of Science subject categories that are assigned to journals, as

[12]Note that the shares of $AllSovietMigrants_{i1990}/AllSovietMigrants_{1990}$ add to less than 1 over the set of cities in my sample since the denominator includes immigrants not assigned to an MSA or in cities not in my sample.

described in Appendix A.

4.2 Difference-in-Differences Analysis

Next, I use a difference-in-differences (DD) analysis at the matched paper-pair level to provide causal evidence on the contribution of immigrants to the flow of ideas at a very micro-level. This approach helps me examine factors associated with knowledge flows that can not be captured with the aggregated data and address endogeneity concerns due to observable and unobservable differences between emigres and non-emigres and their papers.

I use a nonparametric matching method, “Coarsened Exact Matching” (CEM) (Blackwell et al., 2009), which has been used in the recent economics of science and innovation literature to create matched control groups for scientists, articles, and patents. The basic approach of CEM is to temporarily coarsen each covariate, create unique strata based on the coarsened values of the covariates, assign each observation to a stratum, and then drop any observations in strata in which there isn't a control observation for each treatment observation (Blackwell et al., 2009). At the end of the process, there are an equal number of treatment and control observations that are balanced on the covariates selected to match on.

Using CEM, I create pairs of matched papers published by migrants and non-migrants who look very similar on observables at the time the migrant moves. I do this by matching on both scientist-level covariates and paper-level covariates to ensure that migrants and non-migrants look very similar in terms of productivity, renown and openness as well as to ensure that the papers of migrants and non-migrants matched look very similar, including on their pre-migration citation trends.

The scientist-level covariates used in the matching process include scientific field, total citations to articles pre-1990 (not restricted to citations originating in the US), career age (coarsened to several categories), having had a foreign coauthor pre-1990, and originally from Moscow. I also match on article covariates, including the year of publication, the number of coauthors and the pre-move citations originating in the US.[13]

Following Azoulay, Zivin and Sampat (2012) I use the following specification to test for the impact of migration on citations at the matched paper-pair level[14]:

$$Cites_{mt} - Cites_{nt} = \beta_0 + \beta_1 PostMigration_{kt} + f(Age_{kt}) + \phi_{mn} \tag{3}$$

The dependent variable is the difference in total citations in year t between all US citations received to a paper "treated" by migration and the matched control paper, where m denotes an article written by a migrant and n is the matched paper written by a non-migrant.[15] $PostMigration$ is an indicator variable for the years after the migrant (focal) scientist k moves to the US, $f(AGE)$ is a function of scientist k's age, and ϕ_{mn} are the paper-pair fixed effects. Thus, β_1 reflects changes in the pair's citation rate following the migration of scientist k. Robust standard errors are clustered at the scientist-level. All results I present use ordinary least squares (OLS) estimation.[16]

[13] I also tried to refine the match by including additional paper-level covariates such as the journal and the author's position, but including these in the matching process greatly increases the number of strata and decreases the match rate.

[14] Singh and Agrawal (2011) use a different specification for the difference-in-differences estimation. They use patent-level data for focal and control inventors and include patent fixed effects. I also present results in the appendix using this alternate specification.

[15] An advantage of the DD estimation is that it allows me to 'difference out' unobserved factors correlated with citations between migrant and non-migrant articles by controlling for the pre-move trends in citations. This is only possible, however, if there are some citations in the pre-move period. As shown in Figure 4 there were indeed US citations to Soviet work even before 1990, since e.g. many Soviet journals were available and translated to English (see discussion in Section 2).

[16] An alternative specification is the ratio of the citations received by treated and control papers as the dependent variable. However, given that for most years papers receive zero citations, the ratio of citations has more zeroes, and thus I use the difference in citations. Since all matched papers-pairs come from the same scientific field, differential citation rates across fields is not a concern.

Note that the dependent variable does not disaggregate the US citations by city, as there are too many zero values if I only measure citations originating in a particular city. Thus, these results consider the geographic localization of knowledge over a larger geographic unit (country). This approach thus complements the spatial analysis using the city-field panel by essentially moving the geographic unit of analysis to the national level. The comparison of the results using each approach is useful as a robustness measure given discussion in both the innovation and labor literatures about the proper geographic unit of analysis for estimating the localization of knowledge flows (Feldman, 2000) and impacts of immigration due to the adjustment of local labor markets (Borjas, 2003).

5 Results

5.1 City-Field Panel Results

I first present the results of the regressions relating citations to Soviet articles by city and field to the number of migrants in a given year. They provide more aggregated evidence on the link between migration and knowledge flows before moving on to providing causal evidence at a much more micro-level approach with the paper-pair-level analysis, which allows for additional extensions to the analysis.

Table 3 displays the coefficients on the number of migrants from equation (1) using the city-field panel data for years 1992-2002. Column (1) shows that there is a positive and significant relationship between the number of new migrants in a given city, field and year and citations to Soviet-era articles. The coefficient suggests that each additional migrant is associated with approximately 8 additional citations to Soviet publications. When New York and Los Angeles are excluded from the regression in Appendix Table B3, the coefficient becomes smaller, but is still statistically significant.

I next interact the number of migrants with scientific field dummies. The coefficients in column (2) show that there are differences by field in the effect of immigration on citations. The effect is primarily driven by citations in Life Sciences (the omitted category) and to a less extent in Physics. The effect is zero for Chemistry and is even negative for Mathematics and Astronomy & Earth Sciences. This negative result is surprising, but consistent with recent evidence in Borjas and Doran (2012) that the arrival of Soviet mathematicians after the end of the USSR led to lower publishing rates among US mathematicians. If native mathematicians published less when there was increased migration, then citations could also decrease relative to areas with fewer migrants.

To account for different levels and time trends in the number of native papers that could possibility cite the Soviet-era papers, in columns (3) and (4) I control for the total number of native papers published in a given city, field and year. The overall effect of migration is smaller, and the differences in the effect by field persist.

The IV estimates corresponding to the instrument in equation (2) are presented in the last 2 columns of Table 3. The first IV coefficient in column (5), which corresponds to the OLS estimate in column (1), shows that using an instrument increases the coefficient by a large magnitude, from 8 to 55. When excluding the cities with the largest shares of Soviet immigrants in 1990 (New York and Los Angeles), the coefficient is larger (79). While this appears to be a large increase from the OLS to the IV (over six-fold), interpreting the coefficients relative to the total native papers that could potentially cite the Soviet-era papers is helpful. The mean number of native papers across cities, fields and years in the sample is 300. This implies that in the baseline regression, each additional immigrant led to 0.027 greater cites per native paper in a given city, field and year on average. For the IV estimates, each additional immigrant led to 0.184 greater cites per native paper, so that it would take about 5 additional immigrants for one more citation to Soviet-era work per native paper (vs. 37 for the baseline).

While to my knowledge no other study takes a similar approach in analyzing citations to scientific papers at the city-field level, the magnitude of the baseline result is in line with other studies that examine the impact of mobility on citations at the paper level (e.g. Azoulay, Zivin and Sampat (2012)), and we might expect an even larger result in this context given the relative isolationism in the Soviet period and since the dependent variable is cites to any Soviet-era paper, not only those of the migrants themselves. This suggests that the much larger IV estimates are quite plausible, as they capture the effect from migrants with the greatest potential to transfer knowledge through their mobility.

Given the OLS results that field-specific effects matter, I also estimated the IV by field. The first stage estimates of these regressions are presented in Columns 3-7 of Table B2 in the Appendix, and show that the instrument is a significant predictor of the number of migrants only for Physics and Life Sciences, which had the largest shares of migrants in the sample. The IV estimates by field are very imprecisely estimated in these models (not reported here), but the positive effect in Physics remains (9.9).

The overall larger IV estimates in comparison to OLS is unexpected, since the prior was that the location decisions of immigrants would lead to an upward bias in the OLS estimates. This unexpected larger IV estimate is in line with the results of Hunt and Gauthier-Loiselle (2010), who use a similar IV approach to study the impact of immigration on patenting rates. One possible explanation is that immigrants who don't take into account existing immigrant networks in their location decisions would contribute to the transfer of knowledge even if they didn't migrate, while immigrants who move due to these other historic considerations are likely to contribute to knowledge diffusion only through their mobility. For example, immigrants may be more likely to go to and stay in science in university towns like Champaign-Urbana, IL or Ann Arbor, MI, which are places with smaller existing Russian immigrant enclaves. A Russian scientist going to a university town would probably have an employment offer in advance, and the natives there would thus probably already be familiar with their Soviet work. In a non-university town, the Russians would be more likely to be unknown to the natives upon arrival and so the opportunity for immigrants to transmit new knowledge to them would be higher.

It is notable that the IV estimate for Physics is rather similar to the OLS estimate, which suggests that the larger overall IV results in Column 5-6 of Table 3 is driven by other fields. If the hypothesized explanation for the larger IV estimate compared to OLS is correct, then this implies that it is in the other fields that individuals who moved for historic considerations were more likely to transfer knowledge through their mobility. Evidence discussed Section 2 suggests that Americans were more likely to be aware of Soviet-era math and physics research compared to life sciences and chemistry. It also appears there is some differential concentration of immigrants by field across locations, with immigrants in Physics and Math being more likely to move to university towns than immigrants in Astronomy, Chemistry and Life Sciences, which suggests that the opportunity for knowledge transfer was likely higher in these fields, as the natives would have been less familiar with Soviet era work. Such reasoning would be consistent with the IV results, although I can not draw firm conclusions from the IV regressions by field.

However, overall, the results suggest that there are conditions more conducive to the transfer of knowledge by migrants that may differ across fields.

5.2 Difference-in-Differences Results

I now turn to the DD results using the matched paper pairs. Table 2 shows the descriptive statistics for the matched sample of 2,442 total papers (or 1,221 pairs). This implies a match rate of 47.6% of the migrants' articles.[17] Some of the covariates presented were used in the matching process and others were not. Clearly, the match appears to have worked quite well in pairing similar articles published by similar authors. There are no significant differences for any variables across the migrant and non-migrant groups. Note that there are 422 migrants in the sample but 864 non-migrants, since individual papers by a migrant can be matched to control papers written by non-migrants.

The main results are visible in Figure 6. This graph shows the average difference in US citations between the paired papers (solid line) with a 95% confidence interval (dashed line) by the time to migration (in years). The figure shows that the match was effective in pairing articles with similar pre-migration citation trends.[18] A positive effect of migration on citations in the US can be seen beginning after the second year post-migration. The increase is significant for a few years before fading out. The timing of the increase suggests that there is a lag between when a migrant moves to the US and when the Soviet-era paper starts to get cited. Since I am measuring citations to the Soviet publication originating anywhere in the US, this would be consistent with models

[17] Azoulay, Zivin and Sampat (2012) report a match rate of 25.61%, although they begin with more mover articles and match on more article-level covariates, making their matching process more stringent.

[18] Figure 6 only shows the difference between treated and control papers. A similar graph of the levels for each type of paper shows that that the pre-treatment levels are not zero, although the difference between them is. Therefore, it is possible to do a "difference-in-differences" analysis.

of localized knowledge diffusion, since it would take time for the knowledge embodied in the migrant to diffuse throughout the country.

The regression results corresponding to equation (2) are presented in Table 4.[19] These regressions include paper-pair fixed effects and are OLS regressions with robust standard errors clustered at the scientist-level. Column (1) shows that the coefficient on *Post-Migration* is positive and significant. The magnitude is quite small, but considering that these papers are already quite dated, it is notable that there is a significant effect and suggests that migrants did contribute to the diffusion of ideas in the US. I have also run the models after widening the window around the year of migration to 8 years before and after. The coefficient is slightly smaller and still significant.

In column (2), I include interactions of the *Post-Migration* indicator with the scientific field. The effect for Mathematics is negative, suggesting that control articles were cited more. This negative result is consistent with the city-field results and could arise for the same reasons discussed earlier regarding the productivity decline among natives. The negative result at this level of analysis could also reflect the differences Borjas and Doran (2012) point to regarding the differential distribution of US vs. Soviet mathematicians across subfields. For example, it could be that the migrants passed on the ideas of their Russian colleagues in different subfields who stayed in Russia rather than their own. In the 8-year window, the negative effect for Mathematics disappears, and the impact of migration for physicists decreases relative to other fields, so that the effect is close to zero. Since physics research institutes were among the first to get Internet access across Russia, this might have led to greater flows of knowledge to the US and could have

[19] The results using Singh and Agrawal (2011)'s specification on paper-level data and paper fixed effects are in Appendix Table B4. The basic result of an increase in citations to the migrant's papers post-migration holds, but it is only significant at the 10% level.

made the role of migration in the diffusion of ideas less important later as communication costs fell. Overall, these results are consistent with the city-field-level panel results in the previous section, which also showed a negative effect for Mathematics and a smaller effect for Physics.

The DD approach differences out unobserved factors between migrant and non-migrant articles that remain constant over time, but the selection on unobservables issue may still be problematic. In order to partially address this, I do the DD analysis only for pairs where the control article is authored by a migrant to another country. The idea is that there may be something unobserved about migrants in general, which this analysis will capture and will provide reassurance that the reason the citations increase is not something about the migrants themselves, but about where the migrants immigrated to. There are relatively very few migrants to other countries in the sample, so in the CEM match, 466 of the 1224 matched pairs include a control article by a migrant to another country. Running the DD regression only on these pairs, I find a similar point estimate for the effect of migration, although the standard error is large. So while the estimate is noisy, it appears that cites of migrants to the US increased more than cites of migrants to other countries.

I now turn to some additional extensions that are possible with the paper-level data to more closely examine the nature of the knowledge flows. First, in columns (3) I report the results of a specification including an interaction of the *Post-Migration* indicator and a measure of the impact of the paper. Using all citations accrued to the Soviet-era papers from 1980-1990 (including citations originating in the USSR), I define papers as "high impact" if they were in the 80th decile and above within each field and year of publication.[20] The coefficients on the interaction term in column (3) shows that the effect of migration is due to these high-impact papers. There is no significant difference in the 8-year window between high-impact and low-impact papers though. Thus, the initial knowledge being transferred to natives are the best ones, suggesting that immigrants and natives first discuss the most important ideas and then over time discuss lower impact ones. The results also imply that American scientists were not citing some important Soviet-era papers only because the authors did not immigrate to the US, highlighting the important role immigrants can play in the transmission of knowledge across borders.

As discussed in Section 2, many of the papers in the sample were translated to English during Soviet times or were published in international journals. I next investigate whether the availability of Soviet-era papers through these channels impacted the extent of knowledge transfer by immigrants. Since the ISI, the precursor to the Web of Science, indexed the English translations of Soviet journals, I am able to identify the journals that were translated into English in my data.[21] Column (4) shows the results after including an interaction of *Post-Migration* with a dummy for whether the journal was an English translation of a Soviet journal and for whether the journal was a Russian language journal (so the omitted category includes international and western journals). The effect appears to be driven by non-Russian language journals (including Soviet journals that were translated). Since papers in international journals or Soviet journals that were translated would tend to be higher impact, column (5) includes all interactions.

[20] I use the focal paper to define a pair as "high impact". As shown in Table 2, there is no significant difference for this variable between treated and control papers.

[21] In the Web of Science data, these journals are indicated with "ENGL TR" in the journal name. 218 (17%) of the pairs were articles in journals with an English translation available to US scientists.

While not reported here, I also estimated the models separately by field. The only significant effects are for Life Sciences, where both translated and Russian journals were less likely to be diffused through immigration, and Astronomy/Earth Sciences, where the English translations were more likely to be cited with migration. Given that the codified knowledge was already accessible to US scientists (through translated and international journals) these results suggest that immigrants played a role in transmitting tacit knowledge through face-to-face interactions which facilitated knowledge flows. The field-specific results show that in some fields, the importance of these interactions were greater in accessing codified knowledge that was already available.

6 Conclusions

Given the important implications for innovation and economic growth, a large body of literature has focused on analyzing knowledge spillovers and the resulting productivity impacts. A large literature has also sought to quantify the various economic impacts of immigration, and has provided important insights for policy discussions about the number and type of immigrants that should be entering the US. I contribute to these literatures by providing new evidence on the extent to which high-skilled immigrants are a channel for knowledge diffusion by drawing upon a natural experiment provided by the end of the Soviet Union, when an influx of Soviet scientists and engineers entered the US during the 1990s.

Existing related studies have tended to use the mobility of individuals within the US or have focused on how ethnic communities transmit knowledge back to home countries, and most have used patent data. Using aggregate panel data at the city-field-level and microdata at the paper-level on citations to scientific papers and the location of scientists, the results confirm that immigrants do contribute significantly to knowledge flows. Both approaches show that these flows differ by field, with immigrants in the Life Sciences and Physics contributing the most to knowledge flows. The micro-level analysis with matched paper-pairs allows for a deeper understanding of which types of ideas were more likely to be transmitted to natives. The results show that high impact ideas and ideas already accessible as codified knowledge were more likely to "spill over". This suggests that colocation with immigrants may be especially important for the transmission of tacit knowledge.

I close with a few concluding thoughts on implications and limitations of these findings. First, the two empirical approaches measure the effect of immigration at different levels of geography and in terms of the "local" treatment effect. The similar conclusions drawn from both approaches shows that the findings are robust to the method and unit of analysis. The multi-level approach is also useful given discussions about the proper geographic unit of analysis for estimating the localization of knowledge flows and impacts of immigration on local labor markets.

Second, the differences in the impact of migration across scientific fields suggest that there may be certain types of knowledge or conditions that differ by field that allow cross-border knowledge transfer to flourish. A greater understanding is needed of the differences between the scientific fields that are driving these heterogeneous effects, and the results point to the need for further research on identifying conditions more conducive to knowledge flows through migration than others. These effects also highlight the need for more evidence from different fields and settings in order to generalize about knowledge flows and the resulting productivity spillovers.

In this paper I do not address whether the new ideas brought to the US by Russian scientists impacted the productivity of native scientists, but the findings do point to the possible channels through which such productivity spillovers could occur. This paper also relies on linking paper-to-paper citations and knowledge flows, and in some cases, citations may not represent true knowledge flows. However, this 'paper trail' provides useful clues as to the ways that immigrants can contribute to bringing new ideas into a country.

The end of the USSR, and the influx of scientists that followed had important implications for the US labor market and economy. While an influx of highly-skilled immigrants on a similar scale is not likely to recur, this analysis provides compelling evidence that people can play an important role in diffusing knowledge across borders. The results are especially striking since the 'ideas' in this analysis were already quite dated, and thus the role of immigration is likely even larger than these effects imply.

References

Agrawal, Ajay, Devesh Kapur, John McHale, and Alexander Oettl. 2011. "Brain drain or brain bank? The impact of skilled emigration on poor-country innovation." *Journal of Urban Economics*, 69(1): 43–55.

Agrawal, Ajay, Iain Cockburn, and John McHale. 2006. "Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships." *Journal of Economic Geography*, 6(5): 571–591.

Azoulay, Pierre, Joshua S Graff Zivin, and Bhaven N Sampat. 2012. "The diffusion of scientific knowledge across time and space." *The Rate and Direction of Inventive Activity Revisited*, 107–155.

Berry, Michael J. 1988. *Science and technology in the USSR.* Longman London.

Blackwell, Matthew, Stefano M Iacus, Gary King, and Giuseppe Porro. 2009. "cem: Coarsened exact matching in Stata." *Stata Journal*, 9(4): 524–546.

Borjas, George J. 2003. "The labor demand curve is downward sloping: reexamining the impact of immigration on the labor market." *The Quarterly Journal of Economics*, 118(4): 1335–1374.

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

Breschi, Stefano, and Francesco Lissoni. 2009. "Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows." *Journal of Economic Geography*, 9(4): 439–468.

Card, David. 2001. "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration." *Journal of Labor Economics*, 19(1): 22–64.

Cortes, Patricia. 2008. "The effect of low-skilled immigration on US prices: Evidence from CPI data." *Journal of Political Economy*, 116(3): 381–422.

Feldman, Maryann. 2000. "Location and innovation: the new economic geography of innovation, spillovers, and agglomeration." *The Oxford Handbook of Economic Geography*, 373–394.

Ganguli, Ina. 2011. "Saving soviet science: the impact of grants when government r&d funding disappears." *Harvard University, unpublished manuscript.*

Ganguli, Ina. 2012. "Location and scientific productivity: evidence from the soviet brain drain." *Harvard University, unpublished manuscript.*

Garfield, Eugene. 1972. "Cover-to-cover translation of soviet journals - a wrong solution of the wrong problem." *Current Contents*, 1(29): 334–335.

Glaeser, Edward L. 1999. "Learning in cities." *Journal of urban Economics*, 46(2): 254–277.

Graham, Loren R., and Irina Dezhina. 2008. *Science in the new Russia: crisis, aid, reform.* Indiana University Press.

Holden, Constance. 1990. "No American dream for Soviet emigres." *Science*, 248(4959): 1068.

Hunter, Rosalind S, Andrew J Oswald, and Bruce G Charlton. 2009. "The elite brain drain." *The Economic Journal*, 119(538): F231–F251.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How much does immigration boost innovation?" *American Economic Journal: Macroeconomics*, 2(2): 31–56.

Ivakhnyuk, Irina. 2009. "Russian migration policy and its impact on human development." *Human Development Report Office (HDRO), United Nations Development Programme (UNDP).*

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic localization of knowledge spillovers as evidenced by patent citations." *The Quarterly Journal of Economics*, 108(3): 577.

Kahn, Shulamit, and Megan MacGarvie. 2008. "How important is US location for research in science?" *Boston University Working Paper.*

Kerr, William R. 2008. "Ethnic scientific communities and international technology diffusion." *The Review of Economics and Statistics*, 90(3): 518–537.

Kerr, William R, and William F Lincoln. 2010. "The supply side of innovation: H-1B visa reforms and US ethnic invention." *Journal of Labor Economics*, 28(3): 473–508.

Kogut, Bruce, and J Muir Macpherson. 2011. "The mobility of economists and the diffusion of policy ideas: The influence of economics on national policies." *Research Policy*, 40(10): 1307–1320.

Leydesdorff, Loet, and Ismael Rafols. 2009. "A global map of science based on the ISI subject categories." *Journal of the American Society for Information Science and Technology*, 60(2): 348–362.

Locke, William N. 1956. "Translation by machine." *Scientific American*, 194: 29–33.

Marshall, Alfred. 1920. *Principles of economics.* Macmillan London.

Moody, R Adam. 1996. "Reexamining brain drain from the former Soviet Union." *The Nonproliferation Review*, 3(3): 92–97.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German-Jewish Emigres and US Invention." *NBER Working Paper*, , (w19962).

Romer, Paul M. 1990. "Endogenous technological change." *Journal of Political Economy*, 98(S5): S71–S10.

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

Wagner, Caroline, Irene Brahmakulam, DJ Peterson, Inda Staheli, and Anny Wong. 2002. *US government funding for science and technology cooperation with Russia.* Santa Monica:Rand Corporation.

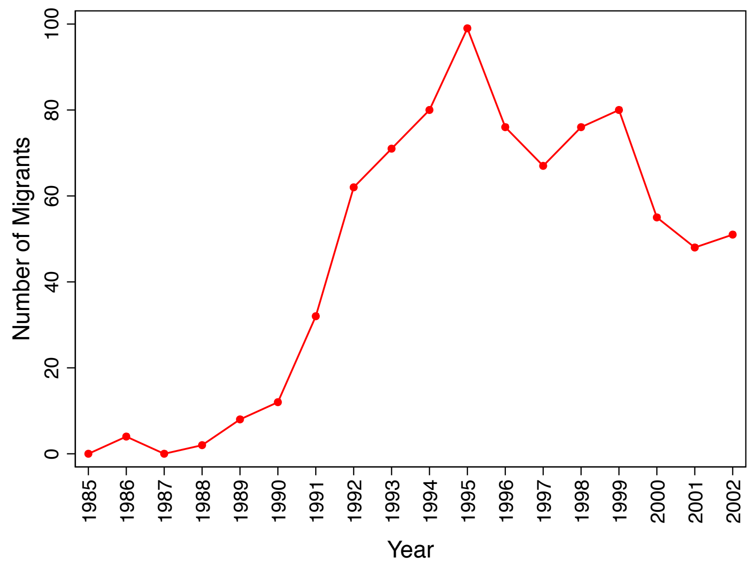
Waldinger, Fabian. 2010. "Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany." *Journal of Political Economy*, 118(4): 787–831.

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

Weinberg, Bruce A. 2011. "Developing science: scientific performance and brain drains in the developing world." *Journal of Development Economics*, 95(1): 95–104.

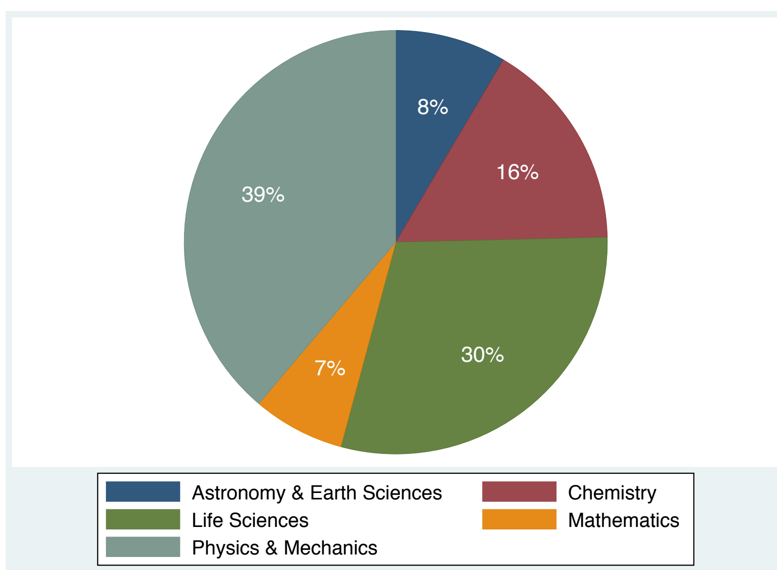
Zucker, Lynne G, and Michael R Darby. 2006. "Movement of star scientists and engineers and high-tech firm entry." *NBER Working Paper Series*, 12172.

Figure 1: Flows of Russian Migrant Scientists



Notes: Only includes Russian scientists who had published in a top Soviet journal from 1986-1990 in the ISI database.

Figure 2: Migrants by Field, post-1992



Notes: Only includes Russian scientists who had published in a top Soviet journal from 1986-1990 in the ISI database.

Figure 3: Year of Publication of Soviet-Era Papers (1980-1990)

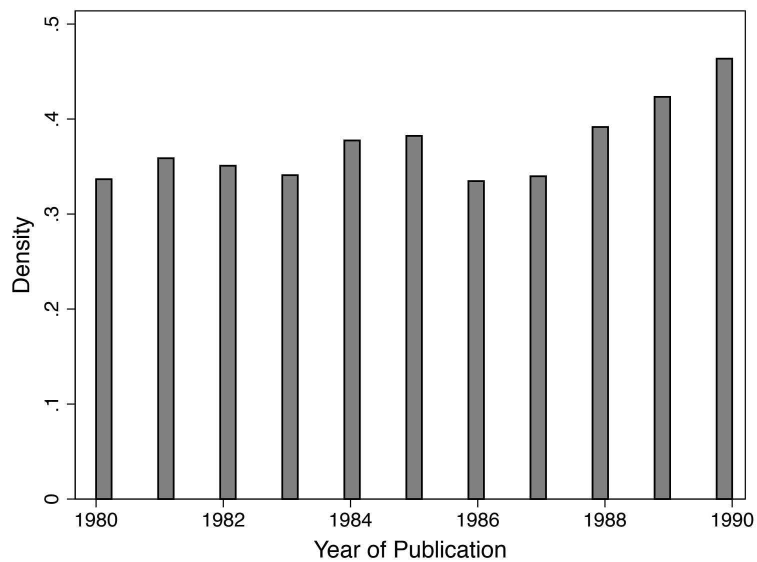
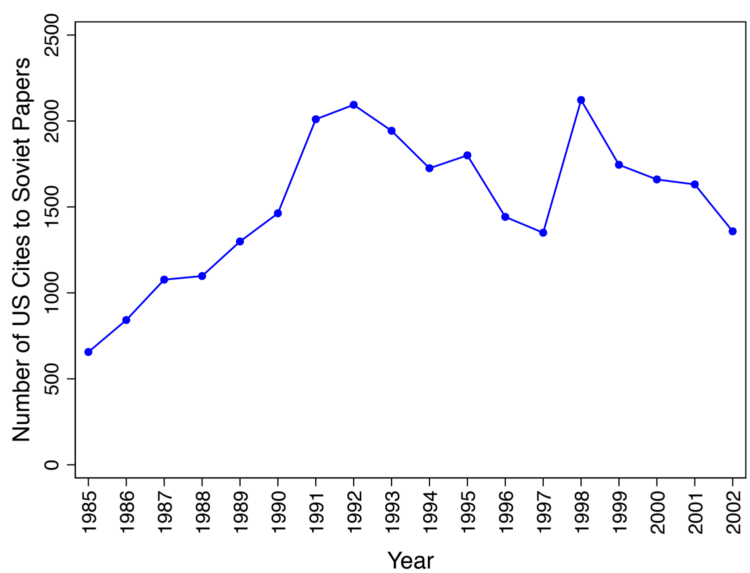
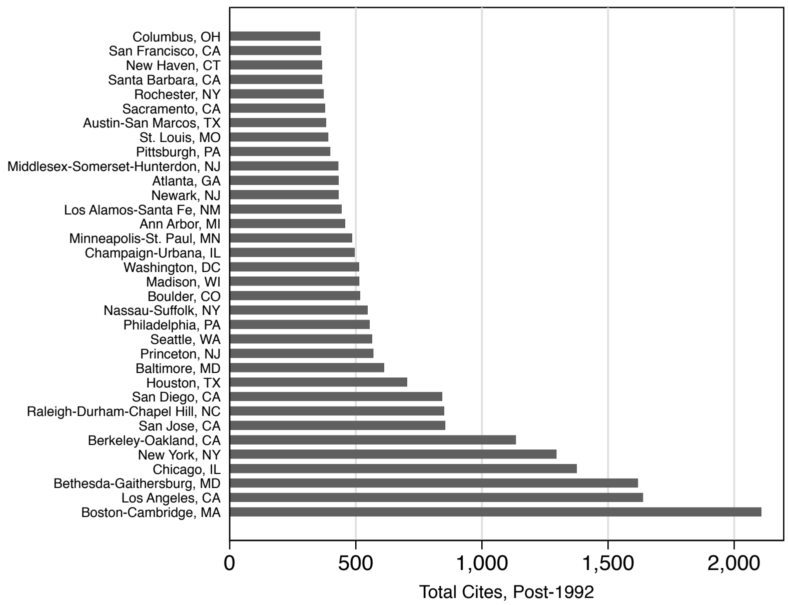


Figure 4: US Citations to Soviet Papers Published 1980-1990



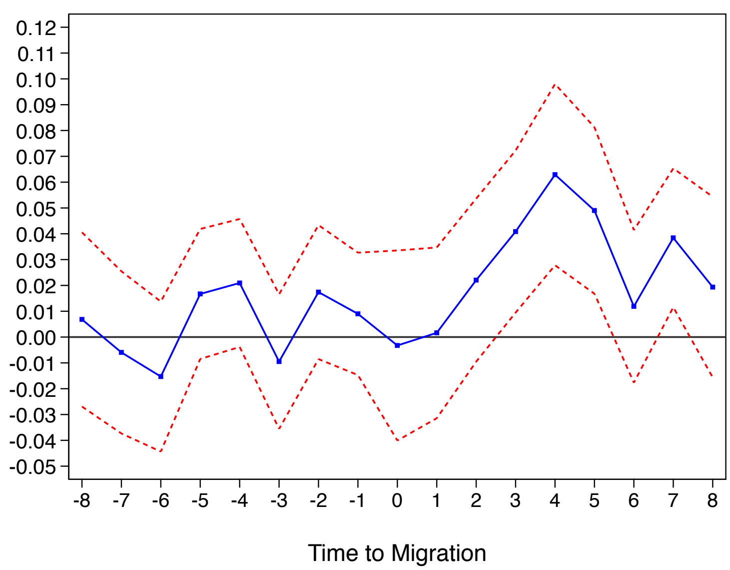
Notes: Citations are to papers by Soviet authors published 1980-1990.

Figure 5: Cities with Most Citations to Soviet Papers, post-1992



Notes: Citations are to Soviet-era papers published 1980-1990.

Figure 6: Effect of Migration on Article Citation Rates



Notes: Difference in US citations between migrant and control matched articles published 1980-1990.

Table 1: Summary Statistics, Full Sample of Articles

	Non-Migrants	Migrants	Difference
Year Article Published	1985.19	1985.77	-0.586***
No. Coauthors	4.40	3.77	0.630***
No. Pubs 1980-1990	34.66	36.07	-1.405**
No. Cites 1980-1990	327.03	783.19	-456.154***
Year of First Pub.	1968.12	1970.67	-2.552***
Non-USSR Coau. Pre-90	0.42	0.59	-0.172***
Astronomy & Earth Sci.	0.08	0.05	0.021***
Chemistry	0.29	0.15	0.141***
Life Sciences	0.26	0.36	-0.093***
Mathematics	0.02	0.05	-0.025***
Physics & Mechanics	0.35	0.39	-0.044***
Moscow Origin	0.69	0.78	-0.087***
Observations (Articles)	19752	2570	

Notes: Stars indicate the results of t-tests for the equality of means.

Table 2: Summary Statistics for CEM Matched Paper Pairs

	Control	Treatment	Difference
Article Age	10.82	10.82	0.000
No. Coauthors	3.86	3.76	0.104
No. Pubs 1980-1990	36.82	35.32	1.496
No. Cites 1980-1990	409.43	419.36	-9.934
Year of First Pub.	1970.46	1970.72	-0.256
Non-USSR Coau. Pre-90	0.50	0.50	0.000
Astronomy & Earth Sci.	0.05	0.05	0.000
Chemistry	0.18	0.18	0.000
Life Sciences	0.31	0.31	0.000
Mathematics	0.03	0.03	0.000
Physics & Mechanics	0.44	0.44	0.000
Moscow Origin	0.79	0.79	0.000
Article's Baseline Stock of US Citations	1.32	1.33	-0.016
Observations (Articles)	1221	1221	

Notes: Stars in the “Difference” column would indicate the results of t-tests for the equality of means.

Table 3: City-Field-Year Regressions: Migrant Inflows and Citations

	OLS				IV	
	(1)	(2)	(3)	(4)	(5)	(6)
No. Migrants	8.127**	16.521**	5.022**	6.256**	55.250**	52.797**
	(1.233)	(2.425)	(0.777)	(1.958)	(12.941)	(16.989)
Field (Life Science omitted)						
No. Migrants X Astronomy & Earth Sci.		-20.667**		-9.268**		
		(2.615)		(2.182)		
No. Migrants X Chemistry		-16.665**		-5.540**		
		(2.516)		(2.111)		
No. Migrants X Math		-24.249**		-11.962**		
		(3.380)		(2.556)		
No. Migrants X Physics & Mechanics		-8.256**		2.052		
		(3.048)		(2.634)		
Tot. Native Papers			0.006**	0.005**		0.001
			(0.001)	(0.001)		(0.002)
Constant	1.581	2.143	-0.306	5.847**	-0.612	-0.746
	(1.639)	(1.507)	(0.821)	(1.447)	(2.206)	(2.046)
R2	0.552	0.590	0.648	0.660	.	.
Nb. of Obs.	9,845	9,845	9,845	9,845	9,845	9,845
Nb. Clust.	895	895	895	895	895	895

Standard errors in parentheses

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

Notes: Dependent variable is citations to Soviet-era publications. Estimates are ordinary least squares (OLS). Observations are at the city-field-year-level for years 1992-2002. All regressions include city, year, field, and year x field fixed effects. Robust standard errors are clustered at the city-field level (179 cities (MSA/PMSA) and 5 fields).

Table 4: Difference-in-Differences, Matched Paper-Pair Regressions

	(1)	(2)	(3)	(4)	(5)
Post-Migration	0.030*	0.050*	0.017	0.039*	0.043+
	(0.013)	(0.025)	(0.013)	(0.016)	(0.025)
Field (Life Science omitted)					
Post-Migration X Astronomy & Earth Sci.		0.024			0.029
		(0.041)			(0.040)
Post-Migration X Chemistry		-0.024			-0.021
		(0.031)			(0.031)
Post-Migration X Math		-0.110*			-0.120*
		(0.052)			(0.051)
Post-Migration X Physics & Mechanics		-0.034			-0.037
		(0.028)			(0.027)
High Impact					
No. Migrants X 80th Decile & Above			0.045+		0.041
			(0.026)		(0.026)
Journal Language (Int'l journals omitted)					
Post-Migration X English Translation				0.002	0.026
				(0.028)	(0.027)
Post-Migration X Russian Language				-0.054*	-0.045*
				(0.022)	(0.021)
Constant	0.013	0.013	0.009	0.011	0.007
	(0.037)	(0.036)	(0.037)	(0.037)	(0.037)
R2	0.001	0.002	0.002	0.002	0.003
Nb. of Obs.	13,016	13,016	13,016	13,016	13,016
Nb. Clust.	422	422	422	422	422

Standard errors in parentheses

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

Notes: Dependent variable is the difference in citations originating in the US for treatment and control matched pairs. Estimates are ordinary least squares (OLS). Sample includes 2,442 total papers (1,221 matched pairs) produced by 422 scientists. Robust standard errors are clustered at the scientist-level. All regressions include pair fixed effects, dummies for age categories and are estimated using 5-year windows around the year of migration (between 1992-2002). "High Impact" indicates cited papers which were in the 80th decile and above for total citations from 1980-1990, calculated by field and year of publication.

Online Appendix A Data Description

In this appendix I include further details about the construction of the dataset.

A.1 Publication Data

To create the sample of scientists who published in the top Soviet and Russian journals, I first identified these journals in the Thomson Reuters ISI Web of Science (ISI). I relied primarily on the list of journals identified by the International Science Foundation (ISF) as "qualifying" journals for potential grantees in Appendix Table B5 (see Ganguli (2011) for more information about the ISF.) In addition to searching journal titles for these titles, I identified the English language translations of these titles. I also checked the journals identified as Russian language journals using the language field in the ISI in case I missed articles published using different English translations.

The ISI publications are not associated with one scientific field, but with rather many scientific subject categories. In order to assign a scientific field to each individual scientist, I did the following. First, I assigned a likely broader scientific field to each of the 221 unique subject categories in the sample of publications based on the search of scientist names. The fields I assigned were one of the 7 major scientific field identified by the ISF (Astronomy, Chemistry, Earth Sciences, Life Sciences, Mathematics, Mechanics, and Physics). Note that when the grantees applied for the grants, they were asked to choose from one of these scientific fields and a number of subfields. Many of the subject categories clearly belonged to a scientific field, e.g. "Cell Biology" was coded as Biology and "Chemistry, Organic" was coded as "Chemistry". For other fields, I used resources that listed field codes along with the scientific field associated with it. For example, Thomson Reuters Journal Citation Report and Essential Science Indicators list subject categories with the broader scientific field, e.g. "Acoustics" is listed under "Physics".

I also compared field codes with the results of analysis presented in Leydesdorff and Rafols (2009), who use exploratory factor analysis of the matrix of field codes in the ISI database to determine the disciplines associated with each subject category. If a subject category could belong to more than one scientific field, I did not code it. Then, to assign a scientific field to each publication, I chose the most common scientific field among the subject categories. Then, for each scientist, I chose the most common scientific field among all the publications he/she published.

A.2 Transliteration & Name Matching Issues

A challenge in matching the scientists to publication data is how names from the Cyrillic alphabet are transliterated into the Latin alphabet. Using a name dictionary Polyglossum 3.71 created by ETS Publishing House (Moscow) that is based on several official standards for transliterations (e.g. ISO 9-1995, OVIR of Russia regulations), I identified possible spellings for each last name and searched for each variant in the publication databases. For example, an example of a surname in my sample in the Latin alphabet is Kuznetsov. This Cyrillic name (Кузнецов) has multiple transliterations, which I identified with the name dictionary:

Кузнецов:
Kuznetsov
Kuznecov
Kouznetsov (*OVIR USSR)

where "OVIR USSR" is the transliteration standard used by the "Office of Visas and Registration" of the USSR. Note that this is not an issue for many names, such as Ivanov, or names from the Baltic countries, as in these countries the languages use the Latin alphabet.

Additionally, typical name ambiguity issues arise, including common names (such as Ivanov, like Smith in the U.S.). I exclude these names from the analysis. I also trim the data by excluding names with more than 500 publications during the period, since the likelihood of common names is higher for these occurrences. The results do not change significantly when these observations are included.

A.3 Comparison of Immigrant Counts with 2000 Census

To benchmark my sample of Soviet scientist migrants with other counts, I used the 2000 Public Use Microdata Sample (PUMS) of the Decennial Census (5% sample). I tallied the number of individuals who immigrated to the US after 1985, were born in the

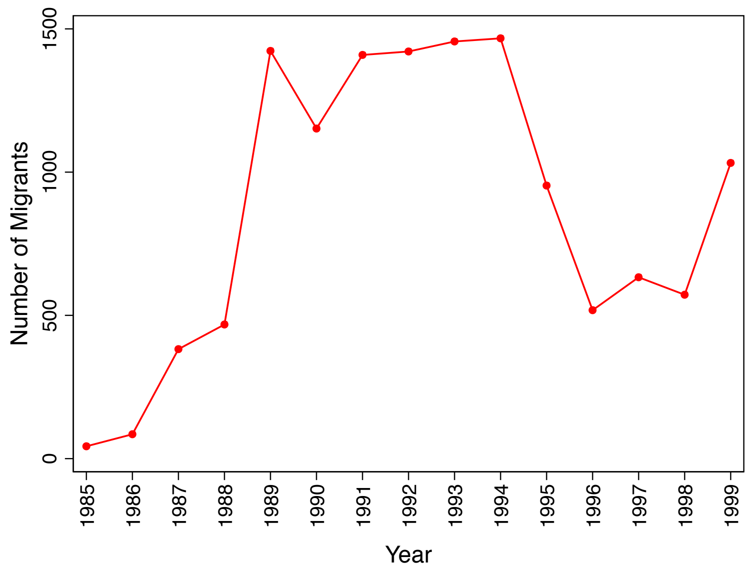
USSR/Russia, were at least 35 at the time of immigration, and reported an occupation of Mathematician, Engineer or Scientist. Appendix Figure B1 shows the counts of these individuals by year of immigration (weighted to be representative of the total population). Comparing the counts to Figure 1, the increase in immigration in the Census sample occurs a few years earlier. Both figures show the decrease in immigrant inflows in 1995/1996, perhaps due to the termination of the Soviet Scientists Immigration Act of 1992, and the subsequent increase in inflows after the 1998 Russian financial crisis.

There are several reasons the Census sample includes many more Soviet scientists than my sample. First, many of the scientists in the Census sample probably were not publishing during the Soviet era. The bulk of researchers in the USSR worked in military or industrial facilities, where publication was not common or allowed, and others in universities, where little research was done due to the separation between teaching and research in the Soviet system (Graham and Dezhina, 2008). According to official Russian statistics, in 1991 there were an estimated 1,079,088 researchers working in R&D organizations in the USSR, and of those, only 134,176 had the equivalent of a PhD (Candidate of Sciences) or higher academic degree (Doctor of Sciences) (**?**). In my sample, scientists had to have published in top Soviet journals or international journals, which meant they were the "cream of the cream". While I could use the Census counts for the city-level analysis, it is not possible to assign these individuals to scientific fields. Moreover, since these individuals did not have publications, it is not possible to track citations to their work after arriving in the US. Finally, my sample includes scientists who were working in the Russian Republic of the USSR, but the Census question asks only the country of birth (USSR/Russia). Thus, the Census sample also likely includes many scientists who were working in other former Soviet Republics.

Online Appendix B Additional Figures & Tables

This section includes additional figures and tables mentioned in the text.

Figure B1: Year of Immigration for Soviet Scientists & Engineers, 5% 2000 PUMS Census Sample



Notes: Counts tallied using the 2000 Public Use Microdata Sample (PUMS) of the Decennial Census (5% sample) using sample weights to be representative of the total population. Included are individuals who immigrated to the US after 1985, were born in the USSR/Russia, were at least 35 at the time of immigration, and reported an occupation of Mathematician, Engineer or Scientist.

Table B1: Cities' Shares of 1990 Soviet-born Immigrants (Top 15)

City (MSA)	Share of Total
New York, NY	28.55%
Los Angeles-Long Beach, CA	18.64%
Chicago, IL	5.15%
Philadelphia, PA	3.80%
Boston, MA	3.48%
San Francisco, CA	3.02%
Newark, NJ	1.66%
Detroit, MI	1.49%
Baltimore, MD	1.48%
Nassau-Suffolk, NY	1.21%
Middlesex-Somerset-Hunterdon, NJ	1.15%
Cleveland-Lorain-Elyria, OH	1.15%
Rochester, NY	0.83%
Anaheim-Santa Ana, CA	0.80%
Seattle-Bellevue-Everett, WA	0.77%

Notes: The numerator includes the total number of 15-65 year olds reporting a birthplace in the USSR in that MSA/PMSA using the 1990 PUMS (State 5% sample). The denominator includes the national total.

Table B2: First Stage Estimates for IV

	All		Life Sci.	Astro./ Earth Sci.	Chem.	Math	Phys./ Mech.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$SovietMigrants_{i1990}$/	0.188**	0.135**	0.083*	0.160	0.071	-0.085	0.264**
$\text{SovietMigrants}_{1990}$	(0.033)	(0.034)	(0.037)	(0.192)	(0.094)	(0.050)	(0.039)
Constant	0.052	0.019	0.303**	-0.016	0.202**	0.104**	0.040
	(0.043)	(0.039)	(0.061)	(0.013)	(0.021)	(0.014)	(0.035)
Total Native Papers	No	Yes	Yes	Yes	Yes	Yes	Yes
Year X Field FE	Yes	Yes	No	No	No	No	No
R2	0.197	0.232	0.443	0.179	0.175	0.180	0.336
Nb. of Obs.	9,845	9,845	1,969	1,969	1,969	1,969	1,969
Nb. Clust.	895	895	179	179	179	179	179

Standard errors in parentheses

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

Notes: Dependent variable is the number of Russian scientist migrants. Observations are at the city-field-year-level for years 1992-2002. Estimates are ordinary least squares (OLS). Columns 1 and 2 are for the entire sample and Columns 3-7 are by field. All regressions include city and year fixed effects, and Columns 1 and 2 include field and year x field fixed effects. Robust standard errors are clustered at the city-field level (179 cities and 5 fields).

Table B3: City-Field-Year Regressions: Excluding New York, NY & Los Angeles, CA

	OLS				IV	
	(1)	(2)	(3)	(4)	(5)	(6)
No. Migrants	7.891**	14.967**	4.733**	4.700*	66.940**	74.347**
	(1.261)	(2.595)	(0.784)	(1.954)	(10.153)	(18.171)
Field (Life Science omitted)						
No. Migrants X Astronomy & Earth Sci.		-19.483**		-8.047**		
		(2.863)		(2.236)		
No. Migrants X Chemistry		-14.973**		-4.046+		
		(2.677)		(2.131)		
No. Migrants X Math		-21.562**		-9.749**		
		(3.235)		(2.461)		
No. Migrants X Physics & Mechanics		-5.972+		4.158		
		(3.226)		(2.697)		
Tot. Native Papers			0.006**	0.006**		-0.002
			(0.001)	(0.001)		(0.002)
Constant	1.714	2.304	-0.172	2.140+	-1.244	-0.916
	(1.698)	(1.553)	(0.817)	(1.099)	(2.563)	(2.894)
R2	0.541	0.575	0.633	0.646	.	.
Nb. of Obs.	9,735	9,735	9,735	9,735	9,735	9,735
Nb. Clust.	885	885	885	885	885	885

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Notes: Dependent variable is citations to Soviet-era publications. Estimates are ordinary least squares (OLS). Observations are at the city-field-year-level for years 1992-2002. All regressions include city, year, field, and year x field fixed effects. Robust standard errors are clustered at the city-field level.

Table B4: CEM Matched Regressions, Alternate Specification at Paper-level

	(1)	(2)
Post-Migration	0.019^{+}	0.019^{+}
	(0.010)	(0.010)
Constant	0.190^{**}	0.183^{**}
	(0.059)	(0.060)
Age Cat Dummies	No	Yes
R2	0.005	0.005
Nb. of Obs.	37,796	37,796
Nb. Clust.	1,283	1,283

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

Notes: This difference-in-differences specification follows Singh and Agrawal (2011) and the data is at the scientist-paper-level. The dependent variable is number of citations in a given year. Estimates are ordinary least squares (OLS). Regressions include paper, year and citation lag fixed effects. Robust standard errors are clustered at the scientist-level. There are 1,283 scientists in the sample, of which 424 are migrants.

Table B5: List of Soviet Journals Used to Identify Soviet Scientists

1. AKUSTICHESKII ZHURNAL
2. ANTIBIOTKI I KHIMIOTERAPIYA
3. ARKHIV PATOLOGII
4. ASTRONOMICHESKII ZHURNAL
5. BIOLOGICHESKIE MEMBRANY
6. BIOLOGIYA MORYA
7. BIOORGANICHESKAYA KHIMIYA
8. BIOFIZIKA
9. BIOKHIMIYA
10. BYULLETEN EKSPERIMENTALNOI BIOLOGII I MEDITSINI
11. VESTNIK AKADEMII MEDITSINSKIKH NAUK SSSR
12. VESTNIK AKADEMII NAUK SSSR
13. VESTNIK MOSKOVSKOGO UNIVERSITEfA SERIYA F1Z1KA I ASTRONOMIYA
14. VESTNIK MOSKOVSKOGO UNIVERSITETA SERIYA KHIMIYA
15. VESTNIK MOSKOVSKOGO UNIVERSITEfA SERIYA MATEMATIKII MEKHAN1KI
16. VOPROSY VIRUSOLOGII
17. VOPROSY MEDITSINSKOI KHIMII
18. VOPROSY ONKOLOGII
19. VYSOKOMOLEKULYARNYE SOEDINENIYA SERIYA A
20. VYSOKOMOLEKULYARNYE SOEDINENIYA SERIYA B
21. GENETIKA
22. GEOMAGNETIZM I AERONOMIYA
23. GEOTEKTONIKA
24. GEOKHIMIYA
25. DIFFERENTSIALNYE URAVNENIYA
26. DOKLADY AKADEMII NAUK BSSR
27. DOKLADY AKADEMII NAUK SSSR
28. ZHURNAL ANALITICHESKOI KHIMII SSSR
29. ZHURNAL VYSSHEI NERVNOI DEYATELNOSTI IMENI I P PAVLOVA
30. ZHURNAL MIKROBIOLOGII EPIDEMILOGII I IMMUNOLOGII
31. ZHURNAL NEORGAATJCHESKOI KHIMII
32. ZHURNAL OBSHCHEI BIOLOGII
33. ZHURNAL OBSHCHEI KHIMII
34. ZHURNAL ORGANICHESKOI KHIMII
35. ZHURNAL STRUKTURNOI KHIMII
36. ZHURNAL TEKHNICHESKOI FIZIKI
37. ZHURNAL FIZICHESKOI KHIMII
38. ZHURNAL EVOLYUTSIONNOI BIOKHIMII I FIZIOLOGII
39. ZHURNAL EKSPERIMENTALNOI I TEORETICHESKOI FIZIKI
40. ZOOLOGICHESKY ZHURNAL
41. IZVESTIYA AKADEMII NAUK SSSR SERIYA BIOLOGICHESKAYA
42. IZVESTIYA AKADEMII NAUK SSSR SERIYA GEOLOGICHESKAYA
43. IZVESTIYA AKADEMII NAUK SSSR SERIYA FIZIKI ATMOSFERI I OKEANA
44. IZVESTIYA AKADEMII NAUK SSSR SERIYA FIZIKI ZEMLI
45. IZVESTIYA AKADEMII NAUK SSSR SERIYA FIZICHESKAYA
46. IZVESTIYA AKADEMII NAUK SSSR SERIYA KHIMICHESKAYA
47. IZVESTIYA VYSSHIKH UCHEBNIKH SERIYA RADIOFIZIKA
48. IZVESTIYA VYSSHIKH UCHEBNIKH ZAVEDENII SERIYA FIZIKA
49. IZVESTIYA VYSSHIKH UCHEBNIKH ZAVEDENII SERIYA KHIMIYA I KHIMCHESKAYA TEKHNOLOGIYA
50. IZVESTIYA SIBIRSKOGO OTIJELENIYA AKADEMII NAUK SSSR SERIYA KHIMICHESKIKH NAUK
51. IZMERITELNAYA TEKHNIKA
52. KARDIOLOGIYA
53. KVANTOVAYA ELEKRONIKA
54. KIBERNETIKA
55. KJNETIKA I KATALIZ
56. KOLLOIDNTI ZHURNAL
57. KOORDINATSIONNAYA KHIMIYA
58. KOSMICHESKAYA BIOLOGIYA I AVIAKOSMICHESKAYA MEDITSINA
59. KRISTALLOGRAFIYA
60. MATEMATICHESKYE ZAMETKI
61. MATEMATICHESKII SBORNJK
62. MIKOLOGIYA I FTIOPATALOGIYA
63. MIKROBIOLOGIYA
64. MOLEKULYARNAYA BIOLOGIYA
65. NEIROFIZIOLOGIYA
66. NEORGANIZHESKYE MATERrALI
67. NEFTEKHIMIYA
68. OKEANOLGIYA
69. OPTIKA I SPEKTROSKOPIYA
70. PARAZITOLOGIYA
71. PISMA V "ASTRONOMICHESKII ZHURNAL"
72. PISMA V "ZHURNAL TEKHNICHESKOI FIZIK"
73. PISMA V "ZHURNAL EKSPERIMENTALNOI I TEORETICHESKOI FIZIKI"
74. POCHVOVEDENIE
75. PRIBORY I TEKHNIKA EKSPERIMENTA
76. PIRKLADNAYA MATEMATIKA I MEKHANIKA
77. RADIOTEKHN1KA I ELEKTRONIKA
78. RADIOKHIMIYA
79. REAKTSIONNAYA SPOSOBNOST ORGANICHESKIKH SOEDINENII
80. SIBIRSKII MATEMATICHESKII ZHURNAL
81. TEORETICHESKAYA I EKSPERIMENTALNAYA KHIMIYA
82. TEORETICHESKAYA I MATEMATICHESKAYA FIZIKA
83. TEORIYA VEROYATNOSTEI I EE PRIMENENIE
84. TEPLOFIZIKA VYSOKIKH TEMPERATUR
85. TERAPEVTICHESKII ARKHIV
86. UKRAINSKII BIOKHIMICHESKII ZHURNAL
87. UKRAINSKII FIZICHESKII ZHURNAL
88. UKRAINSKII KHIMICHESKII ZHURNAL
89. USPEKHI FIZICHESKIKH NAUK
90. USPEKHI KHIMII
91. USPEKHI MATEMATICHESKIKH NAUK
92. FARMAKOLOGIYA I TOKSIKOLOGIYA
93. FIZIKA GORENIYA I VZRYVA
94. FIZIKIA I TEKHNIKA POLUPROVODNIKOV
95. FIZIKA METALLOV I METALLOVEDENIE
96. FIZIKA NIZKIKH TEMPERATUR
97. FIZIKA TVERDOGO TELA
98. FIZIOLOGICHESKII ZHURNAL
99. FIZIOLOGIYA RASTENII
100. FUNKTSIONALNYI ANALIZ I EGO PRILOZHENIE
101. KHIMIKO-FARMATSEVTICHESKII ZHURNAL
102. KHIMICHESKAYA FIZIKA

Notes: List of top Soviet journals reproduced from the International Science Foundation's Individual Emergency Grant application in the Open Society Archives.