Does a Decline in Star Immigration Help or Harm US Science?*

Ajay Agrawal
University of Toronto and NBER
ajay.agrawal@rotman.utoronto.ca

John McHale
National University of Ireland, Galway
john.mchale@nuigalway.ie

Alexander Oettl
Georgia Institute of Technology
alex.oettl@scheller.gatech.edu

October 24, 2013

Abstract

Evidence suggests that the propensity of foreign star scientists to immigrate to the US may be in decline. Would such a decline help or harm US science? We develop a model in which immigrant scientists enhance national welfare by increasing the labor supply. At the same time, they may harm domestic science if there are displacement effects on domestic scientists and if spillovers generated by immigrant scientists are significantly lower than those from their domestic counterparts. Using data from a variety of scientific disciplines, we examine evidence for differential spillovers from star scientists based on an examination of four distinct channels discussed in the literature: 1) citations, 2) collaborations, 3) colleague productivity, and 4) recruiting. Although we do find evidence suggesting that immigrant scientists generate more spillovers internationally, we do not find evidence that they generate less spillovers domestically. Thus, combining these findings with existing evidence of limited displacement effects at the faculty level, we conclude that a decline in star scientist immigration would likely harm US science.

JEL Classifications: J61, O31, O33

Preliminary - please do not cite or distribute

---

*This research was funded by the Centre for Innovation and Entrepreneurship at the Rotman School of Management, University of Toronto, the Martin Prosperity Institute, and the Social Sciences and Humanities Research Council of Canada. All errors are our own.

1 Introduction

Foreign-born scientists play a significant role in US science. Stephan (2012) notes that 48% of the PhDs awarded by US institutions in science and engineering are to foreign-born students on temporary or permanent visas. The share of foreign postdocs is even higher, with 60% held by temporary residents. The presence of foreign-born academics in US science faculties is smaller but still substantial. One study of the composition of chemistry faculties at research-intensive US universities found that almost 25% of faculty members received their undergraduate education outside the US (Stephan, 2012).

Furthermore, the foreign-born are disproportionately represented among elite scientists. In a study that employs a variety of "star" metrics (elected to the National Academy of Sciences or National Academy of Engineering, authors of citation classics, authors of hot papers, the 250 most-cited authors, authors of highly cited patents, scientists who played a key role in launching a biotechnology firm), Levin and Stephan (1999) report that "individuals making exceptional contributions to science and engineering in the United States are disproportionately drawn from the foreign born." This is notable because the distribution of output across scientists is highly concentrated, with a relatively small fraction of researchers producing a substantial fraction of the citation-weighted total output. For example, in the field of evolutionary biology, 80% of total quality-weighted output is produced by the top 20% of scientists. Moreover, evidence indicates that the disproportionate output by "stars" is increasing over time (Agrawal, McHale, and Oettl, 2013a).[1]

While the risk of domestic stars being displaced completely from US science by the

---

[1] The increasing concentration of output produced by stars has occurred even as the base of institutions participating in research has broadened. A possible explanation is that stars gain disproportionately from improvements in collaboration technologies (Agrawal, McHale, and Oettl, 2013a). Improvements in these technologies (e.g., email, electronic file sharing, low-cost teleconferencing) may be especially beneficial to those who develop a rich network of relationships through prior co-location, for example through supervisor-student relationships. To the extent that stars have a larger set of relationships to choose from in matching with collaborators, improvements in collaboration technologies may particularly benefit stars, even as their collaborators become more dispersed.

inflow of immigrant stars is likely to be low, they could be displaced from faculty positions at leading research universities that play a central role in the development of collaborative networks. If, in turn, domestic stars are better able to take advantage of the opportunities for "network centrality" that positions at leading universities give them, then their displacement by immigrant stars could affect the productivity of US academic science.

This paper examines the evidence for differential spillovers between domestic and foreign star scientists working at US universities. We first develop a simple model of the market for scientists that allows for the possibility of displacement of domestic scientists by immigrant scientists and also the possibility of differential spillovers. In the model, the combination of displacement and differential spillovers could harm US science. However, a sufficient condition for the absence of harm is the absence of differential spillovers. After first examining trends in the propensity of foreign stars to immigrate to the US across a range of disciplines, we then look for evidence for differential spillovers for immigrant- and domestic-star scientists. We first compare the relative collaboration and citation patterns of domestic and foreign stars. Turning to the institutional level, we look for evidence of differential productivity and reputational spillovers following domestic and foreign star arrivals on both incumbent productivity at the university and the quality of subsequent faculty recruits.

Our paper relates to a number of literatures: the effects of immigration on domestic wage and employment displacement, the productivity impacts of immigration, and star centrality in peer relationships. Kerr and Kerr (2011) provide a recent survey of native wage and employment displacement effects. Overall, they find displacement effects are relatively small. Using variation in the inflow of immigrants across US states, Peri (2012) finds no evidence of displacement of natives in employment. Looking at the market for new doctorates, Borjas (2005) does find evidence of significant wage displacement effects, concluding that "[a]n immigration-induced 10-percent increase in the supply of doctorates in a particular field at a particular time reduces the earnings of that cohort of doctoral recipients by 3 percent"

(Borjas, 2005, p. 59). Using the natural experiment of the large influx of Soviet mathematicians to the US following the fall of communism, Borjas and Doran (2012) find evidence of US mathematicians being crowded out of fields where the ex-Soviet mathematicians were concentrated, with increased movement to lower-ranked institutions and out of active publishing. Using a shift-share analysis, Stephan and Levin (2007) examine the impact on US citizens in terms of positions on US faculties. Summing up this work, Stephan (2012, p. 196) finds that "one can conclude that some displacement has occurred in academe, but it is fairly minimal and concentrated primarily in postdoc positions."

The literature has tended to find positive aggregate productivity effects as a result of highly skilled immigrant inflows, although the evidence is more limited compared with the work on displacement. 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 1 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 1 percentage point 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 utilise H-1B visa holders increase 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 a 1% increase in the high-skilled population in a state due to immigration, increases income per worker by 1% in that state. Peri, Shih, and Sparber (2013) report that an H-1B visa driven increase in STEM (Science, Technology, Engineering, Mathematics) workers increases the wages of both STEM and non-STEM workers at the city level. In related research using Canadian data, Peri and

Shih (2013) find a positive effect of foreign STEM workers with a college bias in the resulting productivity growth. Borjas and Doran (2012) find that although the total output of US mathematicians shrank following the influx of ex-Soviet mathematicians, the new immigrants filled the gap, although they find "no evidence that the Soviets greatly increased the size of the 'mathematics pie' " (Borjas and Doran, 2012, p. 1143).

Our paper also relates to the literature on peer effects in science and in particular the centrality of stars to peer relationships. Waldinger (2012) finds that the dismissal of scientists in Nazi Germany had long-lasting impacts on German science, with the effects driven by the quality of subsequent hiring. Azoulay et al. (2010), using the exogenous source of variation of stars that died prematurely, find that collaborators experience an average 5 - 8% decline in their quality-adjusted publication rates. Rather than the effects being due to the ability of the star to connect co-authors to funding sources, editorial goodwill, etc., they find that the losses are driven by the absence of an "irreplaceable source of ideas," with losses being larger for those closer to the star in ideas space. Oettl (2012) also uses unexpected star deaths to examine the impact of stars on peer productivity. Utilizing publications and citations to measure productivity and paper acknowledgements to measure "helpfulness," he finds that co-authors of "helpful" scientists experience a decline in paper quality (though not quantity), but that the death of "unhelpful" scientists does not affect co-author output. Using publication and citation data from the field of evolutionary biology, Agrawal, McHale, and Oettl (2013b) find that the arrival of a star benefits incumbents working on topics related to the star but harms those working on unrelated topics. However, they uncover positive effects on the quality of subsequent recruits working in areas both related and unrelated to the star.

We structure the remainder of the paper as follows. In the next section, we develop a simple model of the market for scientists that provides a general framework for an examination of the welfare implications of scientist immigration. The model allows for both domestic scientist displacement and differential spillovers from domestic and immigrant scientists. We describe the data that we use to test for differential spillovers in Section 3. We outline the methods for domestic- and immigrant-star identification as well as our scientist productivity and spillover metrics. We examine the evolution of the propensity of foreign stars to immigrate to the US in Section 4. Then, in Section 5, we compare the co-authorship and citation patterns of domestic and immigrant stars to help determine if they differ significantly in terms of their US-based networks. In Section 6, we use institution-level data to compare the productivity and reputational spillovers of the two types of stars. We conclude in Section 7 with a summary 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 of immigration. The model allows for the displacement – or "crowding out" – of domestic scientists as a result of the immigration of scientists. We adopt the ex ante social welfare perspective of the receiving country, and thus ignore the welfare gains to immigrant scientists. The model also allows for possible differential spillovers from domestic and immigrant scientists. We show it is possible for domestic social welfare to be harmed by immigration as a result of displacement if the difference between domestic and immigrant spillovers is large enough, even if immigration expands the overall size of the active scientific workforce. However, a sufficient condition for immigration to improve domestic social welfare is that there is no difference in the size of per-scientist spillovers between domestic and immigrant scientists.

2.1 Basic market setup

We begin with a specification of the 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 domestic and immigrant labor units.[2] The supply of domestic scientists, $L^s_{domestic}$, is assumed to be a positive linear function of the wage, w:

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

Immigrant labor units, I, are assumed to be supplied perfectly inelastically (possibly due to visa-related limitations), so the total supply of labor is given by[3]:

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

Total labor demand, L^d, is assumed to be 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 assumed to be common across domestic and immigrant scientists. The marginal social value relationship is then given by:

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

---

[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] 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 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 Figure 1. The equilibrium wage and employment levels are given by:

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

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

Total social surplus from trade in the scientific labor market is given by 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:

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

The total social surplus is given by the sum of areas A, B, and C in Figure 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 domestic and immigrant scientists

We next allow for positive immigration but initially assume that spillovers, z, are identical for domestic and immigrant scientists. We graph this case in Figure 2. The new equilibrium wage and employment levels are given by:

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

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

It is also useful to identify the employment level of domestic 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 will be no displacement if ϕ_1 is equal to zero, so that the domestic labor supply is perfectly inelastic. To determine total social surplus, it is useful to separate out the surplus due to domestic versus immigrant scientists. Using Equation (10), the part due to domestic scientists is given by:

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

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

Because we are taking the perspective of the welfare of the receiving country, we exclude the surplus accruing directly to immigrant scientists. Domestic social surplus accruing from immigrants is thus the difference between the marginal social value curve and the post-immigration wage line (Equation (8)), where it is assumed that immigrants are the marginal labor suppliers. This surplus is given by:

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

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

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

Given 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 Figure 2.

2.4 Social surplus with immigration but with differential spillovers for domestic and immigrant scientists

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

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

Compared to the case of equal spillovers, an examination of Figure 3 shows a loss of social surplus on units that would have been supplied by domestic scientists in the absence of displacement. The lower spillovers from immigrant scientists also reduce 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 this loss and gain. If the gap between z^D and z^I is large enough, it is possible that the displacement of domestic scientists reduces social surplus overall, notwithstanding the larger total size of the scientific workforce.

We now can identify from Equation (15) two distinct sufficient conditions for immigration not to reduce domestic social surplus given any level of immigration (i.e., for $S^{**}_{total} \geq S^{*}$). 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 domestic and immigrant spillover, i.e., $z^D - z^I = 0$.

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

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

The "break-even" level of immigrant spillover is then the level of z^I at which Equation (16) holds with equality. We graph the break-even in Figure 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 equal to $2\phi_1 z^D$. Given that the size of the immigrant spillover is assumed to be bounded from below at zero (i.e., we assume the spillover is not negative), any immigration level above this level is associated with a net benefit regardless of the level of domestic displacement.

Summing up this section, we have found in the context of a simple market model with spillovers that it is possible that immigration 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 domestic scientists by immigrants and lower spillovers from immigrants compared with domestic counterparts. However, the size of the spillover required from immigrant scientists to avoid immigration harming social welfare is decreasing in the level of immigration. Notwithstanding displacement effects, a sufficient condition for scientist immigration not to harm 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 domestic 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 domestic 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[4] – and that domestic 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

---

[4] 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, McHale, and Oettl (2013a), 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.

weighed against the direct gains from scientist immigration. The search for evidence on possible differential spillovers from domestic and immigrant stars motivates the empirical work in the remainder of the paper.

3 Data

We use bibliometric data published in the ISI Web of Science (WoS) to construct scientist event histories, identify instances of scientist movement, identify scientific stars, calculate scientist coauthoring relationships, and construct article-level citation patterns.

3.1 Field definitions

We conduct our analysis on six scientific disciplines: Economics, Evolutionary Biology, Immunology, Mathematics, Neuroscience, and Psychology. We identify all papers and subsequently scientists in each respective field by drawing on the field classifications developed by the ISI Journal Citation Reports.

3.2 Identifying scientists

We first disambiguate scientific authors since the WoS data do not provide scientist unique identifiers. This is an important step because it is critical to distinguish between two authors who share the same first initial and surname (the primary method in which authors are identified within the WoS).

To address this issue, we employ heuristics 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.

We identify stars as scientists in the 90th percentile in a given year in terms of their accumulated stock of citation-weighted paper output over the preceding years. We provide a more detailed explanation of how we identify stars and related features of the data in our companion paper that focuses on stars and that uses the same data (Agrawal, McHale, and Oettl, 2013b).

3.3 Identifying scientist locations

Using the unique author identifiers generated in the process described above for each evolutionary biology paper, we then attribute each scientist to a particular institution for every year they are active. A scientist is 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 attribute the location of all authors on these papers to the focal institution.

3.4 Defining immigration

With information on each scientist's location in each year, we identify the country of each scientist's institution. A domestic scientist is a scientist who started their career in the U.S. and never emigrated. An immigrant scientist is a scientist who started their career in a country other than the U.S. and some year after their first year immigrated to the U.S.

3.5 Outcome measures

For each scientist-year, we first identify all papers published by the focal scientist in the focal year. From this set of papers, we identify all unique coauthors of these papers. We then classify each coauthor as a domestic coauthor (residing in the U.S. in the focal year) and an international coauthor (not residing in the U.S. in the focal year).

Next, we return to the list of all papers published by the focal scientist in the focal year. For each paper, we count the number of forward citations (counts of citations made to the focal paper by other papers in the future). We then classify each forward citation as domestic if the first author of the future paper that references the focal paper is from the US.

4 The Changing Pattern of Star Immigration into the US

Before examining the evidence for differential spillovers between domestic and immigrant scientists, we first look at the changing propensity of foreign star scientists to move to the US. Our approach is to use least squares regressions, where the dependent variable is a binary variable that takes the value of 1 if a scientist who previously resided outside the US chooses to move to the US in a given year. The explanatory variables are dummy variables indicating whether the scientist is a star and a linear time trend. We also interact the star dummy and the time trend to identify the star-specific trend in the move probability.

We report the results in Table 1 for all scientists in our sample and also broken down by discipline. Not surprisingly, the results show that stars have a higher move probability, with the probability relatively similar across disciplines (lowest in evolutionary biology and highest in economics). We find evidence of a statistically significant negative trend in the move probability of stars, which is again similar across disciplines.

An obvious concern is that the linear trend might mask a more uneven pattern of move probabilities, possibly driven by period-specific features of US immigration law or the changing relative attractiveness of US institutions as a destination for stars. In Figure 5 we show the results when we use five-year time interval dummies instead of a linear trend. The excluded dummy is for non-stars in the first time period. The results indicate that the move probability for stars was relatively high between the mid-1970s and mid-1980s but then fell steeply. By the mid-2000s, the move probability of stars was almost identical to non-stars and actually below the move probability of non-stars at the beginning of our sample.

While Figure 5 presents immigration probabilities, it may be that the number of at-risk scientists has decreased over time, masking an increase in the absolute number of stars immigrating to the U.S. In Figure 6 we present absolute immigration rates across time revealing a slight increase in star immigration up until the early 1990s, after which the number of stars has continued to decrease.

Is the falling move probability a cause for concern? If displacement effects are large and the spillovers from immigrants are significantly lower than for domestic scientists, then the model of Section 2 suggests the fall in move probability could actually be welfare improving. But if spillovers are not significantly different, then the model suggests welfare will be lower when there is reduced capacity to recruit immigrant stars. We turn next to examining the evidence for differential spillovers.

5 Are There Differential Spillovers? Evidence from Differential Connections

One reason for why spillovers from immigrant stars could be lower than their domestic counterparts is that they are less connected to other scientists in the US. In this section, we look for evidence of differential connectedness to US-resident scientists using co-author and forward citation metrics. We test the null of no difference in mean connectedness. In doing so, we conduct three different sets of mean comparisons: 1) a comparison of mean connectedness measures for eventual immigrants before and after their move, 2) a comparison of domestic and immigrant mean connectedness measures in the set of all star scientists, and 3) a comparison of mean connectedness measures for sets of matched domestic-immigrant star pairs.

5.1 Mean comparisons for immigrant stars pre- and post-move

We begin with an examination of immigrant connection before and after their move to the US. We are particularly interested in how the moves affect the pattern of their connections. No evidence of change in connection patterns would suggest a failure to integrate and immediately raise the likelihood that the spillovers from immigrant stars will be lower than for domestic stars.

We report the results in Table 2. While we find no evidence that the total number of co-authors changes as a result of the move, the mix of co-authors shifts dramatically to US-based co-authors. Prior to the move, these eventual immigrant stars had a mean number of roughly 0.7 US-based co-authors; after the move, this mean number rises to roughly 3.2. For total forward citations, we cannot reject the null of no change in the mean, although the point estimate does rise from roughly 44 to 59. We also cannot reject the null of no significant difference in the mean number of US-resident forward citations, with the number rising from roughly 16 (pre-move) to roughly 29 (post-move). Overall, the evidence strongly suggests an adaptation in connection patterns consistent with integration into the US-based scientific community.

5.2 Mean comparisons for immigrant- and domestic-star samples

We report the mean comparisons for immigrant and domestic stars in Table 3. In terms of total citation-weighted output stocks, we find a statistically significant mean difference in favor of immigrant stars, indicating that immigrant stars are more productive on average. We also find a statistically significant difference in the mean number of total co-author connections and citations in favor of immigrant stars. Turning to US-resident co-authors and US-resident forward-citing scientists – our main focus – we find no statistically significant mean difference between immigrant and domestic stars in the case of co-authors and a slightly larger number of US-resident forward citations accruing to immigrant stars. In terms of these raw connection indicators, the evidence of differential spillovers is mixed.

We also compare the immigrant and domestic mean *shares* of US-resident co-authors and forward citations. These results do show a statistically significant mean difference for both co-authors and forward citations in favor of domestic stars, although the point estimates are reasonably close. The share might matter to the extent that spillovers from a given co-authorship or forward citation are smaller when the star's connections are more thinly spread. In other words, there is a degree of rivalry in spillovers, with more connections meaning less access to spillovers through any given connection. This is likely to matter more for co-authorship connections than for forward-citation connections. There is likely a limited amount of attention that a star can distribute over co-authors. Thus, more co-authors could mean less attention and potentially less spillovers. This "rivalry" is likely to be less important for forward citations as they reflect connections to ideas, which as Paul Romer has emphasised, are inherently non-rivalrous (Romer, 1990). On the other hand, to the extent that social relationships mediate idea flows, a degree of rivalry in access to the star's ideas could still be present.

In Figures 7 and 8 we disaggregate by scientific discipline the mean comparisons that we report in Table 3.

5.3 Mean comparisons for matched star samples

The above mean comparisons focus on differential connection metrics for all stars, where we define stardom as having a citation-weighted output stock in the top decile of the star's discipline. As indicated by the simple stock comparisons, immigrant stars have, on average, higher citation-weighted output than their domestic counterparts. In other words, immigrant stars are more concentrated than domestic stars in the upper reaches of the top decile. A stricter test for differential connections is to match each immigrant star with a domestic counterpart based on citation-adjusted stock and discipline.

We report the results of a comparison of means across the matched samples in Table 4. Interestingly, immigrant stars appear to have more connections overall than their domestic counterparts as measured by mean co-authors and mean forward citations. However, we find evidence of statistically significant means between immigrant stars and their domestic matches in relation to domestic connections for both the co-author (smaller) or forward-citation (larger) metrics.

Turning to the share measures, we again find a statistically significant difference in favor of domestic stars. This raises the possibility that the quality of immigrant connections to domestic scientists may be less conducive to spillovers, especially in relation to co-author mediated connections.

Figures 9 and 10 disaggregate the mean comparisons found in Table 4 by scientific discipline.

6 Evidence on Differential Local-Productivity Effects

6.1 Defining Evolutionary Biology

Defining knowledge in evolutionary biology is not straightforward. On the input side, evolutionary biology, as in many areas of science, draws from many fields, such as statistics, molecular biology, chemistry, genetics, and population ecology. Furthermore, on the output side, some of the most influential papers are published in general interest as opposed to field-specific journals. Therefore, identifying the set of papers that comprise the corpus of the field is complicated because although every paper in the *Journal of Evolutionary Biology* is probably relevant, most papers in *Science* and *Nature* are not, although a significant fraction of the field's most important papers are published in those latter two journals.

Therefore, we follow a three-step process for defining "evolutionary biology papers." First, using bibliometric data from the WoS, we collect data on all articles published during the 29-year period 1980 through 2008 in the journals associated with the four main societies that focus on the study of evolutionary biology: the Society for the Study of Evolution, the Society for Systematic Biology, the Society for Molecular Biology and Evolution, and the European Society of Evolutionary Biology. Their respective journals are: *Evolution*, *Systematic Biology*, *Molecular Biology and Evolution*, and *Journal of Evolutionary Biology*. We focus on these four society journals because every article published within them is relevant to evolutionary biologists. In other words, unlike general interest journals such as *Science*, *Nature*, and *Cell*, which include papers from evolutionary biology but also research from many other fields, these four journals focus specifically on our field of interest. This process yields 15,256 articles.

Second, we collect all articles that are referenced at least once by these 15,526 society journal articles. There are 149,497 unique articles that are referenced at least once by the set of 15,256 evolutionary biology society articles. This set of 149,497 articles includes, for example, papers that are important to the field but are published outside the four society journals, such as key evolutionary biology papers published in *Science* that are cited, likely multiple times, by articles in the four society journals. We call this set of 149,497 papers the corpus of influence because each of these articles has had impact on at least one "pure" evolutionary biology article.

Third, we citation-weight the corpus of influence. We do this by counting the references to each of the 149,497 articles from the original 15,256 society journal articles. There are 501,952 references from the 15,256 society journal articles. So, on average, articles in the corpus are cited 3.4 times. Unsurprisingly, the distribution of citations is highly skewed. The minimum number of citations is one (by construction), the median is one, and the maximum is 906.[5] For most of the analyses in this paper, we use counts of citation-weighted publications. When we do so, we use the 149,497 articles weighted by the 501,952 society article references.

6.2 Results

The absence of significantly different co-authorship and citation patterns between immigrant and domestic stars provides suggestive evidence for the absence of differential spillovers. However, simply counting connections may miss factors affecting the quality of those connections in generating spillovers. The concern here is heightened by the lower share of domestic co-authors for immigrant stars. It is useful, then, to also look for more direct evidence of differential local productivity effects associated with co-location with a star.

Building on our prior work in Agrawal, McHale, and Oettl (2013b), in this section we examine the productivity impacts of immigrant and domestic star arrivals at the institution level for one discipline – evolutionary biology. In addition to examining total effects on

---

[5] This paper is "The neighbor-joining method - a new method for reconstructing phylogenetic trees," published in *Molecular Biology and Evolution* (1987) by Saitou Naruya (University of Tokyo) and Masatoshi Nei (University of Texas).

incumbent productivity, we separately examine the effects on incumbents whose work is related and unrelated to the arriving star. We also investigate the effects of star arrivals on the quality of subsequent department recruits (or "joiners"), again both related and unrelated to the arriving star. The latter can be taken as an indicator of the reputation of the arriving stars, including likely spillover effects but also their reputation in the broader community of scientists.

We report the results in Table 5 for incumbents and Table 6 for joiners. In general, we do not find evidence of statistically significant local impacts between immigrant and domestic stars, although there is a statistically significant difference at the 10% level for the quality of unrelated joiners in favor of domestic stars. Just looking at the size of the point estimates, however, there is evidence of smaller incumbent productivity benefits, the difference being more pronounced for unrelated incumbents. For joiner quality, the point estimate indicates a slight advantage for immigrant stars in relation to related joiners but, as already noted, a disadvantage for unrelated joiners.

7 Conclusions

This paper sets out a model in which it is possible for scientist immigration to harm domestic science if the domestic spillovers are lower for immigrant scientists and significant displacement of domestic scientists takes place. However, a sufficient condition for the absence of harm is the absence of differential spillovers. Focusing on star scientists, the empirical part of the paper then looks at the evidence for such a difference. The first approach is to look for indirect evidence through an examination of connection patterns, both co-authorship and forward citation. We find that immigrant stars dramatically increase their number of US co-authors upon moving to the US. Once in the US, immigrant and domestic stars show no significant difference in US-based co-authorship connections. Nor do we see evidence of differences in the number of US-based forward-citation connections. However, we do observe evidence that immigrants have a lower share of US-based co-authors in their set of total co-authors. This might suggest lower quality connections to the extent that co-authors are rivals for spillovers from the star.

Another interesting finding is that immigrant stars have higher citation-adjusted output stocks than domestic stars, where we define stars as those scientists in the top decile of the citations-adjusted output distribution. This raises the possibility that immigrant and domestic stars are different in the way they interact with other scientists. We thus also examine mean connections for matched samples of immigrant and domestic stars based on citation-adjusted output stocks and discipline. We again do not observe significant differences in the mean number of domestic co-authors or forward citations, but on average the immigrant stars again display a lower share of US-based co-authors than their domestic matches.

With these potential connection-quality differences in mind, we then look for more direct evidence of differential local productivity effects in the discipline of evolutionary biology. While the pattern of point estimates does suggest that domestic star arrivals provide a larger productivity benefit to incumbents, we generally cannot reject the null of no difference, with the exception of the quality of unrelated department joiners.

Given the amorphous nature of knowledge spillovers, it is not possible to be overly definitive about the absence of differential effects. However, the evidence reported in the paper on both indirect connections and more direct local-productivity effects does not indicate that differential spillovers are present. Combined with the limited existing evidence that displacement takes place at the faculty level and recalling from the model that immigrants could directly benefit US science even with differential spillovers and displacement, it seems likely that a decline in the propensity of foreign stars to immigrate to the US would harm US science.

References

Agrawal, A. K., J. McHale, and A. Oettl (2013a, October). Collaboration, stars, and the changing organization of science: Evidence from evolutionary biology. University of Toronto mimeo.

Agrawal, A. K., J. McHale, and A. Oettl (2013b, July). Why stars matter. University of Toronto mimeo.

Azoulay, P., J. Graff Zivin, and J. Wang (2010). Superstar extinction. *Quarterly Journal of Economics 125*(2), 549–589.

Borjas, G. J. (2005). The labor-market impact of high-skill immigration. *The American Economic Review 95*(2), 56–60.

Borjas, G. J. and K. B. Doran (2012). The collapse of the soviet union and the productivity of american mathematicians*. *The Quarterly Journal of Economics 127*(3), 1143–1203.

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

Kerr, S. P. and W. R. Kerr (2011, Spring). Economic impacts of immigration: A survey. *Finnish Economic Papers, Finnish Economic Association 1-32 24*(1), 1–32.

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

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

Peri, G. (2012). The effect of immigration on productivity: Evidence from us states. *Review of Economics and Statistics 94*(1), 348–358.

Peri, G., K. Shih, and C. Sparber (2013). Stem workers, h1b visas and productivity in us cities. Working Paper.

Peri, G. and K. Y. Shih (2013, April). Foreign scientists and engineers and economic growth in canadian labor markets. IZA Discussion Papers 7367, Institute for the Study of Labor (IZA).

Romer, P. M. (1990, October). Endogenous technological change. *Journal of Political Economy 98*(5), S71–S102.

Stephan, P. (2012). *How Economics Shapes Science*. Cambridge: Harvard University Press.

Stephan, P. and S. Levin (2007). Foreign scholars in u.s. science: Contributions and costs. In P. E. Stephan and R. G. Ehrenberg (Eds.), *Science and the University*. Madison: University of Wisconsin Press.

Tang, L. and J. P. Walsh (2010). Bibliometric fingerprints: Name disambiguation based on approximate structure equivalence of cognitive maps. *Scientometrics 84*(3), 763–784.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in nazi germany. *The Review of Economic Studies 79*(2), 838–861.

Figure 1: Market Equilibrium and Total Social Surplus, No Immigration

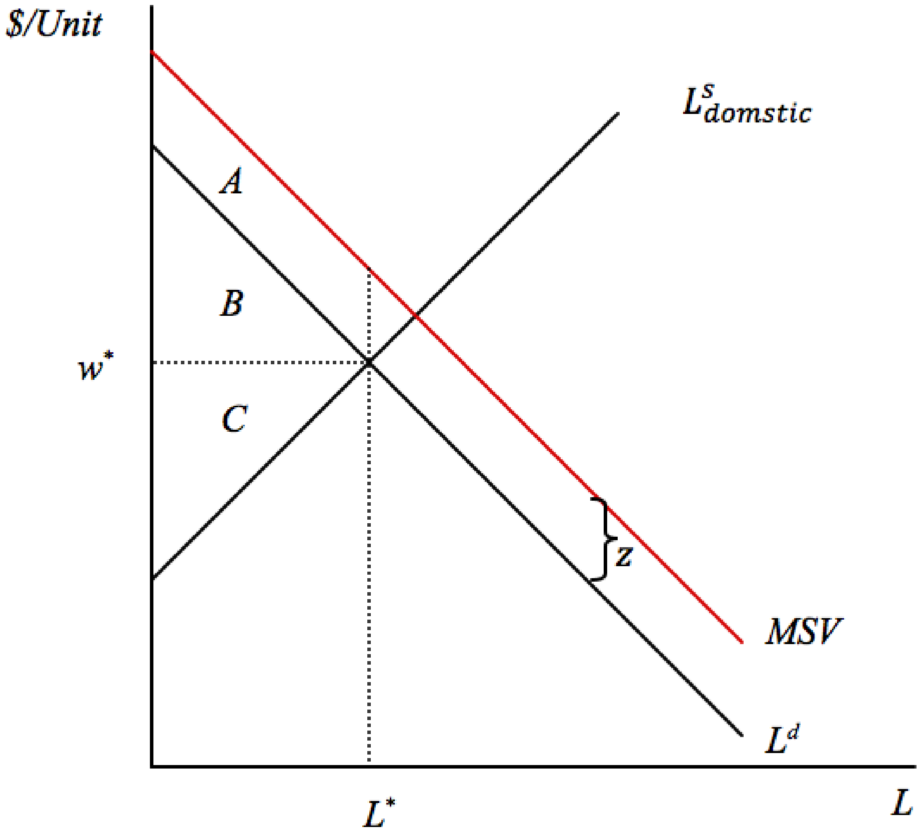
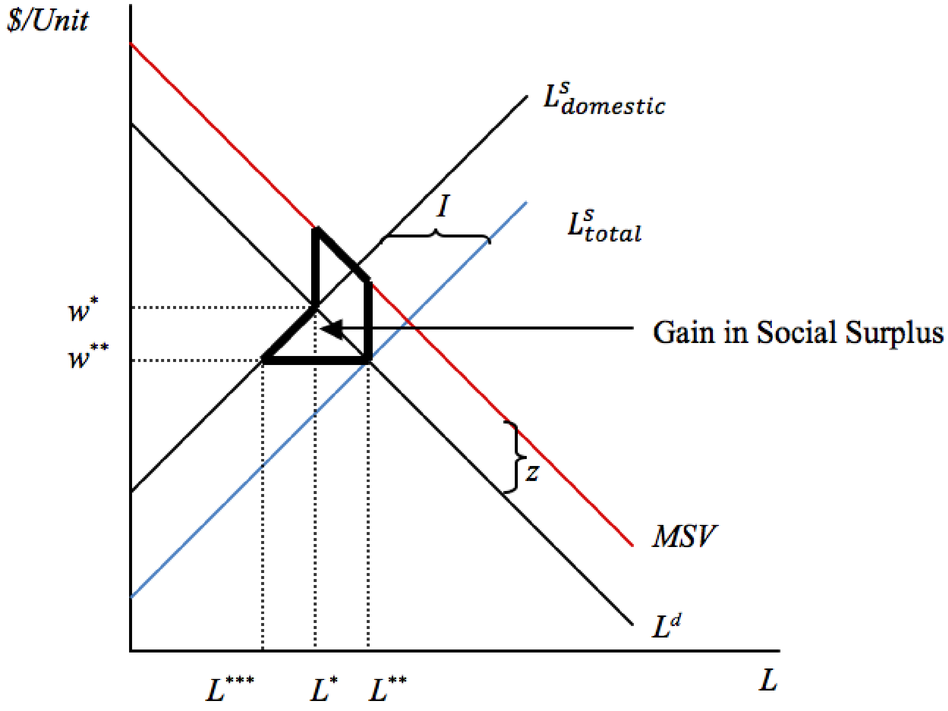


Figure 2: Market Equilibrium and the Gain in Social Surplus from Immigration



Note: The per-scientist externality is assumed to be equal to z for domestic and immigrant scientists.

Figure 3: Market Equilibrium and the Gain and Loss of Social Surplus when the per Scientist Externality is Lower for Immigrant Scientists

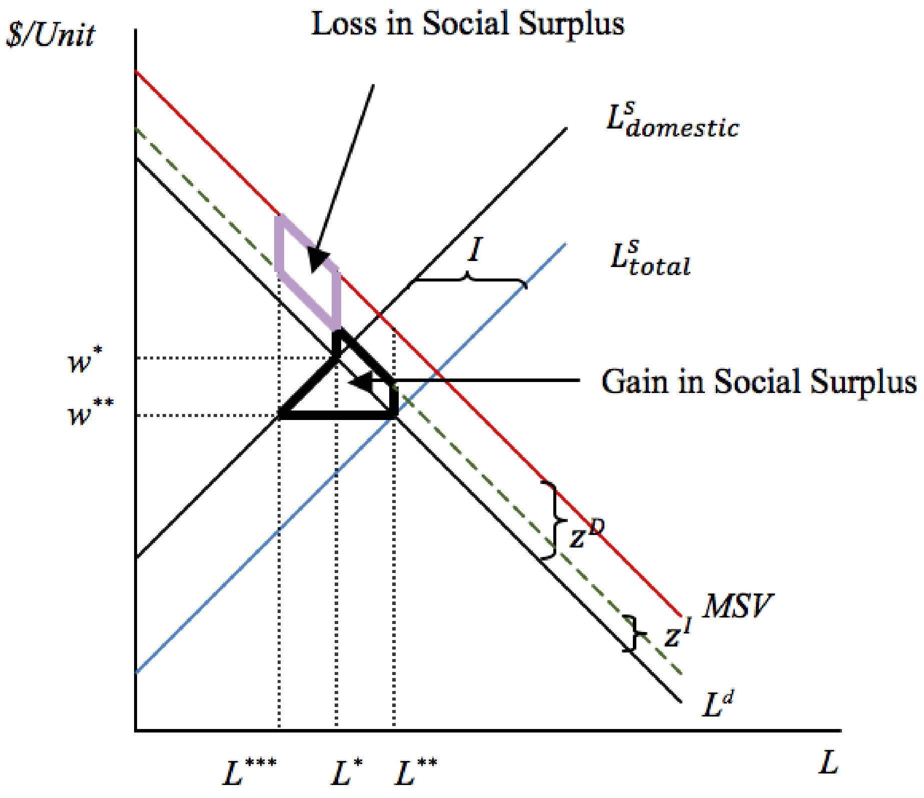


Figure 4: The Level of the Per-Scientist Externality for Immigrant Scientists for No Change in Social Surplus to Occur as a Result of Immigration

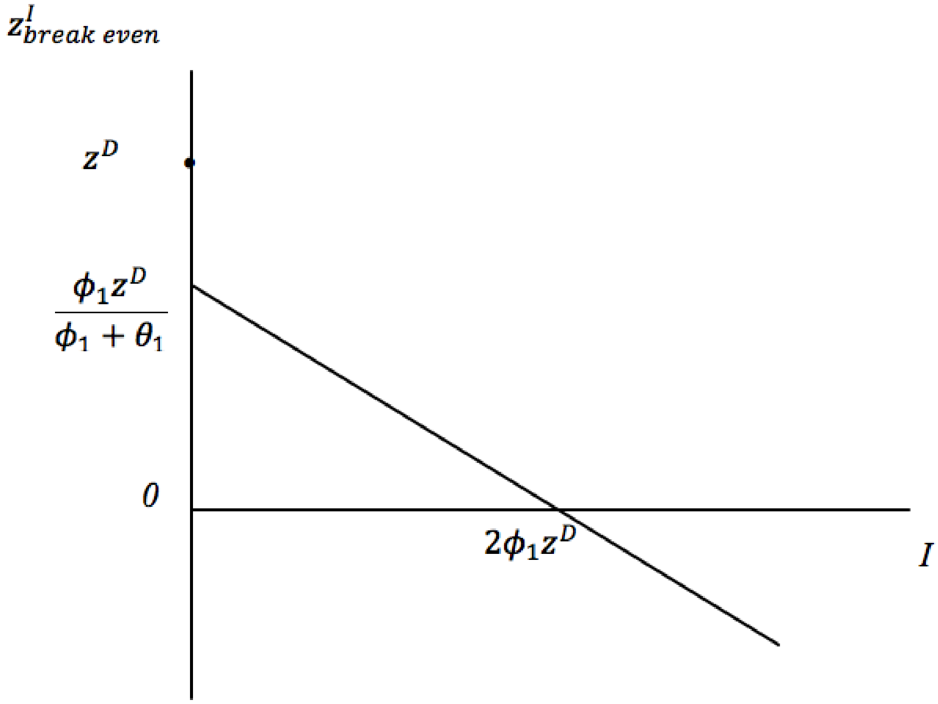


Figure 5: US Immigration Probability by Year across Five-Year Intervals

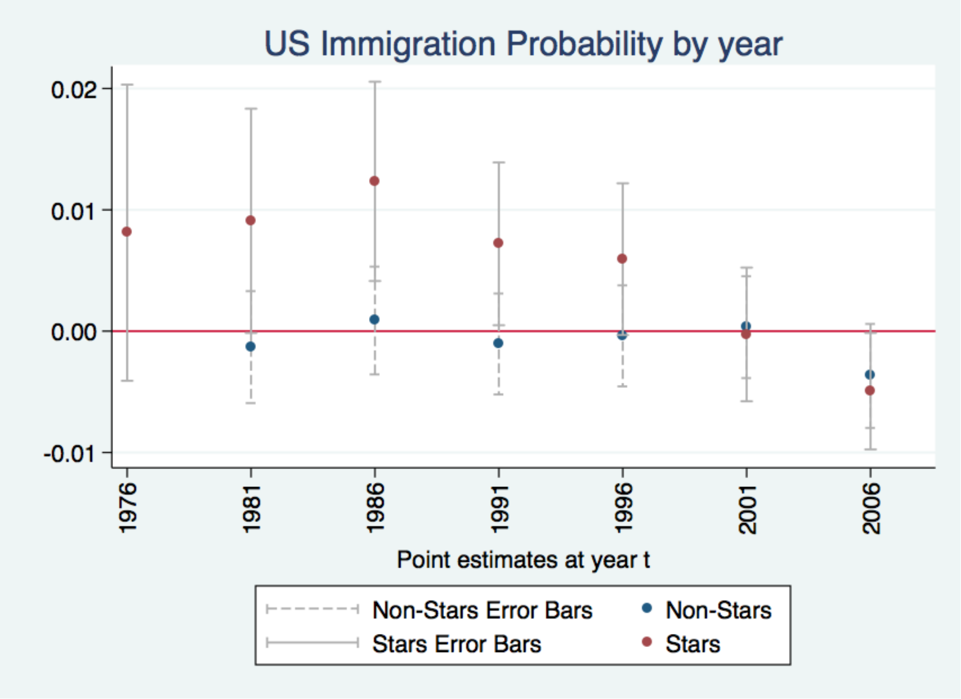


Figure 6: US Immigration by Year

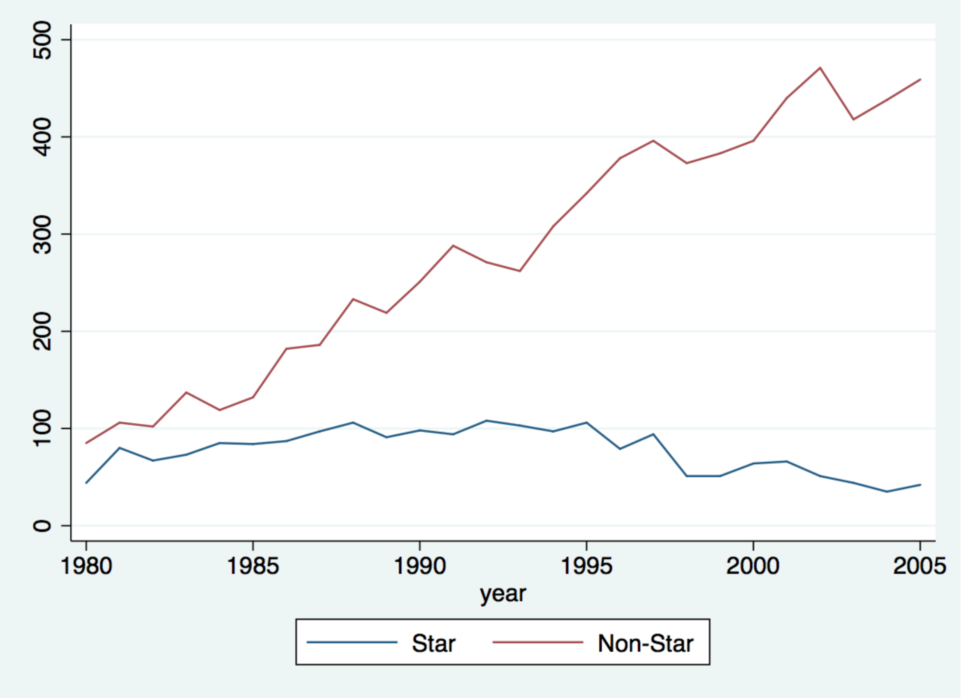
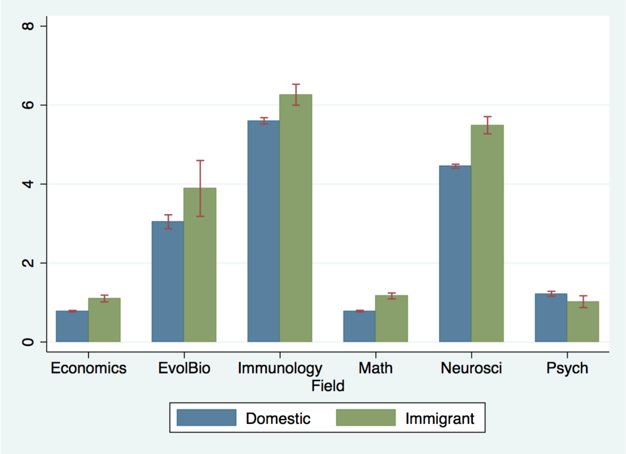
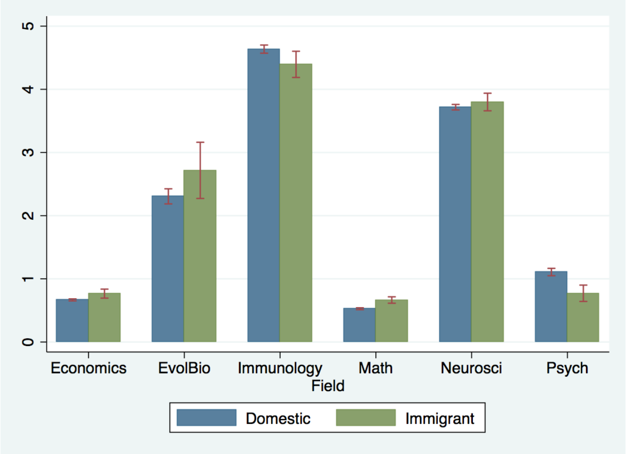


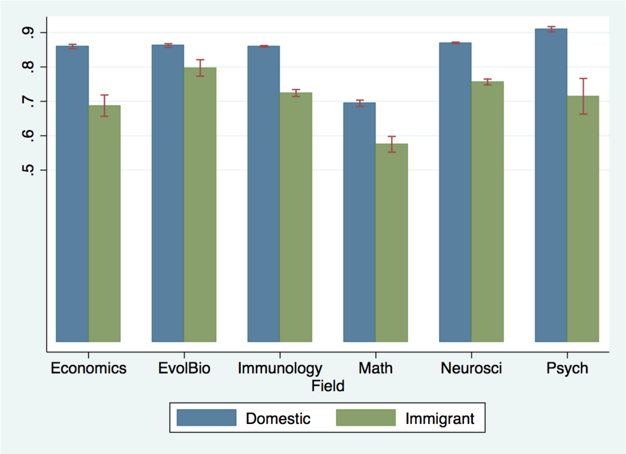
Figure 7: Discipline-Specific Coauthorships



(a) Mean Coauthorships

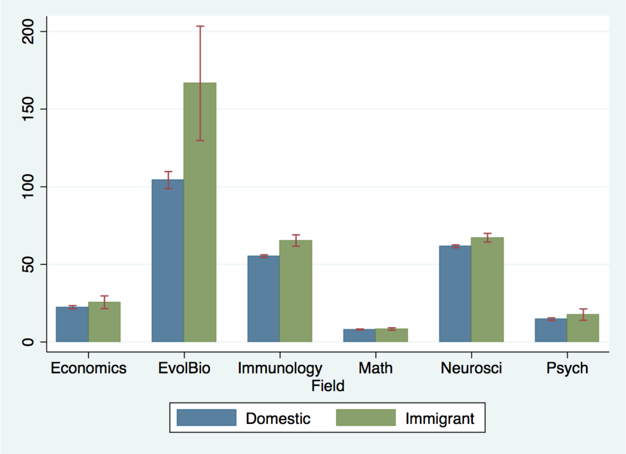


(b) Mean U.S. Coauthorships

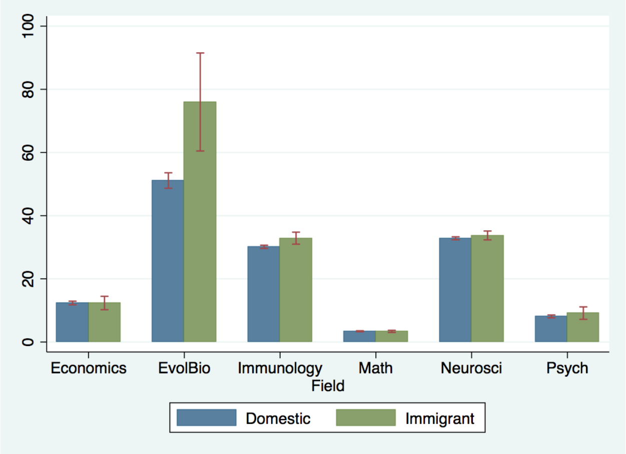


(c) Mean U.S. Coauthorship Share

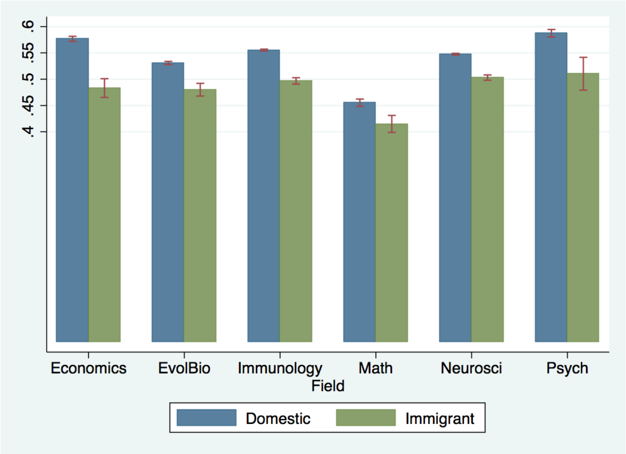
Figure 8: Discipline-Specific Citation Rates



(a) Mean Citations

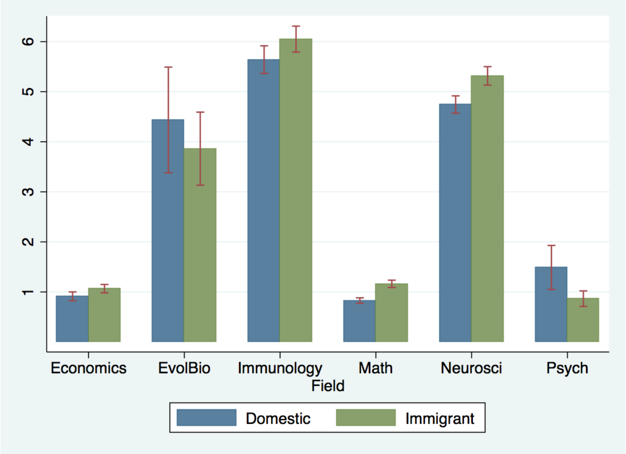


(b) Mean U.S. Citations

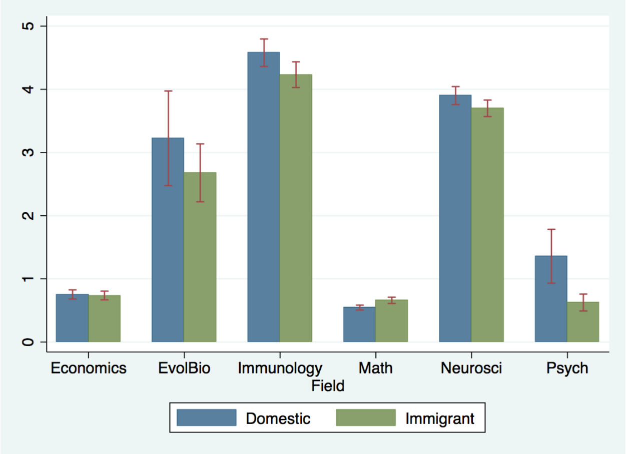


(c) Mean U.S. Citation Share

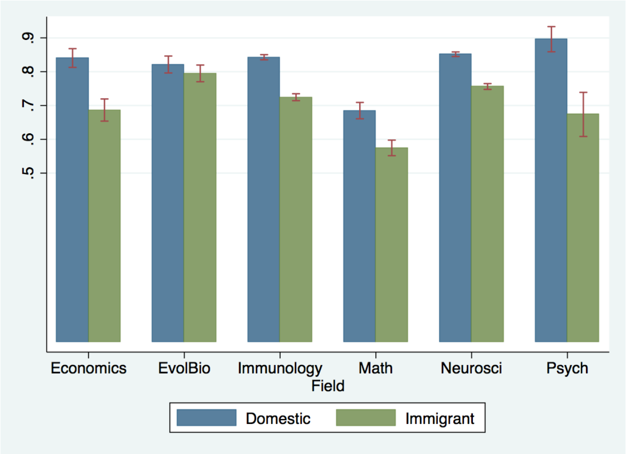
Figure 9: Discipline-Specific Coauthorships for Matched Sample



(a) Mean Coauthorships

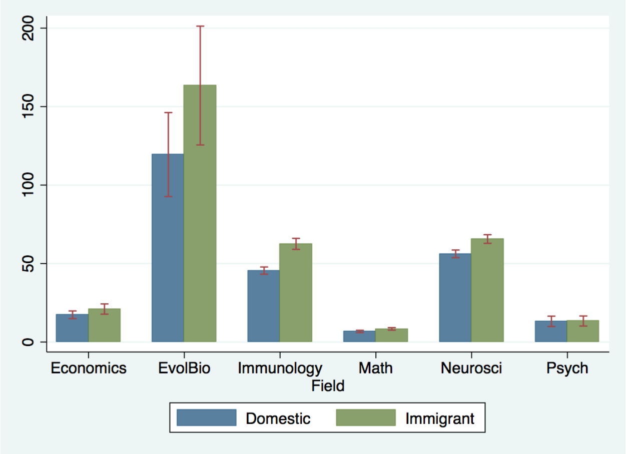


(b) Mean U.S. Coauthorships

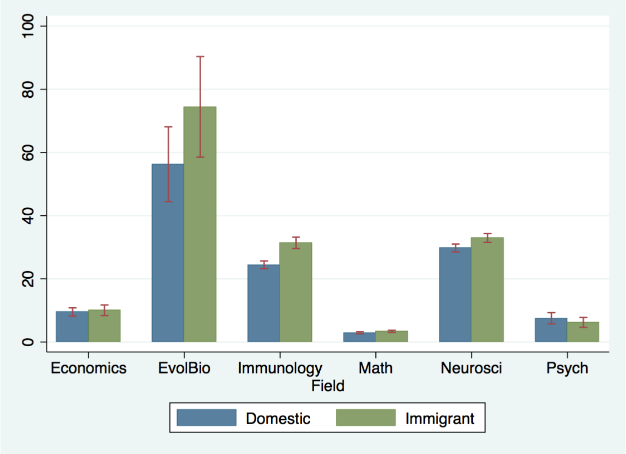


(c) Mean U.S. Coauthorship Share

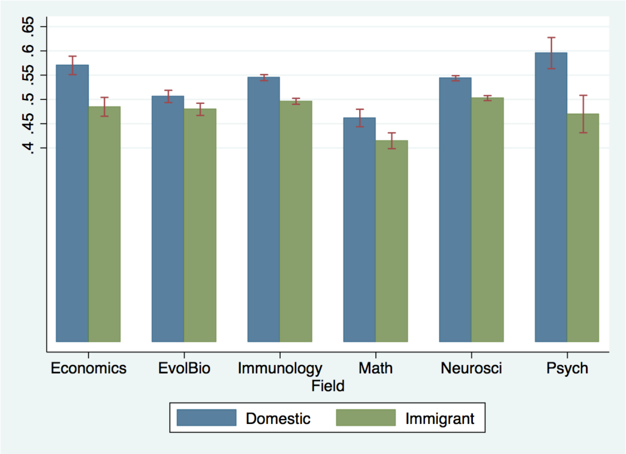
Figure 10: Discipline-Specific Citation Rates for Matched Sample



(a) Mean Citations



(b) Mean U.S. Citations



(c) Mean U.S. Citation Share

Table 1: Propensity of Scientists Who Reside in Rest of World to Immigrate to US

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Field:	All	EvolBio	Math	Economics	Neurosci	Immunology	Psychology
Star	0.017**	0.015**	0.023**	0.028**	0.016**	0.017**	0.016**
	(0.001)	(0.003)	(0.002)	(0.004)	(0.001)	(0.001)	(0.004)
Year	0.000**	−0.000	0.000	0.000	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Star X Year	−0.001**	−0.001**	−0.001**	−0.001**	−0.000**	−0.001**	−0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age FE	✓	✓	✓	✓	✓	✓	✓
Observations	1241775	73174	156982	73826	455023	437149	41780

Notes: The dependent variable is set to 1 if scientist i in year j moves to the United States, and 0 otherwise. Robust standard errors clustered at the scientist are in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 2: Mean Comparisons of Immigrant Stars Pre- and Post-move

	In US (Post-Move) N = 19,786		Not in US (Pre-Move) N = 23,705			
Variable	Mean	Std. Dev.	Mean	Std.Dev	Diff	p-value of diff
Knowledge Stock	1019.43	1283.38	731.79	999.33	287.64	0
Coauthors	4.58	9.22	4.54	9.31	0.04	0.67
US Coauthors	3.17	6.52	0.71	2.86	2.46	0
Share of US Coauthors	0.72	0.34	0.16	0.27	0.57	0
Citations	58.98	186.11	44.23	126.78	14.75	0
US Citations	29.08	85.56	16.27	56	12.81	0
Share of US Citations	0.49	0.2	0.32	0.19	0.17	0

Notes: Each observation is at the scientist-year level. Knowledge Stock is the stock of scientist i's citation-weighted publications up to year $t-1$. Coauthors is the number of unique coauthors in year t. US Coauthors is the number of these coauthors that reside in the US in year t. Citations is the count of the number of forward citations to papers published by scientist i in year t. US Citations is the count of the number of forward citations to papers published by scientist i in year t where the first author of the citing paper resides in the US.

Table 3: Mean Comparisons for Immigrant- and Domestic-Star Samples

	Immigrant		Domestic			
	N = 19,786		N = 215,813			p-value
Variable	Mean	Std. Dev.	Mean	Std.Dev	Diff	of diff
Knowledge Stock	1019.43	1283.38	967.53	1458.52	51.9	0
Coauthors	4.58	9.22	3.82	8.57	0.76	0
US Coauthors	3.17	6.52	3.15	6.8	0.02	0.65
Share of US Coauthors	0.72	0.34	0.86	0.26	-0.13	0
Citations	58.98	186.11	52.01	157.66	6.98	0
US Citations	29.08	85.56	27.52	76.65	1.55	0.01
Share of US Citations	0.49	0.2	0.55	0.2	-0.06	0

Notes: Each observation is at the scientist-year level. Knowledge Stock is the stock of scientist i's citation-weighted publications up to year $t-1$. Coauthors is the number of unique coauthors in year t. US Coauthors is the number of these coauthors that reside in the US in year t. Citations is the count of the number of forward citations to papers published by scientist i in year t. US Citations is the count of the number of forward citations to papers published by scientist i in year t where the first author of the citing paper resides in the US.

Table 4: Mean Comparisons for Immigrant- and Domestic-Star Matched Samples

	Immigrant		Domestic			
	N = 19,216		N = 19,216			p-value
Variable	Mean	Std. Dev.	Mean	Std.Dev	Diff	of diff
Knowledge Stock	903.24	926.27	902.96	926.27	0.28	0.98
Coauthors	4.47	8.51	4.10	9.05	0.37	0
US Coauthors	3.09	6.27	3.32	7.12	-0.22	0
Share of US Coauthors	0.72	0.34	0.83	0.28	-0.11	0
Citations	57.16	184.26	45.2	135.15	11.95	0
US Citations	28.11	84.33	23.61	64.51	4.49	0
Share of US Citations	0.49	0.2	0.54	0.2	-0.05	0

Notes: Each observation is at the scientist-year level. Knowledge Stock is the stock of scientist i's citation-weighted publications up to year $t-1$. Coauthors is the number of unique coauthors in year t. US Coauthors is the number of these coauthors that reside in the US in year t. Citations is the count of the number of forward citations to papers published by scientist i in year t. US Citations is the count of the number of forward citations to papers published by scientist i in year t where the first author of the citing paper resides in the US.

Table 5: Role of Immigrant Stars on Incumbent Output of US Departments

	(1)	(2)	(3)
Dependent Variable:		Incumbent Output	
Sample	Full	Related	Unrelated
Star	0.048	0.605*	−0.045
	(0.124)	(0.294)	(0.129)
Star X Immigrant	−0.105	−0.298	−0.260
	(0.152)	(0.385)	(0.196)
Incumbents	0.061**	0.061**	0.067**
	(0.008)	(0.012)	(0.009)
Year & Dept FE	✓	✓	✓
Observations	2408	1792	2408

Notes: Robust standard errors clustered at the department-level are in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 6: Role of Immigrant Stars on Joiner Quality Output of US Departments

	(1)	(2)	(3)
Dependent Variable:		Incumbent Output	
Sample	Full	Related	Unrelated
Star	0.672**	1.796**	0.505**
	(0.184)	(0.539)	(0.156)
Star X Immigrant	−0.186	0.089	−0.369*
	(0.204)	(0.613)	(0.179)
Scientists	−0.000	0.002	−0.000
	(0.004)	(0.010)	(0.005)
Year & Dept FE	✓	✓	✓
Observations	1722	1173	1722

Notes: Robust standard errors clustered at the department-level are in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$