Research Policy 48 (2019) 1248–1259

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Does scientist immigration harm US science? An examination of the knowledge spillover channel☆

Ajay Agrawal[a,b], John McHale[c], Alexander Oettl[d,b,*]

[a] University of Toronto, Canada
[b] National Bureau of Economic Research, United States
[c] National University of Ireland, Galway and Whitaker Institute for Innovation and Societal Change, Ireland
[d] Georgia Institute of Technology, United States

ARTICLE INFO

JEL classification:
J61
O31
O33

Keywords:
Immigration
Displacement
Spillovers
Knowledge flows
Scientists
Externalities

ABSTRACT

The recruitment of foreign-trained scientists enhances US science through an expanded workforce but could also cause harm by displacing better connected domestically-trained scientists, thereby reducing localized knowledge spillovers. We develop a model in which a sufficient condition for the absence of overall harm is that foreign-trained scientists generate at least the same level of localized spillovers as the domestically-trained scientists they displace. To test this condition, we conduct a hypothetical experiment in which each foreign-trained displaces an appropriately matched domestically-trained scientist. Overall, we do not find evidence that foreign-trained scientists harm US science by crowding out better-connected domestically-trained scientists, measured by citations by the US scientific community to their publications.

1. Introduction

Innovation relies on access to knowledge. Thus, knowledge flow patterns influence innovation and hence productivity and economic growth. As a result, factors that influence knowledge flow patterns are important to understand. One such factor is the immigration of foreign-trained knowledge workers. In particular, foreign-trained workers may have different types of peer networks that influence how their knowledge travels across time and space. An extensive empirical literature documents that knowledge flows are geographically localized (Jaffe et al., 1993; Thompson and Fox-Kean, 2005) and that this is likely because knowledge flow patterns are influenced by intricate networks of peers who are often co-located (Agrawal et al., 2006; Waldinger, 2010). Thus, it is plausible that if foreign-trained scientists displace domestically-trained scientists, then they could cause overall harm to US science by generating fewer localized spillovers because foreign-trained scientists' relationships and thus knowledge flows are more internationally-oriented. This could occur even if foreign-trained scientists are equally or more productive than the domestically-trained scientists they displace. We examine the possibility of differential knowledge flows here.

In recent decades, US science has become increasingly internationalized, with rapid growth in the number of foreign-born scientists and engineers (Stephan, 2012). Between 2003 and 2013, the number of immigrant scientists increased from 0.7 million to 1.1 million (Lan et al., 2015). In the physical sciences, the number of immigrant scientists increased by 17,000 while the number of US-born scientists actually decreased by 13,000 (Table 1). With a downward-sloping demand curve for scientists and an upward sloping supply curve for US-born scientists, standard market analysis predicts that there will be displacement of US-born scientists (Borjas, 2007; Borjas and Doran, 2012).

A central theme of the economics of immigration literature has been the measurement of wage and employment displacement effects (Borjas, 2005; Kerr and Kerr, 2011; Peri, 2012; National Academies of Sciences, Engineering, and Medicine, 2016). A large body of work has also explored the aggregate productivity effects of immigration (Kerr,

☆ We thank Bill Kerr, Joshua Gans, Avi Goldfarb, Will Strange, Scott Stern, William Lincoln, and Mercedes Delgado as well as seminar participants at Universidad Carlos III de Madrid, Harvard, and National University of Ireland, Galway for helpful comments. This research was funded by the Centre for Innovation and Entrepreneurship at the Rotman School of Management, University of Toronto, the Martin Prosperity Institute, the Social Sciences and Humanities Research Council of Canada, and the Kauffman Junior Faculty Fellowship. All errors are our own.

* Corresponding author.
E-mail addresses: ajay.agrawal@rotman.utoronto.ca (A. Agrawal), john.mchale@nuigalway.ie (J. McHale), alex.oettl@scheller.gatech.edu (A. Oettl).

https://doi.org/10.1016/j.respol.2019.01.005
Received 23 August 2017; Received in revised form 4 January 2019; Accepted 6 January 2019
Available online 07 February 2019

Table 1
Estimated prevalence of US born citizens and immigrants working in US science occupations.

	2003				2013			
	All Scientists	US Born	Immigrants	% Immigrant	All Scientists	US Born	Immigrants	% Immigrant
Computer and mathematical scientists	2,008,000	1,521,000	487,000	24%	2,647,000	1,879,000	769,000	29%
Biological, Agr., and Env. life scientists	444,000	342,000	102,000	23%	638,000	459,000	179,000	28%
Physical and related scientists	315,000	251,000	64,000	20%	319,000	238,000	81,000	25%
Social and related scientists	495,000	440,000	54,000	11%	581,000	497,000	84,000	14%
Total scientists	3,262,000	2,054,000	707,000	22%	4,185,000	3,073,000	1,113,000	27%

Source: Lan et al. (2015).

2008; Cooke and Kemeny, 2017). In the "canonical model" (see, e.g., Borjas, 2014), the existence of aggregate gains from immigration depend on the displacement of native workers. The relatively small aggregate gain implied by this model has led researchers to look for evidence of externalities, especially in the form of knowledge spillovers (Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010; Peri, 2012; Peri et al., 2013; Kahn and MacGarvie, 2016). This has in turn led to an emphasis on peer networks that support knowledge exchange, work that connects to the large body of evidence that documents the importance of local knowledge spillovers (Freeman et al., 2015; Agrawal et al., 2015). But if local networks are critical to knowledge exchange within US science, an inflow of foreign-trained scientists that displaces domestically-trained workers could disrupt local knowledge networks if the foreign-trained scientists are less connected to US science than the domestically-trained scientists they displace. This raises an intriguing additional possibility of harm: US science suffers because foreign-trained scientists are less well-connected to US science than the domestically-trained they displace. Essentially, immigration of foreign-trained scientists could weaken the domestic knowledge networks that are critical to US scientific advancement.

We begin by developing a simple model of the market for scientists. A sufficient condition for the absence of foreign-trained scientist-induced harm to domestic science (as opposed to domestically-trained scientists) is that foreign-trained scientists generate as least the same level of localized spillovers to the US scientific community as the domestically-trained scientists they displace. In other words, there is no differential in the localized spillovers generated by foreign-trained versus domestically-trained scientists. Next, we test this condition by conducting a hypothetical experiment in which each foreign-trained scientist is assumed to fully displace an appropriately matched domestically-trained scientist. We then compare the level of citations by the US scientific community to publications by the now-US-residing foreign-trained scientist versus the matched (hypothetically displaced) US domestically-trained scientist.

Our model adapts the canonical model of immigration into a competitive labor market to allow for knowledge spillovers. Of course, in reality the economic effects of immigration are complex and occur through multiple channels in addition to the knowledge spillover channel. For example, allowing for differences in market power between domestically-trained and foreign-trained scientists will introduce wage effects that differ from the competitive model and thus affect the impact of displacement. As another example, domestically-trained and foreign-trained scientists may be drawing on different knowledge bases so that immigration can alter the size and diversity of the knowledge base available to the domestic economy. We thus stress that we are focusing only on a single channel – the way the immigration of foreign-trained scientists affects the connectivity of domestic knowledge networks and thereby the extent of knowledge spillovers within that network.

In the model, the combination of displacement and differential spillovers could harm US science. However, empirically, when we compare the relative citation patterns of domestically-trained and foreign-trained scientists, we find that although the US scientific community is much less likely to cite the matched foreign-trained scientists while they are in their original home country – the differential in spillovers is significant pre-immigration – their propensity to do so increases dramatically after the scientist immigrates to the US.[1] In fact, the US scientific community is equally likely to cite a foreign-trained scientist and a domestically-trained scientist after the foreign-trained scientist has moved to the US. In the context of our model, this absence of differential spillovers is a sufficient condition for the absence of harm. It is important to note that these are aggregate results. We do observe lower spillovers to the US scientific community from foreign-trained scientists who move to universities with more co-nationals and from those who arrive from non-English speaking countries.

We further extend our analysis to focus on scientists in the right tail of the productivity distribution: "star scientists." A growing literature within the economics of science reports evidence that stars generate a disproportionate level of knowledge externalities (Azoulay et al., 2010; Waldinger, 2010; Oettl, 2012). With this subsample of elite scientists, the foreign-trained scientist-generated spillover deficit disappears even more rapidly. In fact, soon after arriving, star foreign-trained scientists generate more localized spillovers than their domestically-trained peers on average.

We structure the remainder of the paper as follows. In the next section we develop a simple model of the market for scientific labor that provides a useful framework for examining the welfare implications of foreign-trained scientist mobility. The model allows for domestically-trained scientist displacement and differential spillovers from domestically-trained versus foreign-trained scientists. We describe our empirical strategy in Section 4 and our data and matching methodology in Section 5. We present our results in Section 6. We conclude in Section 7 with a discussion of the limitations of our findings.

2. A model of the market for scientists with displacement and differential spillovers

We develop a simple model of the market for scientists in a given country and examine factors influencing the social welfare implications

[1] We define a foreign-trained scientist as one who publishes at least their first paper in a country other than the US and then at some later point begins publishing in the US. We define a domestically-trained scientist as one who publishes their first paper in the US and then continues to publish in the US. Thus, a domestic scientist may be foreign-born. These definitions are appropriate for our purpose as we are concerned with scientists' social networks rather than their country of origin.

of immigration (the arrival of foreign-trained scientists). The model allows for the displacement – or "crowding out" – of domestically-trained scientists as a result of the arrival of foreign-trained scientists. We adopt the ex-ante social welfare perspective of the receiving country and thus ignore the welfare gains to foreign-trained scientists. Social welfare is thus measured by aggregate social surplus accruing to domestically-trained domestic residents; we do not focus on the distribution of that surplus. The model also allows for possible differential spillovers from domestically-trained and foreign-trained scientists. We show it is possible for domestic social welfare to be harmed by immigration as a result of displacement if the difference between domestically-trained and foreign-trained spillovers is large enough, even if immigration expands the overall size of the active scientific workforce. Of course, the extent of immigration-induced displacement that occurs is a strongly contested question in the literature (Borjas, 2005, 2007; Kerr and Kerr, 2011; Peri, 2012; National Academies of Sciences, Engineering, and Medicine, 2016). We side-step this important debate when we assume the worst-case scenario of full displacement. However, even with full displacement, we show that a sufficient condition for the absence of harm through the spillover channel is the absence of a difference between the spillovers from a foreign-trained scientist and the domestically-trained scientist they are assumed to displace

2.1. Basic market setup

We begin with specifications for labor supply and labor demand in the market for scientific labor. For simplicity, we assume that the units of labor are homogenous and each unit is a working scientist, although we later allow for differential spillovers between domestically-trained and foreign-trained labor units.[2] The supply of domestically-trained scientists, L^s_{domestic}, is a positive linear function of the wage, w:

$$L^s_{\text{domestic}} = \phi_0 + \phi_1 w. \tag{1}$$

Foreign-trained labor units, I, are supplied perfectly inelastically[3] (possibly due to visa-related limitations), so the total supply of labor is:[4]

$$L^s_{\text{total}} = \phi_0 + \phi_1 w + I. \tag{2}$$

Total labor demand, L^d, is a negative function of the wage:

$$L^d = \theta_0 - \theta_1 w. \tag{3}$$

The inverse of the labor demand function is also the marginal private value function. However, we also assume that there are positive spillovers associated with each unit of scientific labor employed. The per-scientist spillover (or externality) is equal to z (≥ 0), which is initially common across domestically-trained and foreign-trained scientists. The marginal social value relationship is then given by:

$$\text{MSV} = \frac{1}{\theta_1}(\theta_0 - L) + z. \tag{4}$$

[2] The model is easily extended to allow for broader heterogeneity by defining labor units in efficiency (i.e., productivity-adjusted) units. Spillovers then also would be measured per efficiency unit, so that more productive scientists are assumed to generate more spillovers.

[3] The assumption here is that the number of immigrants admitted is determined by a government policy choice (e.g. a cap on visas that are issued) and thus the number admitted is not responsive to the wage. However, all the qualitative results of the model are the same if we assume that the supply curve of immigrants is upward sloping. In particular, it remains true that a sufficient condition for the absence of harm is that there is that there is no difference in the domestic spillovers from domestic and immigrant scientists.

[4] In an efficiency-unit version of the model, the level of immigration is also measured in efficiency units.

2.2. Baseline social surplus in the absence of immigration

As a preliminary step to establishing the effects of immigration of foreign-trained scientists on the market for scientific labor, we first examine the market equilibrium and social welfare in a no-immigration baseline. We graph the market equilibrium in Fig. 1. The equilibrium wage and employment levels are:

$$w^* = \frac{\theta_0 - \phi_0}{\phi_1 + \theta_1}, \tag{5}$$

$$L^* = \frac{\phi_0\theta_1 + \phi_1\theta_0}{\phi_1 + \theta_1}. \tag{6}$$

Total social surplus from trade in the scientific labor market is the area between the inverse labor supply curve and marginal social value curve up to the equilibrium quantity of labor. This surplus is equal to:

$$\begin{aligned} S^* &= \int_0^{L^*} \left[\tfrac{1}{\theta_1}(\theta_0 - L) + z - \tfrac{1}{\phi_1}(L - \phi_0)\right]\text{dL}. \\ &= \left(\tfrac{\phi_0\theta_1 + \phi_1\theta_0}{\phi_1 + \theta_1}\right)\left[\left(\tfrac{\phi_0\theta_1 + \phi_1\theta_0}{2\phi_1\theta_1}\right) + z\right]. \end{aligned} \tag{7}$$

The total social surplus is given by the sum of areas A, B, and C in Fig. 1. The existence of the positive externality means that the market equilibrium employment level is lower than the efficient (i.e., social-surplus-maximizing) level, where the latter is determined by the intersection between the labor supply curve and the marginal social value curve.

2.3. Social surplus with immigration but with identical spillovers for domestically-trained and foreign-trained scientists

We next allow for positive immigration but initially assume that spillovers, z, are identical for domestically-trained and foreign-trained scientists. We graph this case in Fig. 2. The new equilibrium wage and employment levels are:

$$w^{**} = \frac{\theta_0 - \phi_0 - I}{\phi_1 + \theta_1}, \tag{8}$$

$$L^{**} = \frac{\phi_0\theta_1 + \phi_1\theta_0 + \phi_1 I}{\phi_1 + \theta_1}. \tag{9}$$

It is also useful to identify the employment level of domestically-trained scientists at the new equilibrium with immigration:

$$L^{***} = \phi_0 + \phi_1 w^{**} = \frac{\phi_0\theta_1 + \phi_1\theta_0 - \phi_1 I}{\theta_1 + \phi_1}. \tag{10}$$

Notice that the domestic displacement is equal to:

$$L^* - L^{***} = \frac{\phi_1}{\phi_1 + \theta_1} I. \tag{11}$$

There is no displacement if ϕ_1 is equal to zero, so that the domestically-trained labor supply is perfectly inelastic. To determine total social surplus, it is useful to separate out the surplus due to domestically-trained versus foreign-trained scientists. Using Eq. (10), the part due to domestically-trained scientists is given by:

$$\begin{aligned} S^{**}_{\text{domestic}} &= \int_0^{L^{***}} \left[\tfrac{1}{\theta_1}(\theta_0 - L) + z - \tfrac{1}{\phi_1}(L - \phi_0)\right]\text{dL} \\ &= \left(\tfrac{\phi_0\theta_1 + \phi_1\theta_0 - \phi_1 I}{\phi_1 + \theta_1}\right)\left[\left(\tfrac{\phi_0\theta_1 + \phi_1\theta_0}{2\phi_1\theta_1}\right) + z + \tfrac{I}{2\theta_1}\right] \\ &= S^* - \left(\tfrac{\phi_1 z}{\phi_1 + \theta_1}\right)I - \left(\tfrac{\phi_1}{2\theta_1(\phi_1 + \theta_1)}\right)I^2, \end{aligned} \tag{12}$$

where the last line makes use of Eq. (7).

Because we are taking the perspective of the welfare of the receiving country, we exclude the surplus accruing directly to foreign-trained scientists. Domestic social surplus accruing from immigrants is thus the

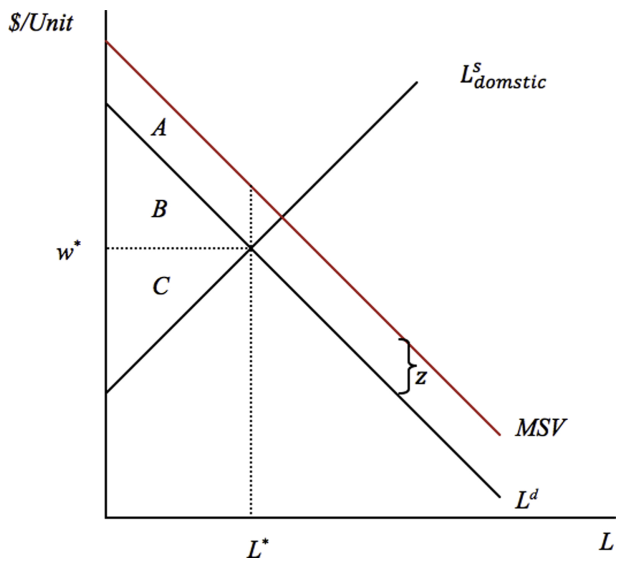


Fig. 1. Market equilibrium and total social surplus, no immigration.

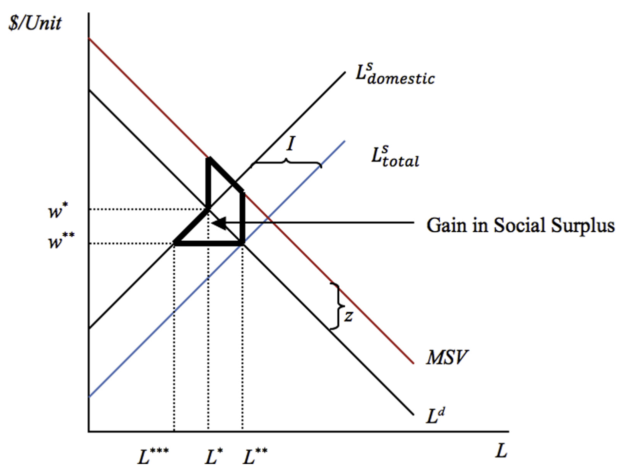


Fig. 2. Market equilibrium and the gain in social surplus from immigration. *Note*: The per-scientist externality is assumed to be equal to z for domestic and immigrant scientists.

difference between the marginal social value curve and the post-immigration wage line (Eq. (8)), where it is assumed that foreign-trained scientists are the marginal labor suppliers. This surplus is given by:

$$\begin{aligned} S^{**}_{\text{immigrant}} &= \int_{L^{***}}^{L^{**}} \left[\frac{1}{\theta_1}(\theta_0 - L) + z - w^{**})\right] \text{dL}. \\ &= \text{z I} + \left(\frac{1}{2\theta_1}\right) I^2. \end{aligned} \tag{13}$$

Total social surplus is found by summing the two components. After some cancellation, this yields:

$$S^{**}_{\text{total}} = S^{**}_{\text{domestic}} + S^{**}_{\text{immigrant}} = S^{*} + \left(\frac{\theta_1 z}{\phi_1 + \theta_1}\right) I + \left(\frac{1}{2(\phi_1 + \theta_1)}\right) I^2. \tag{14}$$

Noting that total social surplus depends positively on both the level and the square of the level of immigration, the surplus is increasing at an increasing rate with the level of immigration. The size of the gain will also depend positively on the size of the per-unit spillover, z, with a positive interaction between the size of the spillover and the level of immigration. The gain in social surplus is shown by the area enclosed by the dark black line in Fig. 2.

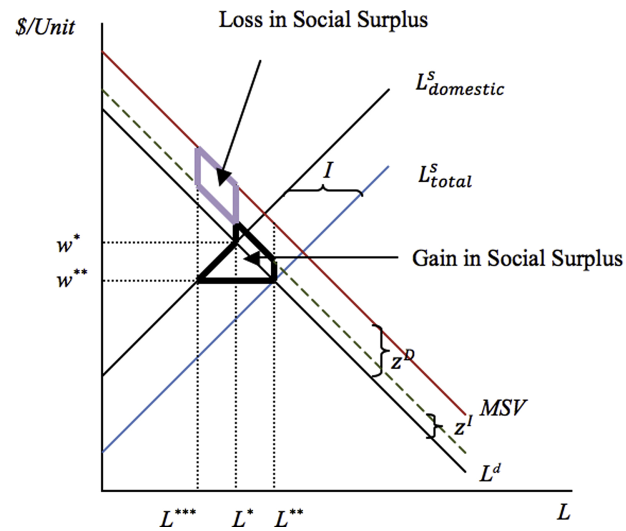


Fig. 3. Market equilibrium and the gain and loss of social surplus when the per scientist externality is lower for immigrant scientists.

2.4. Social surplus with immigration but with differential spillovers for domestically-trained and foreign-trained scientists

We next examine the case where the spillover from domestically-trained scientists, $z^{D}(\geq 0)$, differs from the spillover from foreign-trained scientists, $z^{I}(\geq 0)$, where it is assumed that $z^{D} \geq z^{I}$. The total social surplus is now:

$$S^{**}_{\text{total}} = S^{**}_{\text{domestic}} + S^{**}_{\text{immigrant}} = S^{*} + \left(\frac{\theta_1 z^{I} - \phi_1 (z^{D} - z^{I})}{\phi_1 + \theta_1}\right) I + \left(\frac{1}{2(\phi_1 + \theta_1)}\right) I^2. \tag{15}$$

Compared to the case of equal spillovers, an examination of Fig. 3 shows a loss of social surplus on units that would have been supplied by domestically-trained scientists in the absence of displacement. The lower spillovers from foreign-trained scientists also reduces the size of the gain from immigration, although there is still a direct gain in social surplus that is increasing non-linearly in the level of immigration. The overall impact on social surplus will depend on the relative sizes of these gains and losses. If the gap between z^{D} and z^{I} is large enough, it is possible that the displacement of domestically-trained scientists reduces social surplus overall, notwithstanding the larger total size of the scientific workforce.

We now can identify from Eq. (15) two distinct sufficient conditions for immigration of foreign-trained scientists not to reduce domestic social surplus given any level of immigration (i.e., for $S^{**}_{\text{total}} \geq S^{*}$). First, there will be no harm if there is no domestic displacement, i.e., $\phi_1 = 0$. Second, and central to the empirical part of the paper, there will be no harm if there is no difference between the domestically-trained and foreign-trained spillover, i.e., $z^{D} - z^{I} = 0$.

Using Eq. (15), we also can identify the necessary and sufficient condition for the absence of harm from immigration. This condition is:

$$z^{I} \geq \frac{\phi_1 z^{D}}{(\phi_1 + \theta_1)} - \left(\frac{1}{2(\phi_1 + \theta_1)}\right) I. \tag{16}$$

The "break-even" level of immigrant spillover is then the level of z^{I} at which Eq. (16) holds with equality. We graph the break-even in Fig. 4 as a function of the level of immigration. The break-even level is declining in the level of immigration, reaching zero at an immigration level of $2\phi_1 z^{D}$. Given that the size of the foreign-trained spillover is assumed to be bounded from below at zero (i.e., the spillover is not negative), any foreign-trained immigration level above this level is

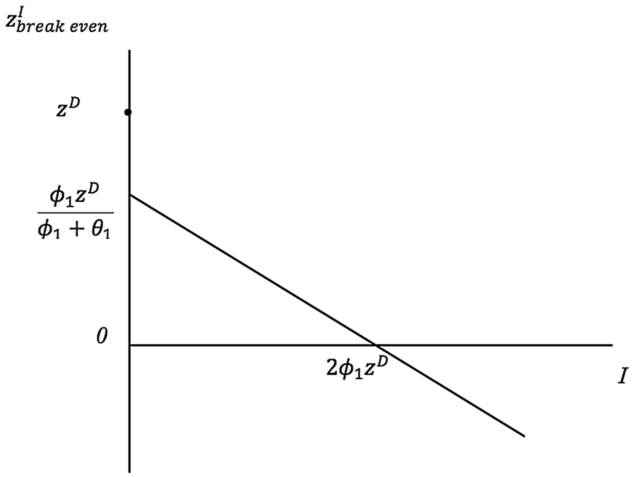


Fig. 4. The level of the per-scientist externality for immigrant scientists for no change in social surplus to occur as a result of immigration.

associated with a net benefit regardless of the level of domestically-trained displacement.

Summing up this section, we have found in the context of a simple market model with spillovers that it is possible that immigration of foreign-trained scientists harms domestic social welfare (as measured by the total surplus accruing to ex ante domestic residents from trade in the scientific labor market). This result requires both the displacement of domestically-trained scientists by foreign-trained scientists and lower spillovers from foreign-trained scientists compared with domestically-trained counterparts. However, the size of the spillover required from foreign-trained scientists to avoid immigration harming social welfare is decreasing in the level of immigration. Notwithstanding displacement effects, a sufficient condition for foreign-trained scientist immigration not to reduce ex ante domestic social welfare in the model is therefore an absence of differential spillovers.

As presented, the model applies to the general market for scientists. One could apply a narrower version to the segment of the market limited to employment at leading research universities. Displacement is then more naturally thought of as domestically-trained scientists moving to lower-ranked universities, as found for example in Borjas and Doran (2012) as a result of the inflow of ex-Soviet mathematicians. In this case, we still would expect spillovers from displaced domestically-trained scientists. However, if we assume that a faculty position in a leading university provides a privileged position in terms of the opportunities for relationship/network development[5] – and that domestically-trained scientists are culturally or linguistically better positioned to take advantage of those opportunities – then downward institutional displacement could still be associated with a loss of aggregate spillovers and social welfare that again must be weighed against the direct gains from immigration of foreign-trained scientists. The search for evidence on possible differential spillovers from domestically-trained and foreign-trained scientists motivates the empirical work in the remainder of the paper.

[5] For example, positions at leading universities may provide faculty members with more graduate students. The pool of former graduate students then becomes a natural pool for matching with collaborators. In Agrawal et al. (2015), we develop a model in which scientists form the best match from the pool of former graduate students. Even where each potential former graduate student collaborator is drawn from a given uniform distribution, simply having more graduate students – and thus more draws – increases the expected value of collaboration. We then show that improvements in collaboration technology, which we assume to scale up the value of collaboration, are more valuable for scientists with more graduate students and thus more draws from which to find the best match.

3. The prevalence and impact of immigrant scientists on US science

Immigrant scientists make up a large and growing proportion of the US scientific workforce. Table 1 shows estimates of the size of the workforce in science occupations in 2003 and 2013. The estimates are drawn from the National Science Foundation's Scientists and Engineers Statistical Data System (SESTAT), which is in turn based on the National Survey of College Graduates (NSCG) and the Survey of Doctoral Recipients (SDR). Overall, immigrant scientists made up an estimated 22% of the US science workforce in 2003, rising to 27% in 2013. The largest immigrant percentages are in non-social science fields, with immigrants in 2013 comprising 29% of computer and mathematical scientists, 28% of biological, agricultural and environmental life scientists, and 25% of physical and related scientists. The largest increase in share over the period was for physical and related scientists, which rose from 20% in 2003 to 25% in 2013.

A significant literature has explored the performance of immigrant scientists and engineers in the US labor market and also their impact on US economic performance. Hunt and Gauthier-Loiselle (2010) observe both positive direct and indirect effects of highly skilled immigrants on US innovation. Using the 2003 National Survey of College Graduates they find that immigrants patent at roughly double the rate of natives. A one percentage point increase in the college-educated immigrant share in the population leads to a 6% increase in patents per capita. However, they note that this could overstate the effects of immigration if there are displacement effects, or understate it if there are positive spillovers. Using a state-level analysis to correct for these biases, they find that a one percentage increase in the share of immigrant college graduates in the population leads to a 9–18% increase in patents per capita. Kerr and Lincoln (2010) find that cities and firms that disproportionately utilized H-1B visa holders increased employment and patenting relative to peers. Peri (2012) finds evidence that immigrants to the US have increased total factor productivity. Combining various impacts, he finds that an increase in 1% in high-skilled population in a state due to immigration, increases income per worker by 1% in that state. Peri et al. (2013) report that a H-1B driven increase in science, technology, engineering, and mathematics (STEM) workers increases the wages of both STEM and non-STEM workers at the city level. Using data from the American Community Surveys of 2009 and 2010, Hunt (2015) finds that immigrant engineers that are successful in finding work in engineering occupations outperform the native-born based on both educational attainment and earnings. Piecing together various data sources, Stephan (2010) documents the high prevalence and impact of immigrant scientists at US Universities. A "conservative estimate" is that immigrants comprise "at least 25% of tenure-track faculty and make up over 43% of graduate students and 60% of post docs" (Stephan, 2010, p. 85). She also estimates that 44% of first authors on papers are immigrants. This impact is also apparent for the most prestigious publications; for example, 44% of first authors of US papers in Science are immigrants.

Recognizing that foreign-trained scientists may be less connected to other US scientists than the domestically-trained counterparts they potentially displace, the focus of our empirical analysis is a hypothetical experiment in which a foreign-trained scientist is assumed to fully displace a matched US scientist with the same observed characteristics. In doing so we are able to hold constant all other characteristics that may affect the likelihood of producing spillovers (career age, productivity, etc.) and focus solely on the variation in peer networks between foreign-trained and domestically-trained scientists. Indirect evidence suggests that US scientists do successfully integrate into US-based knowledge networks. In an important paper, Ganguli (2015) finds that when Soviet-era scientists moved to the US, citations by US scientists to their Soviet-era work increased, indicating they were able to successfully form connections to the US-based scientific community.

We finally reiterate that the assumption of full displacement of a

Table 2
Descriptive statistics.

Discipline	Journals	Papers	Scientists	Domestic trained	Foreign trained	Citations/scientist/year	Coauthors/scientist/year
Panel A: full sample							
Economics	214	105,305	18,466	10,302	552	8.38	0.39
Evol. biology	42	55,035	9,619	4,497	286	18.76	0.74
Immunology	175	586,424	84,649	35,281	3,311	16.17	2.59
Mathematics	190	126,535	22,156	7,644	1,065	3.67	0.42
Neuroscience	247	678,572	91,405	38,074	4,209	19.14	2.14
Psychology	71	49,316	9,805	5,495	218	6.9	0.67
Total	939	1,601,187	236,100	101,293	9,641	12.17[a]	1.16[a]
Panel B: star sample							
Economics	214	29,727	1,324	1,058	101	34.45	0.72
Evol. biology	42	14,866	755	458	49	59.72	1.21
Immunology	175	131,385	7,220	4,094	687	53.71	4.71
Mathematics	190	39,369	1,653	893	214	12.06	0.76
Neuroscience	247	144,420	7,129	3,902	799	61.72	3.83
Psychology	71	16,530	801	548	46	20.58	1.00
Total	939	376,297	18,882	10,953	1,896	49.69[a]	3.34[a]

Notes: *Scientists* refers to the total number of scientists active in the world. *Domestics* refers to the number of US-based scientists who started their careers in the US. *Immigrants* refers to the number of US-based scientists who emigrated to the US. Note that *Domestics* and *Immigrants* do not sum to *Scientists* because we do not report counts of scientists in the rest of the world who do not emigrate to the US during our study period. The last two columns count the mean number of citations received/unique coauthors per scientist per year.

[a] Means, instead of sums, are reported for these two columns.

domestically-trained US scientist is an extreme assumption. It is likely that the number of displaced scientists is less than the number of foreign-trained scientists immigrating. Based on the findings of Borjas and Doran (2012), an empirically relevant but less dramatic mode of displacement is for the displaced US scientists to find employment at lower-ranked institutions or to less competitive fields Borjas and Doran (2015). To the extent that working at highly ranked institutions provides a privileged position of access to US knowledge networks and US-born scientists are better able to avail of the resulting opportunities for knowledge exchange, this form of displacement could also negatively impact on the performance of US science. However, such downward displacement should have less of an adverse effect than the full displacement baseline, again suggesting that the assumption of full displacement provides a reasonable estimated upper bound of the extent of harm through impaired US knowledge networks.

4. Empirical strategy

In the model, a sufficient condition for the absence of harm is that the US scientific community draws equally from knowledge generated by foreign-trained recruits as they do from knowledge generated by the domestic scientists the foreign-trained immigrants may displace. This holds true even with full displacement. Our empirical strategy is to conduct a hypothetical experiment in which a foreign-trained recruit displaces a matched domestically-trained scientist, where we match scientists based on career age (years since first publication), productivity (number of citation-weighted publications), and field (six distinct fields described below). We then examine the number of citations from the US scientific community – our measure of localized knowledge flows – to foreign-trained scientists compared to their matched domestically-trained counterparts.

The measurement of knowledge flows through forward citation data has long been debated in the literature (see, e.g., Jaffe and Trajtenberg (2002), and – for a skeptical view – Breschi and Lissoni (2001)). The chief advantage of citation data is that they leave a "paper trail" Feldman (2000). We recognize that that a citation is a noisy (if still informative) measure of knowledge flow. A knowledge flow can take place even if there is no citation, as when an author fails to acknowledge their use of prior work; and a citation may take place even when no knowledge flow has occurred, as when the citation is made as an "appeal to authority" for an argument. However, for our core empirical question the noisiness of citations should not be a significant problem provided that there is no difference in how well they proxy for knowledge flows for domestically-trained compared to foreign-trained scientists.

As a preliminary step, we first examine the number of citations from the US scientific community to immigrants before they move compared to their domestic matches. Next, turning our attention to how knowledge flows from immigrants to the US scientific community change over time, we compare the number of citations from the US scientific community to immigrants before versus after the immigrants move to the US. Finally, focusing on differential knowledge flows that may provide evidence of immigrant harm to US science – the core of our analysis – we examine the number of citations from the US scientific community to immigrants compared to their domestic matches. A finding of no difference would be consistent with the hypothesis of no harm to domestic science even with full displacement.

5. Data and matching methodology

Our primary objective is to compare spillover patterns between domestically-trained and foreign-trained scientists. Thus, we must identify scientists, their type (domestically-trained or foreign-trained), and their spillovers. We use bibliometrics data to do this. Our primary source is the ISI Web of Science (WoS). We begin by collecting publications in six fields: (1) evolutionary biology, (2) mathematics, (3) economics, (4) neuroscience, (5) immunology, and (6) psychology. We collect all publications in the journals classified by the ISI Journal Citation Reports as being associated with each of those fields. In Table 2, Panel A, we list the number of journals associated with each field and the number of papers we collect from this set of journals over the period 1979–2008. In terms of the number of publications, neuroscience and immunology are the two largest fields (825,048 and 639,439 papers, respectively) and evolutionary biology and psychology are the two smallest (114,190 and 191,333 papers, respectively). We identify 9,641 foreign-trained scientists that moved into the US, which

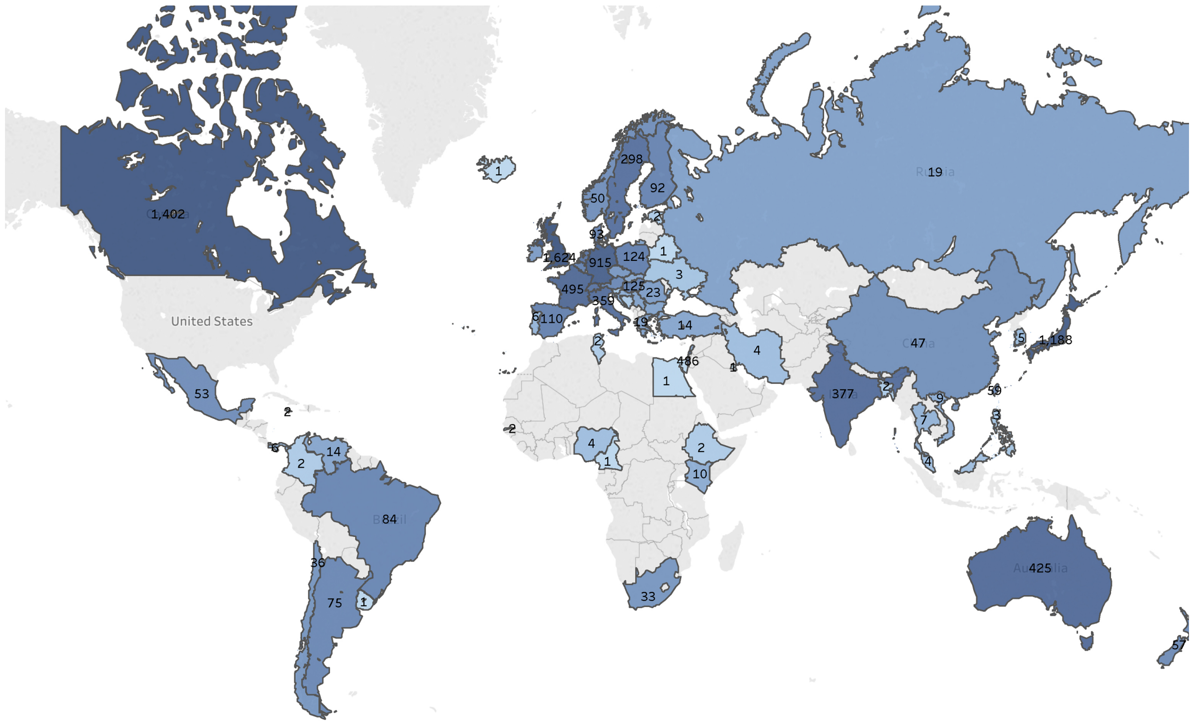


Fig. 5. US scientist immigration by source country. $N = 9641$.

represents 4.1% of our sample of all active scientists across the globe. We present descriptive statistics of our star subsample in Panel B of Table 2. For the star subsample, we identify 1896 foreign-trained scientist who make up 10% of the sample of all global star scientists. This is consistent with the US attracting a disproportionate share of high-quality foreign-trained scientists.

Fig. 5 presents a map showing the origins of all foreign-trained scientists in our sample. While the top 5 origin countries (in order: United Kingdom, Canada, Japan, Germany, and France) are responsible for over 58% of all US foreign-trained scientists, we observe emigration from a total of 69 countries in our sample.

5.1. Identifying scientists

We conduct most of our analyses at the scientist-year level. So, using the publication data described above, we identify the set of scientists in each of the six fields. One data challenge with this process is that WoS data do not provide unique identifiers for scientists. In other words, the data do not distinguish between two different people who have the same name. Thus, we must disambiguate scientific authors. To do so, we employ an approach developed by Tang and Walsh (2010). The heuristic utilizes backward citations of focal papers to estimate the likelihood of the named author being a particular person. For example, if two papers reference a higher number of the same papers (weighted by how many times the paper has been cited, i.e., how popular or obscure it is), then the likelihood of those two papers belonging to the same author is higher. We attribute two papers to the same author if both papers cite two or more rare papers (fewer than 50 citations) in both papers. We repeat this process for all papers that list non-unique author names (i.e., same first initial and last name). We exclude scientists who do not have more than two publications linked to their name. In Table 2, we list the number of unique scientists we identify in each field. Once again, immunology and neuroscience are the two largest fields (84,649 and 91,405 scientists, respectively). The two smallest fields are evolutionary biology and psychology (9619 and 9805 scientists, respectively). Scientists enter the panel when they publish their first paper. We identify their location and status (star or not) on an annual basis.

5.2. Defining stars

We define stars as scientists in the 90th percentile in a given year and discipline in terms of their accumulated stock of citation-weighted paper output over the preceding years. To calculate a scientist's accumulated stock of citation-weighted paper output, we begin by identifying the set of papers they published in the years preceding the focal year. We then weight these papers by the number of citations they receive during our study period. For example, if a scientist published four papers by 1990 and these papers received 10, 20, 15, and 40 citations by 2008 (the final year of our study period), then that scientist's accumulated stock of citation-weighted paper output would be 85 in 1990. While we define a scientist's contribution on an annual basis, our measure of stardom is time-invariant whereby we classify a scientist as a star if the scientist has ever been above the 90th percentile (approximately 15% of scientists).[6] Furthermore, stars are defined relative to the other scientists in our sample in the same discipline. When we do analyses of the full sample (across all disciplines), we utilize the star categorization determined from the within-field analysis. Although citation practices vary across fields, scientists in the 90th percentile are disproportionately more productive than the median scientist across all fields as seen in Fig. 6.

5.3. Identifying scientist locations

Using the unique author identifiers generated in the process described above for each paper, we then attribute each scientist to a

[6] Results are very similar if we conduct our analyses using a time-varying definition of star scientists whereby we only classify scientists as stars in years in which their stock of citation-weighted paper output exceeds the 90th percentile.

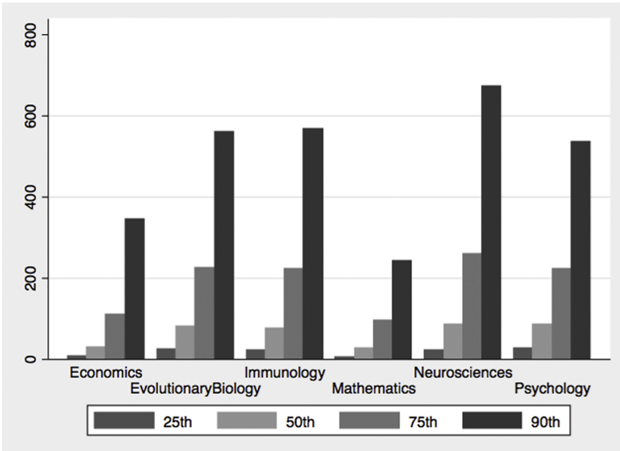


Fig. 6. Citation stock by percentiles and field in 1995.

particular institution for every year of activity. Scientists are active from the year they publish their first paper to the year they publish their last paper. Here again, we must overcome a data deficiency inherent within the WoS data; until recently, the WoS did not link institutions listed on an article to the authors. Instead, we impute author location using reprint information that provides a one-to-one mapping between the reprint author and the scientist's affiliation. In addition, we take advantage of single institution publications that allow us to directly link authors to institutions.

5.4. Defining foreign-trained scientists

With information on each scientist's location in each year, we identify the country of each scientist's institution. Domestically-trained scientists are those who start their career in the US and never emigrate. Foreign-trained scientists are those who start their career in a country other than the US and some year after their first publication emigrate to the US.

5.5. Outcome measure

Our outcome measure of interest is knowledge flows. We identify all papers published by the focal scientist in the focal year for each scientist-year. From this set of papers, we count the number of forward citations (citations made to the focal paper by other papers in the future). We classify each forward citation as domestic if the first author of the future paper that references the focal paper is from the US and not-domestic otherwise.[7]

While a large literature exists on the localization of knowledge flows (Feldman, 2000; Ganguli, 2015), we choose to circumscribe the flow of citations at national borders for two reasons. First, this study directly speaks to the debate in immigration policy which is set at the national level. The US federal government as an entity cares about growing the spillover "pie" within the US and is less concerned about how the benefits of this pie are allocated across space within the US, per se. Second, and relatedly, federal science funding also adopts a very similar model by which it cares about advancing the US scientific enterprise in aggregate and is less focused on distributional concerns.

5.6. Matching

Foreign-trained and domestically-trained scientists may systematically differ along a range of dimensions hindering insightful comparisons between the two groups. As such, we identify a subset of both foreign-trained and domestically-trained scientists who are on the common support of a vector of covariates related to scientific productivity in the year of the immigrant's move to the US. More specifically, for all foreign-trained scientists who move to the US in year *t*, we identify a domestically-trained scientist match in year *t* who is in the same field, has a similar quality-weighted stock of publications, was equally as productive in year *t*, and has a similar career age.[8] We make use of the of the Coarsened Exact Matching (CEM) methodology first developed by Iacus et al. (2012). Table 3 shows balance between foreign-trained and domestically-trained scientists of our matched covariates across both the full and star sample.

6. Results

6.1. Comparisons of matched pairs

Our knowledge flow measure is the number of times the US scientific community cites the focal scientist. Under the hypothetical scenario of full displacement of an equivalent domestically-trained scientist, we test for significant differences between the number of citations to the work of the foreign-trained scientists and their (hypothetically displaced) domestically-trained match. We look at all foreign-trained scientists and also, separately, the subset of foreign-trained stars.

For each sample, we make three distinct comparisons. First, we compare pre-move foreign-trained scientists with their domestically-trained matches. This allows us to understand spillover differentials to the US scientific community before the move takes place. Second, we compare spillovers of foreign-trained scientists before and after their move to the US. This allows us to understand the way the foreign-trained scientists' connection to US science changes upon moving to the US. Third, we compare post-move foreign-trained scientists with their domestically-trained matches. This is our main comparison, and it allows us to understand how localized knowledge flows would be affected even with full displacement of an equivalent domestically-trained scientist.

In Fig. 7 we provide a graphical depiction of all three comparisons and also compare across the full and star samples. The general picture that emerges is that pre-move foreign-trained scientists produce significantly fewer knowledge flows to the US scientific community than their domestically-trained matches. However, this gap tends to disappear with the move as foreign-trained scientists appear to quite rapidly integrate with the US scientific community. Post move, the number of citations from the US scientific community to the foreign-trained scientists' work is at least as large as to their domestically-trained matches.

Tables 4 and 5 provide formal tests for our three comparisons for the full and star-only samples, respectively. The top panel of each table compares pre-move foreign-trained scientists with their domestically-trained matches. We look separately at total citations, US citations, and the share of US citations of total citations. Indicating the success of the matching procedure, there is no statistically significant difference in total citations for the foreign-trained scientists and the domestically-trained matches in the pre-move period indicating that the foreign-trained and domestically-trained scientists are of equal productivity. However, the domestically-trained matches have significantly higher US citations and higher shares of US citations in total citations. This difference is particularly pronounced for the star sample, where on average domestically-trained matches receive more than 10 additional

[7] Results are qualitatively similar if we adopt a much looser definition of a domestic forward citation wherein a citing paper is classified as domestic if there is at least one US author in the author list.

[8] By making use of publication data to infer scientist location, we consequently also use publication data to infer mobility. That is, we can only identify the first year that a scientist immigrated to the US in the year that they first published a paper in the US.

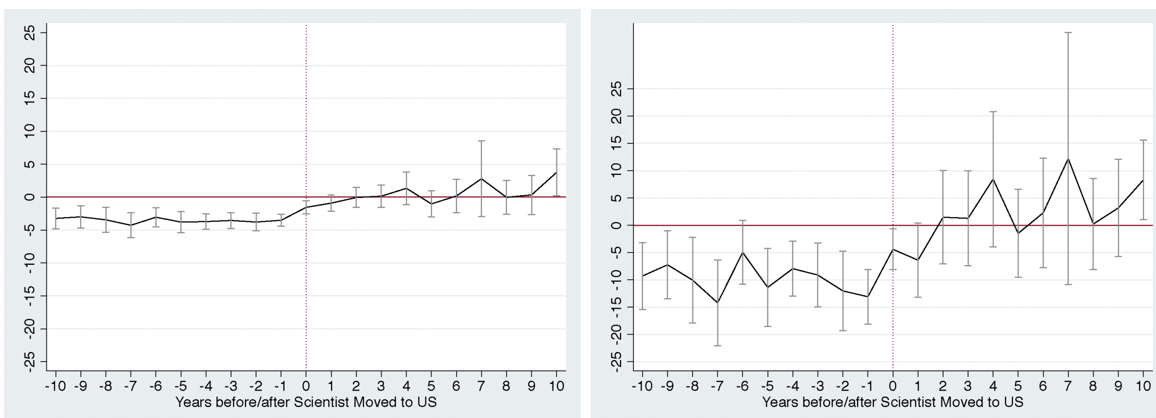
Table 3
Descriptive statistics of matched domestically-trained and foreign-trained scientists.

Variable	Domestically-trained scientists mean	Foreign-trained scientists mean	Difference	*p*-value of difference
Panel A				
Career age	7.43	7.48	−0.05	0.73
Ever a star	0.14	0.14	0.00	1.00
Σ^{t-1} Cites	154.11	153.92	0.19	0.97
Cites	34.36	35.19	−0.83	0.37
Observations	4,623	4,623		
Panel B: star sample				
Career age	9.99	10.1	−0.11	0.79
Ever a star	1.00	1.00	0.00	1.00
Σ^{t-1} Cites	449.9	442.67	7.23	0.78
Cites	78.7	80.83	−2.13	0.54
Observations	640	640		

6.2. *Factors mediating the integration of foreign-trained scientists into US science*

Recognizing that not all foreign-trained scientists will be equally well-positioned to generate knowledge flows for the US scientific community, we next explore how sensitive our main result is to plausible factors mediating the connection of foreign-trained scientists to US science networks. Where a factor is plausibly linked to a weaker (stronger) relationship to US scientists, a finding of a smaller (larger) "foreign-trained immigrant premium" gives us greater confidence that the difference between the matched pairs provides a meaningful measure of different spillover potential between foreign-trained immigrants and the domestically-trained US scientists they (hypothetically) displace.

We examine two candidate-mediating factors. The first is the prevalence of co-nationals at the destination institution. A higher prevalence of co-nationals is likely to be associated with more limited connections to US scientists (McPherson et al., 2001). Such differential



(a) Full Sample　　　　(b) Star Sample

Fig. 7. Number of US citations: foreign-trained relative to domestically-trained scientists. *Notes*: This figure plots point estimates for leading and lagging indicators for the migration of a scientist to the US. Both panels plot the point estimates of the following specification estimated using OLS: $\text{USCitations}_{it} = \sum_{\tau=0}^{10} \alpha_{-\tau}\text{Arrival}_{i,t-\tau} + \sum_{\tau=1}^{10} \alpha_{+\tau}\text{Arrival}_{i,t+\tau} + \sum_{\tau=0}^{10} \beta_{-\tau}\text{Arrival}_{i,t-\tau}\times \text{immigrant}_i + \sum_{\tau=1}^{10} \beta_{+\tau}\text{Arrival}_{i,t+\tau} \times \text{immigrant}_i + \theta(\text{Age}_{it}) + \delta_t + \varepsilon_{it}$. $USCitations_{it}$ is the number of citations received by scientist i in year t from US-authored papers. The α parameters (21 in all) controls for the US citation patterns of the matched domestic scientists for each year 10 years prior and post to the matched immigrants arrival. The β parameters are our point estimates of interest and are the ones plotted in the above figure. These reflect the differences in US citation patterns between immigrants and domestic scientists for each year around the move year (± 10 years). θ flexible controls for scientist i's age and δ is a full set of year dummies. There is no constant in this specification. The vertical bars correspond to 95% confidence intervals with scientist-clustered standard errors.

citations from the US scientific community compared to pre-move foreign-trained scientists.

The bottom panel of each table tests for differences in citations to the work of foreign-trained scientists pre- and post-move. Post-move foreign-trained scientists receive significantly more total citations and US citations and also receive a higher share of US citations of their total citations for both the full and star-only samples.

The middle panel of each table compares post-move foreign-trained scientists with their domestically-trained matches – our central comparison. Post-move foreign-trained scientists receive a larger number of citations from the US scientific community compared to their domestic matches (difference in full sample = 0.47, p-value = 0.15; difference in star sample = 2.58, p-value = 0.06).

Overall, using citations as our measure of knowledge spillovers, foreign-trained scientists are found to produce at least an equivalent level of knowledge flows to the US scientific community as the matched domestically-trained scientists who they hypothetically displace. At least by this measure, there is no evidence that the immigration of foreign-trained scientists would harm US science even with full displacement.

integration is supported by findings that co-ethnicity supports knowledge flows (see, e.g., Agrawal et al., 2008), so that the close proximity of co-nationals could reduce the incentive for the foreign-trained scientists to form connections with US scientists. The second is where the use of English is common in the foreign-trained scientist's country of training. Proficiency in English should be positively associated with the ability of the foreign-trained scientists to connect with US scientists. A large literature has documented that proficiency in English is positively associated with success in English-speaking destination-country labour markets (see, e.g., Chiswick and Miller, 1995; Dustmann and Fabbri, 2003).

We show the results of these difference-in-difference analyses in Tables 6 and 7 . We focus in particular on the difference in post-move US cites between foreign-trained scientists and their domestically-trained matches for both the full and star samples (Columns 2 and 5). In Table 6, we first compare the size of this "foreign-trained scientist premium" when the foreign-trained scientists have at most a single diaspora colleague with the case where they have two or more such colleagues. For the full sample, when foreign-trained scientists are relatively isolated, there is a statistically significant positive foreign-

Table 4
Mean comparisons of citations.

Variable	Foreign-trained Mean (1)	Std. Dev. (2)	Domestically-trained Mean (3)	Std. Dev (4)	Column Diff (5)	*P*-value of diff (6)
Pre-move period	$N = 28,449$		$N = 28,449$			
(1) Citations	17.68	54.25	17.92	49.64	−0.24	0.58
(2) US citations	6.55	24.58	10.13	28.30	−3.57	0.00
(3) Share of US citation	0.33	0.22	0.57	0.23	−0.24	0.00
Post-move period	$N = 21,008$		$N = 21,008$			
(4) Citations	20.45	74.40	18.08	52.19	2.36	0.00
(5) US citations	10.25	39.06	9.77	28.26	0.47	0.15
(6) Share of US citations	0.49	0.25	0.54	0.24	−0.06	0.00
	Row Diff	*p*-value of diff				
(7) Citations	2.77	0.00				
(8) US citations	3.70	0.00				
(9) Share of US citations	0.15	0.00				

Notes: Each observation is at the scientist-year level. Citations is the mean sum of the number of forward citations to papers published by the scientist in the specific time period (pre or post move). US Citations is the mean annual count of the number of forward citations to papers published by scientist *i* in the time period where the first author of the citing paper resides in the US. Foreign-trained and domestically-trained scientists are matched using coarsened exact matching along the following dimensions: scientist age, total citations within the US, and discipline.

Table 5
Mean comparisons of citations (star sample).

Variable	Foreign-trained Mean (1)	Std. Dev. (2)	Domestically-trained Mean (3)	Std. Dev (4)	Column Diff (5)	*p*-Value of diff (6)
Pre-move period	$N = 5,103$		$N = 5,103$			
(1) Citations	52.28	108.92	53.08	96.50	−0.79	0.70
(2) US citations	19.95	50.90	30.17	55.18	−10.21	0.00
(3) Share of US citations	0.35	0.18	0.58	0.18	−0.23	0.00
Post-move period	$N = 4,611$		$N = 4,611$			
(4) Citations	59.04	144.97	50.01	95.22	9.03	0.00
(5) US citations	29.79	76.53	27.22	51.49	2.58	0.06
(6) Share of US citations	0.49	0.20	0.55	0.20	−0.06	0.00
	Row Diff	*p*-value of diff				
(7) Citations	6.75	0.00				
(8) US citations	9.84	0.00				
(9) Share of US citations	0.15	0.00				

Notes: Each observation is at the scientist-year level. Citations is the mean sum of the number of forward citations to papers published by the scientist in the specific time period (pre or post move). US Citations is the mean annual count of the number of forward citations to papers published by scientist *i* in the time period where the first author of the citing paper resides in the US. Foreign-trained and domestically-trained scientists are matched using coarsened exact matching along the following dimensions: scientist age, total citations within the US, and discipline.

trained scientists premium; but there is a statistically significant negative premium (or discount) when the foreign-trained scientists is co-located with two or more diaspora colleagues. The null of no difference between these premiums is strongly rejected (p-value = 0.001). For the star sample, the size of the positive premium for the relatively isolated foreign-trained scientists is even more pronounced than in the full sample. However, the effect is not statistically significant when there are two or more diaspora colleagues. The null of no difference between these premiums is again strongly rejected (p-value = 0.001).

In Table 7 we repeat these comparisons of the "foreign-trained scientists premium" based on whether the foreign-trained scientists comes from a country where the use of English is common or not. For the full sample, the premium is only marginally statistically significant where the foreign-trained scientists comes from a country where English is common. However, there is no statistically significant effect for foreign-trained scientists from countries where English is uncommon. The p-value for the null of no difference between the two cases is 0.072. Interestingly, for the star sample, we cannot reject the null of no difference between the two cases (p-value = 0.251). This may reflect the fact that strong English ability is common among stars regardless of whether they come from a country where the use of English is common or not.

Overall, the results of these difference-in-difference analyses are generally consistent with our priors. Foreign-trained scientists tend to perform better in terms of connections to US science when they are relatively isolated from co-nationals and also come from countries where the use of English is common, although the latter effect is not evident for stars.

7. Concluding comments

The search for evidence of native wage and employment displacement effects has been a major theme of the immigration literature. More recently, in an attempt to better identify the benefits of high-skilled immigration, more attention has focused on knowledge spillovers to native workers. But this raises a new possibility of harm if local knowledge networks are disrupted by arrivals who displace domestic workers who are better embedded in knowledge-sharing networks. To explore the possibility of such displacement, we use citation patterns to answer a simple question: Are foreign-trained scientists less connected to the US scientific community than the domestically-trained scientists they potentially displace? We find that although foreign-trained scientists are significantly less connected to US science than their domestically-trained matches pre-immigration, convergence is rapid post-immigration. Overall, we do not find evidence of harm to domestic science through a knowledge network disruption channel.

We conclude by noting some possible limitations of our findings and important areas for further research. First, while we use state-of-the-art matching techniques to identify our domestically-trained matches for foreign-trained scientists, there is an inevitable residual concern that actual scientists displaced by foreign-trained arrivals are better connected to domestic scientists than these identified matches. It may also be that universities engaged in recruiting foreign-trained scientists are selecting those who are most likely to increase their productivity after arrival, increasing both the total knowledge spillovers they produce and also those that flow to the US.

Table 6
Difference-in-differences, diaspora effect.

Sample	Full			Star		
Dependent variable	(1) Cites	(2) US cites	(3) US cites share	(4) Cites	(5) US cites	(6) US cites share
Foreign-trained with ≤ 1 diaspora colleagues†	5.280*	2.010+	−0.031**	27.430**	12.056*	−0.043**
	(2.169)	(1.199)	(0.006)	(9.406)	(5.313)	(0.014)
Foreign-trained with ≥ 2 diaspora colleagues‡	−1.857	−1.752**	−0.046**	−0.132	−1.966	−0.062**
	(1.136)	(0.600)	(0.005)	(4.334)	(2.272)	(0.012)
University fixed effects	✓	✓	✓	✓	✓	✓
p-Value of H_0 : † = ‡	0.001	0.002	0.017	0.001	0.004	0.159
Observations	41,970	41,970	41,970	9,181	9,181	9,181

Notes: The unit of analysis is the scientist-year. The sample consists of domestically-trained and foreign-trained scientists in the US. All specifications are estimated using OLS. Robust standard errors clustered at the scientist level are in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.

Table 7
Difference-in-differences, countries where English is common.

Sample	Full			Star		
Dependent variable	(1) Cites	(2) US cites	(3) US cites share	(4) Cites	(5) US cites	(6) US cites share
Foreign-trained from country where English is common†	2.445+	0.531	−0.029**	4.820	0.564	−0.057**
	(1.415)	(0.730)	(0.006)	(4.730)	(2.462)	(0.014)
Foreign-trained from country where English is common‡	−1.008	−1.320	−0.052**	14.162+	5.293	−0.054**
	(1.647)	(0.912)	(0.006)	(8.189)	(4.632)	(0.013)
University fixed effects	✓	✓	✓	✓	✓	✓
p-Value of H_0 : † = ‡	0.078	0.081	0.001	0.251	0.296	0.838
Observations	41,970	41,970	41,970	9,181	9,181	9,181

Notes: The unit of analysis is the scientist-year. The sample consists of domestically-trained and foreign-trained scientists in the US. All specifications are estimated using OLS. Robust standard errors clustered at the scientist level are in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.

Second, while we believe that forward citations provide the best measure of knowledge connections between scientists, other possibilities exist. One alternative is co-authorships with US scientists. Preliminary results suggest that foreign-trained scientists have fewer post-arrival co-authorship relationships with US scientists than their domestically-trained matches. But conditional on a co-authorship relationship with a US scientist, the quality of the output as measured by forward citations to the work is higher for immigrant-domestic collaborations. The nature of this quantity/quality tradeoff and also the relative importance of citation and co-authorship metrics as measures of connections between scientists requires further exploration.

Third, the diaspora and English-language results point to the kind of variables that mediate the integration of foreign-trained scientists into US knowledge networks. More work is needed to better understand the integration process and the public or organizational policies that might facilitate it.

Finally, although scientists who publish are a key component of US knowledge networks, further work is required to confirm that foreign-trained scientist-related network disruption effects do not cause greater harm in other knowledge sectors. An advantage of examining scientific papers is that a natural paper trail of connections is provided through citation patterns. Patent citations may allow for an extension of the approach used here to explore network disruption effects in other parts of the US knowledge system.

We close by briefly considering the possible policy implications of our findings. Of course, the value to a receiving country of more open admission policies for scientists and other highly skilled workers depends on multiple factors. In addition to the potential implications for scientific and innovative productivity, policy makers will weigh, for example, distributional and fiscal effects. Our paper focuses on one possible source of harm from recruiting foreign-trained scientists – the displacement of better connected domestically-trained scientists and thus the disruption of domestic knowledge networks. Our findings suggest that this should not be a significant concern, as foreign-trained scientists appear to quickly embed into US knowledge networks as indicated by the flows of their US-knowledge creation to other US scientists relative to their US-trained peers. This occurs despite evidence of a lower connection to US scientists prior to their arrival. Even so, full integration can take time and some may never fully integrate. Selection polices that factor in predicted speed of integration and policies that help aid the integration process – e.g., avoiding restrictions on free circulation – will help limit harm through any displacement-related network-disruption effects.

References

Agrawal, Ajay, Kapur, Devesh, McHale, John, 2008. How do spatial and social proximity influence knowledge flows? Evidence from patent data. J. Urban Econ. 64 (2), 258–269.

Agrawal, A., Cockburn, I., McHale, J., 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. J. Econ. Geogr. 6 (5), 571–591.

Agrawal, A., McHale, J., Oettl, A., 2015. Collaboration, stars, and the changing organization of science: evidence from evolutionary biology. The Changing Frontier: Rethinking Science and Innovation Policy. The University of Chicago Press, Chicago University of Toronto mimeo.

Azoulay, P., Zivin, J.G., Wang, J., 2010. Superstar extinction. Q. J. Econ. 125 (2), 549–589.

Borjas, G., 2005. The labor-market impact of high-skill immigration. Am. Econ. Rev. 95 (2), 56–60.

Borjas, G., 2007. Do foreign students crowd out native students from graduate programs. In: Ehrenberg, R.G., Stephan, P.E. (Eds.), Science and the University. University of Wisconsin Press, Madison.

Borjas, G., Doran, K.B., 2012. The collapse of the Soviet Union and the productivity of American mathematicians. Q. J. Econ. 127 (3), 1143–1203.
Borjas, G.J., 2014. Immigration Economics. Harvard University Press.
Borjas, G.J., Doran, K.B., 2015. Cognitive mobility: labor market responses to supply shocks in the space of ideas. J. Labor Econ. 33 (S1), S109–S145.
Breschi, S., Lissoni, F., 2001. Knowledge spillovers and local innovation systems: a critical survey. Ind. Corporate Change 10 (4), 975–1005.
Chiswick, R., Miller, P.W., 1995. The endogeneity between language and earnings: international analyses. J. Labor Econ. 13 (2), 246–288.
Cooke, A., Kemeny, T., 2017. Cities, immigrant diversity, and complex problem solving. Res. Policy 46 (6), 1175–1185.
Dustmann, C., Fabbri, F., 2003. Language proficiency and labour market performance of immigrants in the UK. Econ. J. 113 (489), 695–717.
Feldman, M.P., 2000. Location and innovation: the new economic geography of innovation, spillovers, and agglomeration. In: Clark, G., Feldman, M., Gertler, M. (Eds.), The Oxford Handbook of Economic Geography. Oxford University Press, Oxford, pp. 373–394.
Freeman, R., Ganguli, I., Murciano-Goroff, R., 2015. Why and wherefore of increased scientific collaboration. The Changing Frontier: Rethinking Science and Innovation Policy. The University of Chicago Press, Chicago.
Ganguli, I., 2015. Immigration and ideas: what did Russian scientists "bring" to the United States? J. Labor Econ. 33 (S1), S257–S288.
Hunt, J., 2015. Are immigrants the most skilled US computer and engineering workers? J. Labor Econ. 33 (S1), S39–S77.
Hunt, J., Gauthier-Loiselle, M., 2010. How much does immigration boost innovation? Am. Econ. J. Macroecon. 2 (2), 31–56.
Iacus, S.M., King, G., Porro, G., 2012. Causal inference without balance checking: coarsened exact matching. Polit. Anal. 20 (1), 1–24.
Jaffe, A.B., Trajtenberg, M., 2002. Patents, Citations, and Innovations: A Window on the Knowledge Economy. MIT Press.
Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Q. J. Econ. 108 (3), 577–598.
Kahn, S., MacGarvie, M., 2016. Do return requirements increase international knowledge diffusion? Evidence from the Fulbright program. Res. Policy 45 (6), 1304–1322.
Kerr, S.P., Kerr, W.R., 2011. Economic impacts of immigration: a survey. Finnish Economic Papers, vol. 24, no. 1. Finnish Economic Association, pp. 1–32 1-32.
Kerr, W.R., 2008. Ethnic scientific communities and international technology diffusion. Rev. Econ. Stat. 90 (3), 518–537.
Kerr, William R., Lincoln, William F., 2010. The supply side of innovation: H-1B visa reforms and US ethnic invention. J. Labor Econ. 28 (3), 473–508.
Lan, F., Hale, K., Rivers, E., 2015. Immigrants' Growing Presence in the U.S. Science and Engineering Workforce: Education and Employment Characteristics in 2013, NSF 15-328. National Science Foundation.
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annu. Rev. Sociol. 27 (1), 415–444.
National Academies of Sciences, Engineering, and Medicine, 2016. The Economic and Fiscal Consequences of Immigration. The National Academies Press, Washington, DC.
Oettl, A., 2012. Reconceptualizing stars: scientist helpfulness and peer performance. Manage. Sci. 58 (6), 1122–1140.
Peri, Giovanni, 2012. The effect of immigration on productivity: evidence from US states. Rev. Econ. Stat. 94 (1), 348–358.
Peri, G., Shih, K., Sparber, C., 2013. STEM workers, H1B Visas and Productivity in US Cities. (Working Paper).
Stephan, Paula, 2010. The "I" s have it: immigration and innovation, the perspective from academe. Innov. Policy Econ. 10 (1), 83–127.
Stephan, P., 2012. How Economics Shapes Science. Harvard University Press, Cambridge.
Tang, L., Walsh, J.P., 2010. Bibliometric fingerprints: name disambiguation based on approximate structure equivalence of cognitive maps. Scientometrics 84 (3), 763–784.
Thompson, P., Fox-Kean, M., 2005. Patent citations and the geography of knowledge spillovers: a reassessment. Am. Econ. Rev. 95 (1), 450–460.
Waldinger, F., 2010. Quality matters: the expulsion of professors and the consequences for PhD student outcomes in Nazi Germany. J. Polit. Econ. 118 (4), 787–831.