STEM Workers, H-1B Visas, and Productivity in US Cities

Giovanni Peri, *University of California, Davis*

Kevin Shih, *University of California, Davis*

Chad Sparber, *Colgate University*

Science, technology, engineering, and mathematics (STEM) workers are fundamental inputs for innovation, the main driver of productivity growth. We identify the long-run effect of STEM employment growth on outcomes for native workers across 219 US cities from 1990 to 2010. We use the 1980 distribution of foreign-born STEM workers and variation in the H-1B visa program to identify supply-driven STEM increases across cities. Increases in STEM workers are associated with significant wage gains for college-educated natives. Gains for non-college-educated natives are smaller but still significant. Our results imply that foreign STEM increased total factor productivity growth in US cities.

I. Introduction

Science, technology, engineering, and mathematics (STEM) workers are the primary contributors to the creation and adoption of technological in-

We thank Nick Bloom, Hilary Hoynes, William Kerr, Enrico Moretti, Sarah Turner, and seminar participants at the University of California, Davis and Berkeley, Université Catholique del Louvain, Institute for the Study of Labor (IZA), and the National Bureau of Economic Research for helpful comments and suggestions. Peri gratefully acknowledges an Interdisciplinary Frontiers in the Humanities and Arts grant from the University of California, Davis, for partially funding this project. Sparber gratefully acknowledges a Major grant from Colgate University for funding his research. Contact the corresponding author, Giovanni Peri, at

[*Journal of Labor Economics,* 2015, vol. 33, no. 3, pt. 2]
 0734-306X/2015/33S1-0011$10.00
Submitted January 31, 2013; Accepted May 7, 2014; Electronically published June 29, 2015

novation, the fundamental driver of sustained economic growth. The importance of STEM innovations has long been recognized by growth economists. Griliches (1992) and Jones (1995), for example, have used measures of scientists and engineers to identify research and development (R&D) contributions to idea production, with the latter study arguing that scientists and engineers are responsible for 50% of long-run US productivity growth. A related literature (e.g., Katz and Murphy 1992; Acemoglu 2002; Autor, Katz, and Kearney 2006) has noted that technological innovation during the past 30 years has not increased the productivity of all workers equally. The development of new technologies—especially information and communication technologies (ICT)—significantly increased the productivity and wages of college-educated workers. They had a much smaller effect on the demand for non-college-educated workers, which has remained rather stagnant.

Importantly, while technological and scientific knowledge is footloose and spreads across regions and countries, STEM workers are less mobile. Tacit knowledge and face-to-face interactions influence the speed with which new ideas are locally adopted. Several studies (e.g., Moretti 2004a, 2004b; Iranzo and Peri 2009) have illustrated that concentrations of college-educated workers spur local productivity. Others have shown the tendency for innovation- and idea-intensive industries to agglomerate (Ellison and Glaeser 1999; Glaeser 2011; Moretti 2012) and for ideas to remain local generators of virtuous innovation cycles (Jaffe, Trajtenberg, and Henderson 1993; Saxenian 2002).

This article sits at the intersection of these literatures. We quantify the long-run effect of increased city-level STEM employment on labor market outcomes for STEM, college-educated, and non-college-educated native-born workers. Sections II and III describe our empirical specification and data. The challenge of the exercise is to identify variation in the growth of STEM workers across US metropolitan statistical areas (MSAs, or cities) that is supply driven and hence exogenous to other factors that affect local wages, employment, and productivity. We do this by exploiting the introduction of the H-1B visa in 1990 and the differential effect that these visas had in bringing foreign-born college-educated workers (mostly STEM workers) to 219 US cities from 1990 to 2010. The H-1B policy changes were national in scope but had differentiated local effects because foreign STEM workers were unevenly distributed across US cities before the inception of the H-1B visa program. Migrant preferences and the availability of information spread by ethnic networks led subsequent inflows of H-1B workers to concentrate in areas with a large preexisting foreign STEM presence.

---

gperi@ucdavis.edu. Information concerning access to the data used in this article is available as supplementary material online.

Our identification strategy is rooted in methods used by Altonji and Card (1991), Card (2001), and Kerr and Lincoln (2010). First, we measure foreign STEM workers as a share of employment in each MSA in 1980. This share exhibits large variation. Next, we predict the number of new foreign STEM workers in each city by allocating the H-1B visas to 14 foreign nationality groups in proportion to their city-level presence in 1980. This H-1B-driven imputation of future foreign STEM is a good predictor of the actual increase of both foreign STEM and overall STEM workers in a city over subsequent decades. Thus, we use this prediction as an instrument for the actual growth of foreign STEM workers in order to obtain causal estimates of the impact of STEM growth on the wages and employment of college-educated and non-college-educated native-born workers.

The 1980 distribution of foreign STEM and the overall inflow of H-1B workers between 1990 and 2010 could be correlated with unobservable city-specific shocks that affect employment and wage growth, so Section IV explores the power and validity of our instrumental variable strategy. We check that the initial industrial structure of the metropolitan area, the 1980 distribution of other types of foreign-born workers (e.g., less educated and manual workers), and the subsequent inflow of non-STEM immigrants do not predict foreign STEM employment growth. We also show that the trends of native outcomes prior to the inception of the H-1B program (1970–80) were uncorrelated with the H-1B-driven growth in STEM workers from 1990 to 2010. Finally, our demanding regression specifications always include both city and period fixed effects while relying on changes in growth rates of H-1B-driven STEM workers within MSAs over time for identification.

The main regression estimates are in Section V. Our preferred specifications reveal that a rise in foreign STEM growth by 1 percentage point of total employment increases wage growth of college-educated natives by 7–8 percentage points. The same change had a smaller but usually significant effect on non-college-educated native wage growth equal to 3–4 percentage points. We find no statistically significant effects for native employment growth.

Section VI closes the analysis by introducing a simple model of city-level production and combining it with our estimated parameters to simulate the effect of STEM on total factor productivity and skill-biased productivity. When we aggregate at the national level, inflows of foreign STEM workers explain between 30% and 50% of the aggregate productivity growth that took place in the United States between 1990 and 2010. This range is consistent with Jones's (2002) analysis of science and engineering contributions to productivity growth. We also find that foreign STEM inflows account for a more modest 4%–8% of US skill-biased technological change.

II. Empirical Framework

Our empirical analysis uses variation in foreign-born STEM workers across US cities (c) and time periods (t) to estimate their impact on native wages and employment. We discuss identification and its challenges in Section IV. The basic specifications we estimate in Section V take the form

$$y_{ct}^{\text{Native},X} = \phi_t + \phi_c + b_{y,X} \cdot \frac{\Delta \text{STEM}_{ct}^{\text{Foreign}}}{E_{ct}} + b_3 \cdot \text{Controls}_{ct}^{X} + \varepsilon_{ct}. \quad (1)$$

The variable $y_{ct}^{\text{Native},X}$ is the period change in outcome y (either employment or average weekly wages) for the subgroup of natives with skill X (either STEM workers, college-educated workers, or non-college-educated workers), standardized by the initial year outcome level. The term ϕ_t captures period fixed effects, while ϕ_c captures city fixed effects. The variable $\Delta\text{STEM}_{ct}^{\text{Foreign}}/E_{ct}$ is the change of foreign STEM over a period, standardized by a city's initial total employment (E_{ct}). The term Controls_{ct}^{X} includes other city-specific controls, and ε_{ct} is a zero mean idiosyncratic random error. The specification implies that identification relies on variation in the growth of foreign STEM workers within cities over time periods.

Our analysis spans 1990–2010, and we choose to partition these two decades into three specific time periods: 1990–2000, 2000–2005, and 2005–10. This enables us to exploit the large variation in national H-1B policy that occurred between 2000 and 2005 relative to the other periods. Additionally, this facilitates (unreported) robustness checks that remove the 2005–10 period to avoid influence from the Great Recession.[1]

The coefficient $b_{y,X}$ captures the elasticity of outcome y, for worker group X, to an exogenous increase in STEM workers. Interpreting these estimates as causal requires changes in $\text{STEM}_{ct}^{\text{Foreign}}$ that are exogenous to productivity shocks and other unobservable determinants of city-level wage and employment changes. Before turning attention to this challenge, we describe our data, STEM employment measures, the construction of the H-1B-driven foreign STEM instrument, and our instrument's power.

III. Data: STEM Workers in US Cities

We develop two separate methods of defining STEM occupations. Each method also uses both a more inclusive and a more restrictive STEM identification criterion, resulting in four possible STEM definitions. The first method is based on skills that workers use in their occupations. We use the US Department of Labor's (2012) O*NET database, which measures the occupation-specific importance of several dozen skills required

[1] Estimates are robust to removing the Great Recession. Similarly, they remain robust when constructing variables over 1990–2000 and 2000–2010. Results are available on request.

to perform the job. We select four O*NET skills that involve STEM use, namely, mathematics in problem solving, science in problem solving, technology design, and programming. We then compute the average score of each occupation across the four skills and rank the 331 occupations consistently identified in the 1980–2010 census according to the average STEM skill value defined above.[2] We classify STEM occupations as those employing the top 4% (strict definition) or 8% (broad definition) of workers in that ranking in the year 2010; O*NET 4% (or 8%) STEM workers are the individuals with these occupations.

Our second method for identifying STEM occupations is based on the skills workers possess before employment—the college majors found among workers within occupations. The US State Department recognizes a list of STEM majors for the purpose of granting foreign students extended time to work under the Optional Practical Training (OPT) program.[3] We rank occupations on the basis of the 2010 ACS share of individuals with a college degree in a STEM major. We then classify STEM occupations as those employing the top 4% (strict) or 8% (broad) of workers following that ranking in 2010. Major-based 4% (or 8%) STEM workers are the individuals within those occupations. Both the O*NET and major-based strict definitions include mainly census occupations with "scientist" or "engineer" in the title. Major-based STEM occupations largely coincide with O*NET STEM occupations.

A. H-1B Visa Policy Changes

Our analysis exploits large shifts in national H-1B visa policy between 1990 and 2010 as an exogenous source of variation in the inflow of foreign STEM workers across US cities to identify the effect of STEM workers on the wages and employment of native-born workers. The H-1B visa, introduced in 1990, provides temporary permits for college-educated foreign "specialty" workers. The visa has been a crucial channel of admission for many college-educated foreign-born workers employed in STEM occupations.[4] Set initially at 65,000 H-1B visas annually, the cap rose to 115,000 for

[2] We make small refinements to the census occupational classification in order to ensure complete time consistency in the availability of occupations over the 1980–2010 period. A detailed description of both of our STEM definitions, as well as the refinement of occupations, is available in the online appendix.

[3] There is no direct crosswalk between majors listed under the OPT STEM classification and major categories in the 2010 American Community Survey (ACS). Thus, our list is consistent with, but not identical to, OPT STEM degree fields.

[4] Lowell (2000) notes that 70% of H-1B visas have been awarded to people employed as computer analysts, programmers, electrical engineers, university professors, accountants, other engineers, and architects. Similarly, US Citizenship and Immigration Services (various years) reports that for all years between 2004 and 2011, more than 85% of new H-1B visa holders worked in computer science, health science, accounting, architecture, engineering, and mathematics.

fiscal years 1999 and 2000 and then to 195,000 per year for 2001, 2002, and 2003. It reverted to the original 65,000 beginning in 2004. Though the limit officially remains at 65,000, the first 20,000 H-1B visas issued to individuals who have obtained a graduate degree in the United States became exempt from H-1B limits beginning in 2005, effectively raising the cap to 85,000.[5]

Not only has the size of the H-1B program varied greatly since its inception, but the ensuing inflow of foreign STEM workers has been heterogeneously distributed across US cities as well. Part of these cross-city differences was certainly due to varying economic conditions, industrial structures, and labor demand influencing wage and employment growth. Importantly, however, a portion of this variation was due to persistent immigrant preferences to locate in cities with historical communities of past immigration. The 1980 distribution of STEM workers by nationality proxies for these historical settlements. Our analysis needs to capture only the heterogeneity in foreign STEM created by this differential initial presence (in 1980) of foreign enclaves by nationality that are exogenous to other determinants of future city-level native wage and employment growth. To do this we construct an H-1B-driven instrument that retains only the portion of growth in foreign STEM attributable to national policy fluctuations, and our regressions account for city-specific factors that may have attracted foreign STEM and native workers alike.

B. The H-1B-Driven Increase in STEM

Our data on the occupations, employment, wages, age, and education of individuals come from the Ruggles et al. (2010) Integrated Public Use Microdata Series (IPUMS) 5% census files for 1980, 1990, and 2000; the 1% ACS sample for 2005; and the 2008–10 3% merged ACS sample for 2010. We use data only on 219 MSAs consistently identified from 1980 through 2010. These span a range of US metropolitan sizes, including all the largest cities in the United States down to MSAs with close to 200,000 people (Danville, VA, Decatur, IL, Sharon, PA, Waterbury, CT, Muncie, IN, and Alexandria, PA, are the six smallest). Data on aggregate H-1B flows by nationality and year are publicly available from the US Department of State (2012).

We construct our H-1B-driven increase in STEM workers variable for each city between 1990 and 2010. This captures supply-driven variation in the growth of foreign STEM workers, which we use as an instrumental variable to estimate equation (1). To create this instrument, we first impute the number of foreign STEM workers in city c and year t:

$$\widehat{\mathrm{STEM}}_{ct}^{\mathrm{FOR}} = \sum_{n=1}^{14} \mathrm{STEM}_{c1980}^{\mathrm{FOR}_n} \left(\frac{\widehat{\mathrm{STEM}}_{t}^{\mathrm{FOR}_n}}{\mathrm{STEM}_{1980}^{\mathrm{FOR}_n}} \right). \tag{2}$$

[5] Kerr and Lincoln (2010) and Kato and Sparber (2013) provide more discussion on the H-1B visa and its economic effects.

The term $\mathrm{STEM}_{c1980}^{\mathrm{FOR}_n}$ is the number of foreign STEM workers of nationality n in city c in 1980.[6] The growth factor of all foreign STEM workers for each nationality in the United States between 1980 and year t is represented by $\widehat{\mathrm{STEM}}_t^{\mathrm{FOR}_n}/\mathrm{STEM}_{1980}^{\mathrm{FOR}_n}$. This is calculated by adding the inflow of STEM workers from each nationality between 1980 and t to its initial 1980 level. For the period 1980–90, we simply add the net increase in STEM workers from nationality n as recorded in the US census ($\Delta\mathrm{STEM}_{1980\text{–}90}^{\mathrm{FOR}_n}$). For later periods we use the cumulative H-1B visas allocated to each nationality ($\#\text{ of H1B}_{1990-t}^{\mathrm{FOR}_n}$).[7] The imputed growth factor for STEM workers for each foreign nationality in year t is therefore

$$\frac{\widehat{\mathrm{STEM}}_t^{\mathrm{FOR}_n}}{\mathrm{STEM}_{1980}^{\mathrm{FOR}_n}} = \frac{\mathrm{STEM}_{1980}^{\mathrm{FOR}_n} + \Delta\mathrm{STEM}_{1980\text{–}90}^{\mathrm{FOR}_n} + \#\text{ of H1B}_{1990-t}^{\mathrm{FOR}_n}}{\mathrm{STEM}_{1980}^{\mathrm{FOR}_n}}. \tag{3}$$

The H-1B-driven change in foreign STEM workers that we use as our instrument is the time period change in $\widehat{\mathrm{STEM}}_{ct}^{\mathrm{FOR}}$ standardized by the initial imputed city employment ($\widehat{E_{ct}}$).[8]

Our identification strategy is closely related to those used by Altonji and Card (1991) and Card (2001), who exploit the initial distribution of foreign workers across US cities. We use the initial distribution of foreign STEM workers across cities rather than all immigrants. In this regard, our

[6] We aggregate to 14 nationality groups: Canada, Mexico, rest of the Americas (excluding the United States), western Europe, eastern Europe, China, Japan, Korea, Philippines, India, rest of Asia, Africa, Oceania, and other. We choose 1980 as the base year in the imputation of foreign STEM for three reasons. First, it is the earliest census that allows the identification of 219 metropolitan areas. Second, it occurs well before the creation of the H-1B visa and hence does not reflect the distribution of foreign STEM workers affected by the policy. Third, it predates most of the ICT revolution so that the distribution of STEM workers was hardly affected by the geographic location of the computer and software industries.

[7] Data on visas issued by nationality begin in 1997. While we know the total number of H-1B visas issued in each year from 1990, we must estimate the total number of visas issued by nationality between 1990 and 1996 as

$$\#\text{ of }\widehat{\text{H1B}}_{n,1990-t} = \#\text{ of H1B}_{1990-t}\left(\frac{\#\text{ of H1B}_{n,1997\text{–}2010}}{\#\text{ of H1B}_{1997\text{–}2010}}\right),$$

where $\#\text{ of H1B}_{n,1997\text{–}2010}/\#\text{ of H1B}_{1997\text{–}2010}$ is the share of visas issued to nationality group n among the total visas issued from 1997 to 2010. For t larger than 1997, we have the actual number of yearly visas by nationality.

[8] To avoid endogenous changes in total employment at the city level, we also impute city employment by augmenting employment by nativity and skill level in 1980 by the corresponding growth factor in total national employment. Hence, $\widehat{E_{ct}^x} = E_{c1980}^x \times (E_t^x/E_{1980}^x)$, where x is native college-educated workers, native non-college-educated workers, foreign college-educated workers, and foreign non-college-educated workers. Thus, $\widehat{E_{ct}} = \Sigma_x \widehat{E^x}$, and the instrument is $\Delta\widehat{\mathrm{STEM}}_{ct}^{\mathrm{FOR}}/\widehat{E_{ct}}$.

methodology is more similar to Kerr and Lincoln's (2010) examination of the impact of H-1B flows on innovation. We distinguish our approach by using the foreign STEM presence in 1980, rather than in 1990, and by further differentiating immigrant groups by nationality, instead of using aggregate immigrants. We also use a more demanding panel specification, measuring variables in growth rates while including both city and time period effects. Before discussing the validity of our instrumental variables approach in detail, we present descriptive statistics that illustrate the significance of foreign-born STEM workers and the importance of the H-1B program in transforming the US STEM workforce.

C. Foreign STEM Summary Statistics

Foreign-born individuals have been persistently overrepresented in STEM occupations and have contributed substantially to the aggregate growth of STEM jobs in the United States.[9] Table 1 displays the foreign-born share of four different employment groups. Columns 1–4 represent the foreign-born percentage among total employment, college-educated workers, STEM occupations, and college-educated STEM workers—all calculated for the aggregate of 219 MSAs that we analyze. While foreign-born individuals represented about 16% of total US employment in 2010, they counted for more than 27% of college-educated STEM workers in the MSAs we analyze. This percentage has more than doubled since 1980.

Columns 1 and 2 of table 2 show that college-educated STEM workers have increased from 1.7% of total employment in 1980 to 3.2% in 2010. The share of college-educated foreign STEM workers has grown from 0.2% to 0.87%. Of the 0.78 point increase in college-educated STEM as a percentage of employment between 1990 and 2010, 0.53 percentage points (two-thirds of the total) were due to foreigners.

Columns 3–5 display changes in STEM employment and H-1B visas between periods. Column 3 reports the net total increase in college-educated STEM workers in the United States over the periods, and column 4 displays the rise in college-educated foreign STEM workers. While only one-fifth of the net increase in STEM workers between 1980 and 1990 was driven by foreigners, they were responsible for 77% of the net STEM growth between 1990 and 2000 and for more than the total growth from 2000 to 2010. Column 5 displays the cumulative number of H-1B visas issued between periods. It is clear that enough H-1B visas were issued to cover the whole growth in college-educated foreign STEM workers in the United States. Remarkably, H-1B issuances were three to four times as large as the net increase in college-educated STEM between 2000–2005 and 2005–10. This implies that many foreign STEM workers,

[9] In the summary statistics and in the empirical analysis we mainly use the O*NET 4% STEM definition unless we note otherwise.

Table 1
Summary Statistics: Percentage of Foreign-Born by Group, 219 MSAs

	Total Employment (%) (1)	College-Educated Employment (%) (2)	Employment in STEM Occupations (%) (3)	College-Educated Employment in STEM Occupations (%) (4)
1980	6.15	6.81	8.14	11.09
1990	8.82	8.95	10.98	14.24
2000	13.31	12.80	17.47	22.69
2005	15.37	14.81	20.03	25.76
2010	16.37	15.46	21.19	27.15

Note.—The figures are obtained by the authors' calculations using IPUMS census data from 1980–2010. The relevant population includes only noninstitutionalized individuals between ages 18 and 65 who have worked at least 1 week in the previous year and report identified occupations. The statistics exclude those with unknown, unreported, or military occupations and individuals without a clearly identified birthplace who do not possess US citizenship through parents with US citizenship. STEM occupations are defined according to the O*NET 4% definition. College-educated workers have a bachelor degree or higher. The sample comprises 219 consistently identified MSAs from 1980–2010.

Table 2
Shares and Absolute Net Changes in STEM Employment, 219 MSAs

	Share of Employment (%)		Net Absolute Change from Previous Period (1,000s)		
Period	College-Educated Total STEM (1)	College-Educated Foreign STEM (2)	College-Educated Total STEM (3)	College-Educated Foreign STEM (4)	H-1B Visas Issued (5)
1980	1.76	.19			
1990	2.42	.34	915	218	0
2000	2.99	.68	670	518	574
2005	3.01	.77	109	208	659
2010	3.20	.87	164	146	653

Note.—The figures in cols. 1–4 are obtained by the authors' calculations on data from 219 consistently identified MSAs in IPUMS census data from 1980–2010. The relevant population includes only noninstitutionalized individuals between ages 18 and 65 who have worked at least 1 week in the previous year and report identified occupations. The statistics exclude those with unknown, unreported, or military occupations and individuals without a clearly identified birthplace who do not possess US citizenship through parents with US citizenship. STEM occupations are defined according to the O*NET 4% definition. College-educated workers have a bachelor degree or higher. Data on the total number of H-1B and TN visas issued (col. 5) are from the Department of State (2012). H-1B numbers also include TN visas and are relative to the whole United States.

including H-1B recipients, have left the United States.[10] Overall, table 2 highlights the importance of foreign workers within STEM jobs and confirms that the scope of the H-1B program was large enough to substantially contribute toward foreign STEM growth since 1990.

[10] Depew, Norlander, and Sorensen (2013) provide a detailed analysis of quit and return rates for temporary skilled employees of six large Indian ICT firms. During the course of the survey period (2003–11), 29% of their sample returned to India.

IV. Identification: Power and Validity of the Instruments

Our identification strategy relies on the H-1B supply-driven instrument. Its validity is based, in large part, on the assumption that the 1980 employment share of foreign STEM workers varied across cities because of factors related to the persistent agglomeration of foreign communities in some localities. These historical differences—after controlling for an array of other city characteristics and shocks—affected the change in the supply of foreign STEM workers but were unrelated to shocks affecting city-level native wage and employment growth. Though our modeling choices aim to reduce the risk of correlation between the instrument and unobserved determinants of wage and employment growth, such confounding factors are of great concern. For example, the initial distribution of foreign STEM may be correlated with persistent city factors that influenced future labor market outcomes, resulting in omitted variables bias. Alternatively, aggregate inflows of H-1B workers might have been driven by a few specific cities. The presence of measurement error, more likely in cities with small populations, could lead to attenuation bias. This section tests our instrument's validity and addresses key challenges to our identification strategy.

The following first-stage regression provides a framework to explore these issues:

$$\frac{\Delta \mathrm{STEM}_{ct}^{\mathrm{FOR}}}{E_{ct}} = \phi_t + \phi_c + \beta \cdot \frac{\widehat{\Delta \mathrm{STEM}}_{ct}^{\mathrm{FOR}}}{\widehat{E_{ct}}} + \varepsilon_{ct}. \tag{4}$$

The coefficient β measures the impact of H-1B-driven STEM inflows—our instrument—on the measured increase in foreign STEM workers, the explanatory variable in our second-stage regression (1). This coefficient and its power are the main objects of interest for causal interpretation. The terms ϕ_t and ϕ_c capture period and MSA fixed effects. Changes refer to the periods 1990–2000, 2000–2005, and 2005–10. The zero-mean random error (ε_{ct}) is uncorrelated with the explanatory variable.

A. Basic Specifications and Checks

We tackle several threats to the identification assumptions and begin by showing that the 1980 presence of foreign STEM workers in cities did not always mirror the presence of native STEM workers. Table 3 shows the estimated coefficient (β) and the partial F-statistic from first-stage regression equation (4). The coefficients reported in the first and the second rows are the β and the F-statistics of the instrument when using the O*NET STEM definition for both the endogenous variable and the instrument. Those in the third and fourth rows are the corresponding statistics when using the major-based STEM definition.

Column 1 includes period effects, state effects, and the 1980 employment share of native STEM. Imputed H-1B-driven STEM growth has a

Table 3
First Stage: Power and Validity of H-1B-Driven STEM as an Instrumental Variable

Explanatory Variable	Strict (4%) Definition of STEM (with State Fixed Effects) (1)	Baseline: Strict (4%) Definition of STEM (with City Fixed Effects) (2)	Broad (8%) Definition of STEM for Both Endogenous Variable and Instrument (3)	As Col. 2, Excluding the 5 Cities with Largest Number of STEM Workers (4)	As Col. 2, Excluding STEM from India (5)	As Col. 2, Imputation Using Aggregate H-1B Visas (Not by Nationality) (6)	As Col. 2, Excluding Cities with Population ≤400,000 (7)	As Col. 2, Controlling for Imputed Non-College-Educated Immigrants (8)	As Col. 8, Controlling for Bartik Employment and Wage Growth (9)
H-1B-driven growth in foreign STEM, O*Net:									
Coefficient	.48***	2.56***	4.06***	1.82**	3.53***	3.02***	3.22***	2.48***	2.34***
	(.18)	(.88)	(1.29)	(.83)	(.91)	(.75)	(.86)	(.90)	(.92)
F-statistic	6.57	8.51	9.95	4.86	15.09	16.43	14.04	7.59	6.46
H-1B-driven growth in foreign STEM, major-based:									
Coefficient	.44***	2.83***	4.23***	2.13***	3.59***	3.26***	3.34***	2.79***	2.42**
	(.16)	(.84)	(.93)	(.63)	(1.00)	(1.04)	(1.08)	(.86)	(.88)
F-statistic	7.73	11.32	20.53	11.27	12.91	9.86	9.48	10.64	7.50
Fixed effects	State and period	City and period	City and period	City and period	City and period	City and period	City and period	City and period	City and period
Observations	657	657	657	642	657	657	354	657	657
Metro areas	219	219	219	214	219	219	118	219	219

NOTE.—Each cell shows the coefficient from a different regression. The dependent variable is the growth in foreign STEM as a percentage of the labor force. The units of observations are 219 US metropolitan areas over the periods 1990–2000, 2000–2005, and 2005–10. The explanatory variable is the H-1B-driven growth of foreign STEM jobs, as a percentage of initial employment. The top 2 rows use the O*NET-based definition of STEM occupations. The third and fourth rows use major-based STEM definitions. Baseline models use the narrow (4%) definition of STEM. Column 1 also controls for a city's native STEM employment in 1980. Standard errors (in parentheses) are always clustered at the metro area level.

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

highly significant impact on foreign STEM growth. This implies that even controlling for the initial native STEM share, the foreign STEM share has significant explanatory power.[11]

The next two columns introduce MSA fixed effects to control for all other initial city-specific conditions so that our identification relies only on deviations in MSA growth rates from MSA-specific trends. We include city fixed effects in all subsequent specifications. Column 2 uses the narrow 4% (STEM or major-based) definitions for both the endogenous variable and the instrument, whereas column 3 uses the broader 8% definitions. The power of the instrument in these specifications is stronger than in column 1. The F-statistics are close to or above 10, emphasizing that our H-1B-based instrument is good at capturing changes in the inflow of STEM workers within cities over time. Moreover, we find that the two definitions of STEM produce similar results, though some small differences exist.

Columns 4 and 5 of table 3 address two important concerns. The first is that the correlation between the instrument and the actual change in foreign STEM could be driven by the large high-tech boom in a few large MSAs rather than by the exogenous initial distribution of immigrants. If large metropolitan areas drove most of the country's R&D and produced a large increase in demand for foreign H-1B visas and STEM workers, the instrument and the endogenous variable for large R&D-intensive cities could be spuriously correlated. Alternatively, the presence of a few particular industries (e.g., the ICT sector) might have attracted particular types of immigrants whose growth simply proxies for the success of those industries. The current population of foreign STEM workers from India, for example, is strongly associated with information technology since most of them are employed in computer, software, and electrical engineering occupations. Moreover, Indians have always accounted for at least 40% of H-1B visas.

Column 4 excludes the five metro areas with the largest number of STEM workers in 1980.[12] Column 5 excludes Indian STEM workers from the calculations of the instrument. The coefficients are still highly significant (although somewhat reduced in col. 4 for O*NET STEM), indicating that the correlation between H-1B-driven STEM growth and a city's actual foreign STEM growth is not driven by top STEM cities or by a specific nationality group.

An alternative way to ensure that the predictive power of our instrument is not driven by individual nationality groups—whose location preferences

[11] One reason for the power of foreign STEM after controlling for native STEM is that cities with large native STEM shares in 1980 were associated with traditional sectors that attracted scientists and engineers in the 1970s but did not predict the presence of information technology and computer sectors that dominated R&D in the 1990s and 2000s.

[12] New York, Los Angeles, Chicago, San Jose, and San Francisco account for 24% of STEM workers in our sample.

may be affected by specific industries—is to remove the nationality dimension. We construct an instrument similar to the one used by Kerr and Lincoln (2010) by exploiting only variation in the aggregate number of H-1B visas over time, interacted with the initial overall presence of foreign STEM workers. First-stage results using this instrument are shown in column 6. The estimates remain similar, and *F*-statistics confirm that the instrument retains its power.

Column 7 accounts for another potential weakness of our instrument. The use of 1%–5% population samples may introduce measurement error. Aydemir and Borjas (2011) show how measurement error can produce attenuation bias when estimating the causal effect of immigrants on native outcomes. Small census and ACS samples might fail to record small foreign STEM communities in small cities. In order to see whether this measurement error affects the power of our instrument, column 7 shows the first-stage estimates when eliminating all metropolitan areas with fewer than 400,000 people. This cutoff eliminates all cities from our sample that have a measured zero foreign STEM (or imputed foreign STEM) employment share. Although we retain only 118 of the 219 cities, the coefficient estimates remain significant and stable, while the instrument is still reasonably powerful (more so for the O*NET STEM definition). While we will discuss the potential impact of measurement error on attenuation bias when presenting the second-stage estimates (in table 5 below), it is reassuring that the exclusion of the cities in which measurement error is most likely hardly affects the power of the instrument and the first-stage coefficient estimate.

B. Confounding Shocks

Two types of shocks at the MSA level might be correlated with the inflow of STEM workers, wages, and employment, thereby creating omitted variable bias. The first is a change in the skill distribution of workers related to the inflow of non-STEM immigrants. The second is an industry-driven change in productivity affecting native employment and wages. Directly controlling for such shocks would introduce endogeneity. Instead, we include predicted values formed by interacting the 1980 immigrant and industry distributions with national immigrant and industry shocks, respectively.

As STEM immigrants usually earned a college degree, we introduce a control for the imputed number of non-college-educated immigrants ($\widehat{\text{NoColl}}_{ct}^{\text{FOR}}$) based on their 1980 distribution, by nationality, across metropolitan areas ($\text{NoColl}_{c1980}^{\text{FOR}_n}$) and their subsequent aggregate growth in the United States ($\text{NoColl}_{t}^{\text{FOR}_n}/\text{NoColl}_{1980}^{\text{FOR}_n}$). Using notation similar to (2), we use equation (5) to calculate $\widehat{\text{NoColl}}_{ct}^{\text{FOR}}$ and then construct our control by taking the change over time relative to total initial imputed employment ($\Delta\widehat{\text{NoColl}}_{ct}^{\text{FOR}}/\widehat{E}_{ct}$):

$$\widehat{\text{NoColl}}_{ct}^{\text{FOR}} = \sum_{n=1,14} \text{NoColl}_{c1980}^{\text{FOR}_n} \left(\frac{\text{NoColl}_t^{\text{FOR}_n}}{\text{NoColl}_{1980}^{\text{FOR}_n}} \right). \tag{5}$$

To control for shocks driven by a city's industrial structure, we construct Bartik instruments (from Bartik [1991]) that predict the wage and employment growth of college- and non-college-educated workers based on each city's industrial composition in 1980. Specifically, let $s_{ic,1980}$ denote the share of total city employment in each three-digit census industry classification sector ($i = 1, 2, \ldots, 212$) in 1980. Then let $\Delta y_t^{i,X}/y_t^{i,X}$ be the real growth of $y = \{\text{Wage, Employment}\}$ over the decade for group $X = \{\text{College, NoCollege}\}$ in sector i. We define our sector-driven Bartik variables as

$$\left(\frac{\Delta y^X}{y^X} \right)_{ct}^{\text{Sector-Driven}} = \sum_{s=1}^{212} \left(s_{ic,1980} \frac{\Delta y_t^{i,X}}{y_t^{i,X}} \right). \tag{6}$$

Column 8 of table 3 adds the imputed growth of non-college-educated immigrants to the basic first-stage regression of column 2. Cities with large communities of less educated immigrants might also have large communities of highly educated immigrants, although usually from different nationalities. Controlling for these flows will also be important to account for complementarities between college- and non-college-educated workers and their possible effect on wages in the second-stage regressions. Nonetheless, the imputed H-1B-driven instrument retains its power when controlling for the imputed number of non-college-educated immigrants. Column 9 further adds the employment and the wage Bartik instruments. This still leaves the H-1B imputed STEM growth instrument with significant, albeit somewhat reduced, explanatory power, especially when using the O*NET definition.

C. Falsification and Extensions

Our instrument is predicated on two assumptions. First, from the perspective of each metropolitan area, the H-1B visa policy significantly and exogenously affected the inflow of foreign STEM workers to the United States from 1990 to 2010. Second, the initial distribution of foreign STEM was crucial in determining the subsequent city-level inflow of H-1B immigrants and was uncorrelated with other city-level shocks affecting native wages and employment. Columns 1–4 of table 4 test these assumptions.

The aggregate inflow of H-1B workers in the United States could simply be a proxy for aggregate labor demand growth and not policy-driven supply changes. This could induce a positive correlation between the instrument and the explanatory variable even in the presence of city and period effects. Note, however, that this scenario would also imply a positive correlation between the explanatory variable and a falsified instrument constructed by substituting non-H-1B immigrant flows (or non-college-educated immigrant flows) for H-1B flows. Columns 1 and 2 show that the first-stage point estimates are insignificant and close to zero when we impute foreign STEM growth by interacting the 1980 distribution of foreign STEM with subsequent noncollege immigrant flows (col. 1) or with aggregate immigrant flows net of H-1B flows (col. 2). Hence, the aggregate variation of H-1B visas over time is crucial for predicting subsequent STEM variation across cities. The two "falsified instruments" used in these specifications, therefore, do not covary with foreign STEM changes because they do not incorporate the variation in H-1B aggregate visas. Column 3 similarly finds no evidence of correlation when we substitute the initial presence of foreign workers in manual-intensive jobs (rather than in STEM) across metropolitan areas in the construction of the instrument. Therefore, less skilled immigration—though possibly correlated with STEM immigration—did not drive the explanatory power of the instrument. These results reassure that our preferred policy-driven instrument is not simply reflecting aggregate labor demand or aggregate migration.[13]

Column 4 tests the correlation between the instrument—calculated for the 1990–2000 decade—and the preexisting growth in native college wages from 1970 to 1980. Reassuringly, there is no correlation between the H-1B imputed STEM growth after 1980 and pre-1980 native wage growth despite, as will be seen in Section V, the strong relationship between increased STEM during the 1990s and 2000s and concurrent wage growth. This test ensures that the pre-H-1B (pre-1980) outcomes across MSAs were not correlated with the post-1990 H-1B-driven STEM growth.

As a final check in this section, we explore how H-1B policy affects the total number of STEM workers and, specifically, whether metropolitan areas with large foreign STEM inflows substitute foreign STEM for native STEM or instead increase the overall STEM labor force. If the latter is true, we can consider the H-1B policy as an exogenous shock to assess the impact of total STEM on native wages, employment, and productivity. Columns 5 and 6 examine this by regressing native plus foreign STEM worker growth on the H-1B-predicted inflow of foreign STEM (the instrument). The estimated coefficient is even larger than in the basic specification, implying, as we will see below, a positive response of native STEM to foreign inflows. In column 5, we use the stricter 4% STEM definition (based on O*NET in the top rows and on college major in the two lower rows) for both the endogenous and instrumental variables. In column 6, we use the broader 8% definition of STEM for the endogenous and instrumental variables. The power of the instrument is relatively strong in most cases.

Overall, the specifications and falsifications shown in this section demonstrate that our H-1B imputed instrument has significant power in

[13] The online appendix details the construction of these falsified instruments.

Table 4
First Stage: Falsification and Extensions

Explanatory Variable	Falsification: Endogenous Variable			Dependent Variable		
	Growth of Foreign STEM O*NET 4% (1)	Growth of Foreign STEM O*NET 4% (2)	Growth of Foreign STEM Major-Based 4% (3)	1970–80 College-Educated Native Wages; Explanatory Variable, 1990–2000 (4)	Total STEM Growth (Native + Immigrant), 4% STEM Definition (5)	Total STEM Growth, 8% STEM Definition (6)
Predicted Foreign STEM, O*NET definition:						
Coefficient				1.66	5.03***	8.50**
				(1.02)	(1.81)	(4.40)
F-statistic				2.62	7.70	3.65
Predicted foreign STEM, major-based definition:						
Coefficient				1.12	5.29***	9.63***
				(1.10)	(1.71)	(2.56)
F-statistic				1.05	9.54	14.05

	(1)	(2)	(3)	(4)	(5)	(6)
Predicted growth in foreign STEM using flows of noncollege immigrants:						
Coefficient	.04					
	(.03)					
F-statistic	2.24					
Predicted growth in foreign STEM using flows of total immigrants minus H-1B:						
Coefficient		.043				
		(.024)				
F-statistic		3.08				
Predicted growth in foreign STEM using 1980 distribution of manual immigrants:						
Coefficient			.41			
			(.27)			
F-statistic			2.42			
Observations	657	657	657	116	657	657
Metro areas	219	219	219	116	219	219

NOTE.—Each cell shows the coefficient from a different regression and below it the F-test of significance. The units of observations are 219 US metropolitan areas over the periods 1990–2000, 2000–2005, and 2005–10. The dependent variable is the growth in foreign STEM in cols. 1–3; the growth in native college-educated wages, 1970–80, in col. 4; and total STEM growth in cols. 5 and 6. The explanatory variables are described at the beginning of the row. Standard errors (in parentheses) are always clustered at the metro area level.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

predicting foreign STEM and total STEM growth, which is not driven by top cities, one ethnic group, or labor demand and survives the inclusion of city effects and controls for industrial composition and low-skilled immigration. The instrument's predictive power is crucially driven by the H-1B program and by the initial distribution of STEM immigrants across cities.

V. The Effect of STEM on Native Outcomes

A. Basic Results

The empirical specifications estimated in this section follow the regression described in equation (1) to identify the impact of STEM workers on native labor market outcomes ($y_{ct}^{\text{Native},X}$) by group X (STEM, college-educated, or non-college-educated) in city c. Outcomes measure growth either in average weekly wages or in employment. The explanatory variable in each regression is the change in foreign STEM relative to the initial level of total employment, $\Delta\text{STEM}_{ct}^{\text{Foreign}}/E_{ct}$. All two-stage least-squares (2SLS) regressions use the H-1B-driven change in foreign STEM relative to initial imputed employment ($\Delta\widehat{\text{STEM}}_{ct}^{\text{FOR}}/\widehat{E_{ct}}$) as an instrument for the actual change.

Each of the six columns of table 5 reports the $b_{y,X}$ coefficient of interest, as defined in equation (1), corresponding to the differing outcome variables. The basic specification includes time period effects, 219 MSA fixed effects, and the Bartik instruments for the relevant wage and employment changes. We always cluster standard errors at the MSA level.

In columns 1–3, the dependent variable is the percentage change of the weekly wage ($\Delta w_X^{\text{Native}}/w_X^{\text{Native}}$) paid to STEM, college-educated, and non-college-educated native-born workers, respectively.[14] We define college-educated workers as individuals who completed 4 years of college, while non-college-educated are those who did not. Columns 4–6 show the effect of STEM on the employment change of these native-born groups as a percentage of total city employment (respectively, $\Delta\text{STEM}_{ct}^{\text{Native}}/E_{ct}$, $\Delta H_{ct}^{\text{Native}}/E_{ct}$, and $\Delta L_{ct}^{\text{Native}}/E_{ct}$).

The different rows of table 5 represent different specifications to test the robustness of the estimates, mirroring in large part the first stage in table 3. Row 1, the baseline specification, shows the results when the O*NET 4% definition of STEM workers is used for both the explanatory variable and the instrument. Row 2 instead uses the major-based 4% definition of

[14] Weekly wages are defined as yearly wage income divided by the number of weeks worked. Employment includes all individuals between 18 and 65 years old who have worked at least 1 week during the previous year and do not live in group quarters. We convert all wages to current 2010 prices using the Bureau of Labor Statistics Inflation Calculator. See the online appendix for full details on the sample selection process.

Table 5
The Effects of Foreign STEM on Native Wages and Employment

Explanatory Variable: Growth Rate of Foreign STEM	Weekly Wage, Native STEM (1)	Weekly Wage, Native College-Educated (2)	Weekly Wage, Native Non-College-Educated (3)	Employment, Native STEM (4)	Employment, Native College-Educated (5)	Employment, Native Non-College-Educated (6)
1. Baseline 2SLS; O*NET 4% definition	6.65	8.03***	3.78**	.53	2.48	5.17
	(4.53)	(3.03)	(1.75)	(.56)	(4.69)	(4.20)
2. 2SLS; major-based 4% definition	6.64	10.95**	3.22**	.60	1.05	7.82
	(5.08)	(4.34)	(1.67)	(.63)	(3.99)	(4.90)
3. 2SLS; O*NET 8% definition	7.23**	5.64***	2.55**	.53	1.85	4.14
	(3.52)	(1.95)	(1.08)	(.75)	(3.21)	(3.32)
4. Omitting top 5 STEM cities	11.35	12.78***	5.03	1.65***	8.46	2.51
	(8.63)	(4.99)	(3.42)	(.53)	(7.04)	(7.46)
5. Controlling for imputed noncollege immigrants	7.94	7.00**	4.95**	.76	3.29	3.39
	(5.38)	(2.98)	(2.09)	(.61)	(4.85)	(4.15)
6. Dropping small cities (population <400,000)	5.70	7.18***	4.28***	.34	−.60	5.20
	(3.51)	(2.61)	(1.45)	(.58)	(1.51)	(3.18)
7. Dropping Indians	3.48	9.38**	3.46*	.47	1.31	6.44*
	(5.07)	(4.37)	(2.08)	(.51)	(3.61)	(3.54)
8. Aggregate H-1B IV	5.76	6.04**	4.13***	.31	1.64	5.56
	(4.05)	(2.75)	(1.34)	(.48)	(4.20)	(3.63)
9. Controlling for imputed college natives	2.72	7.58**	2.39	−.32	−.62	7.29
	(4.68)	(3.78)	(2.00)	(.47)	(4.19)	(5.15)
10. OLS version of specification 5	3.32	4.10***	1.16	.92**	4.97	2.11
	(2.99)	(1.86)	(1.24)	(.34)	(3.69)	(2.51)

Note.—The instrument is the H-1B imputed growth of foreign STEM. Each cell shows the estimate of the coefficient on the growth in foreign STEM (relative to employment) when the dependent variable is the one described at the top of the column. Each regression includes period effects, metropolitan area effects, and the Bartik for employment or wage of the relevant group. Rows 1 and 4–8 are 2SLS regressions using the O*NET 4% definition of STEM. Rows 2 and 3 use alternative definitions of STEM. Row 10 shows the OLS estimates. Standard errors (in parentheses) are clustered at the metro area level. Units of observations are 219 metro areas over three periods: 1990–2000, 2000–2005, and 2005–10.

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

STEM workers, and row 3 uses the broader O*NET 8% definition. Row 4 omits the top five metropolitan areas in terms of STEM employment but is otherwise identical to the specification in row 1. Row 5 adds the growth of imputed non-college-educated immigrants, defined in (5), as a control to the baseline specification. Row 6 excludes MSAs with populations below 400,000. Row 7 excludes Indian STEM workers from the construction of the instrument. Row 8 uses the instrument constructed using aggregate H-1B flows and the initial foreign STEM distribution, thus removing the nationality dimension. Row 9 controls for growth in native college-educated employment by including a shift share instrument for the growth of college-educated natives, constructed by interacting the 1980 number of college-educated natives in each city with the national growth of college-educated natives. Finally, row 10 shows the ordinary least squares (OLS) estimates of the basic specification.

The main results are relatively consistent across specifications. First, there is a large, positive, and significant effect of foreign STEM workers on wages paid to college-educated natives. The estimated effect is significantly different from zero at the 5% significance level in all specifications and is significant at the 1% level in most. The point estimates from the 2SLS specifications are mostly between 5.6 and 9.3, with some larger values. This implies that a rise in foreign STEM growth by 1 percentage point of initial employment increases college-educated native wage growth between 5.6 and 9.3 percentage points.[15]

Second, the estimates of the effects on native STEM wages are comparable to, but less precisely estimated than, the effects on native college-educated wages. While we can never rule out the hypothesis that the estimated effects for the two groups are equal, the native STEM wage effect is only occasionally different from zero at the 5% significance level.[16] As there are fewer STEM natives (about 4% of employment) than college-educated natives (about 25% of employment), measurement error in the average wage of the first group reduces the precision of the estimates.

The third regularity of table 5 is that foreign STEM workers had a positive and usually significant effect on wages paid to non-college-educated natives. Point estimates are mostly between 2.4 and 4.3—results that are both smaller and less significant than those for college-educated natives. This implies that STEM workers generate a productivity effect that is skill

[15] Note that 1 percentage point of employment is a very large increase of STEM workers, comparable to the increase over the whole 1990–2010 period, as shown in table 2.

[16] For instance, a formal test that the estimated coefficient on STEM wages in row 1 is equal to 8.03 (the point estimate for the effect on the college-educated) has a p-value of .76. At no level of confidence can we reject the hypothesis that they are equal. Similarly for the other specifications, we can never reject the hypothesis of equality at the 10% confidence level.

biased. Foreign STEM workers are closer substitutes for college-educated natives than for non-college-educated natives, yet they generate a larger increase in the wages paid to college-educated natives.

Fourth, the inflow of STEM workers did not significantly affect the employment of any native group. The point estimates are mainly positive for native STEM and college-educated workers and mainly negative for non-college-educated natives. However, they are usually not significant, even at the 10% level. Given the mobility of college-educated natives and their city-level wage gain from STEM flows, this weak employment response is somewhat surprising and suggests the potential existence of additional adjustment mechanisms for college-educated workers at the metropolitan area level. In section 5.4 of the working paper version of this study (Peri, Shih, and Sparber 2014), we argue that STEM flows are also associated with increased housing rents for college-educated natives and that this increase in nontradables prices might absorb up to 50% of the college-educated native wage gain. This might help explain the small employment response while cautioning against interpreting the wage gains of table 5 as full increases in total purchasing power.

B. Robustness Checks

We now comment on the robustness checks performed in table 5. To mitigate endogeneity concerns discussed earlier, row 4 omits the top five STEM-dependent cities and row 7 removes Indian workers. The estimated effects of STEM on native wages remain stable and even increase in some cases, albeit at the cost of larger standard errors. On one hand, this suggests that the fixed effects, instrumental variable strategy, and Bartik controls in the baseline model largely address endogeneity bias. On the other hand, the increase in standard errors indicates that the omitted cities, when included in regressions, afford precision in the estimates due to larger data variation.

Row 5 adds a control for imputed low-skilled immigrants. As above, this also results in minimal changes in the coefficient estimates when compared to row 1. The estimated STEM effect on college-educated wages is somewhat smaller (down to 7.00 from 8.03), and the coefficient for non-college-educated wages is somewhat larger (up to 4.95 from 3.78). This could indicate that the inflow of less educated immigrants, as predicted by the 1980 MSA distribution, was slightly correlated with foreign STEM and that less educated labor inflows complemented college-educated natives but substituted for non-college-educated ones. Explicitly controlling for such imputed inflows helps to isolate the effect of STEM and identifies more balanced productivity effects for college- and non-college-educated natives.

Similarly, a large initial share of foreign STEM in a city might proxy for high initial education levels. If such cities also experienced wage and employment growth during periods of sizable foreign STEM inflows, it would generate spurious regression results. Row 9 includes a shift share predictor of college-educated native growth to help address this issue. The estimated STEM impact on wages paid to college-educated natives remains quantitatively similar to baseline estimates and is still statistically significant.

Row 6 omits small cities to examine measurement error issues. The point estimates are similar to those in row 5, but the standard errors decrease. Hence, measurement error does not seem to bias the coefficients, but the focus on large MSAs reduces measurement error and improves precision.

Finally, it is worth commenting on the difference between the OLS estimates in row 10 and the corresponding 2SLS results in row 5. Interestingly, while the estimated employment effects have an upward bias in OLS relative to 2SLS, the wage effects have a downward bias. This may be due to the correlation between unobserved shocks and the inflow of foreign STEM. It is likely that foreign STEM inflows are positively correlated with employment growth and a city's openness to new workers. Hence, the cities endogenously attracting foreign STEM workers could be those with fast inflows of workers in general, which could moderate wage growth. Thus, the correlation between STEM growth and omitted employment determinants could be positive, and the correlation between openness and wage growth could be negative, thereby resulting in the observed biases.

Before extending the findings, we provide a sense of the magnitude of the estimated effects. Foreign STEM growth, measured as a percentage of total initial employment in aggregate, was only about 0.53% between 1990 and 2010. Applying the 7.00 2SLS estimates of row 5 to the national growth in foreign STEM implies that the foreign-driven net growth in STEM increased real wages of college-educated natives by around 3.71 percentage points ($= 7.00 \times 0.53$) during this period. For reference, census data suggest that the cumulative growth of college-educated wages in this period equaled about 13 percentage points. Thus, almost one-third of that growth can be attributed to the increased presence of foreign STEM workers. We return to these implications in Section VI when we analyze the implied productivity and skill-bias effects of STEM.

C. Extensions

As shown in the first-stage results in columns 5 and 6 of table 4, our H-1B-driven increase in the STEM instrument raises overall STEM employment, not just foreign STEM. Table 6 generalizes the main second-stage results by replacing the foreign STEM growth explanatory variable with total STEM growth. The estimates confirm that STEM workers generate wage gains for college-educated and non-college-educated natives. More specifically, using the estimates in row 1 of table 6, a 1 percentage point increase in STEM as a share of employment caused a 4 percentage point

Table 6
The Effects of Total STEM on Native Wages and Employment

Explanatory Variable: Growth Rate of Total STEM	Weekly Wage, Native STEM (1)	Weekly Wage, Native College-Educated (2)	Weekly Wage, Native Non-College-Educated (3)	Employment, Native College-Educated (4)	Employment, Native Non-College-Educated (5)
1. 2SLS; O*NET	4.50	3.97***	2.44**	1.86	−1.67
4% definition	(2.94)	(1.42)	(1.02)	(2.31)	(2.34)
2. 2SLS; major-based	4.90	5.68**	2.40**	1.15	−2.92
4% definition	(3.41)	(2.42)	(1.00)	(2.10)	(2.82)
3. 2SLS; O*NET,	4.55	2.64*	1.67**	1.46	−1.23
8% definition	(3.01)	(1.43)	(.76)	(1.25)	(1.79)
4. Same as row 1 but	4.50	4.03**	1.97*	3.23	−.28
omitting top 5 STEM cities	(3.39)	(1.74)	(1.12)	(2.38)	(2.33)
5. OLS; O*NET	.37	.73	.75*	2.72***	4.60***
4% definition	(1.08)	(.54)	(.40)	(.77)	(.79)

NOTE.—Each cell shows the estimate of the coefficient on the growth in total STEM (relative to employment) when the dependent variable is the one described at the top of the column and the instrument is the H-1B driven STEM growth. Each regression includes period effects, metropolitan area effects, the Bartik for employment and wage of the relevant group, and the imputed growth of non-college-educated immigrants. Standard errors (in parentheses) are clustered at the metro area level. Units of observations are 219 metro areas over three periods: 1990–2000, 2000–2005, and 2005–10.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

increase in college-educated native wage growth and about a 2.4 percentage point wage growth for non-college-educated natives. There is no evidence that either group experiences an employment effect.

These results are robust to using the major-based definition of STEM (row 2), using the broad (8%) definition of STEM (row 3), and omitting top STEM cities (row 4). Also, the OLS estimates continue to exhibit a positive bias (relative to the 2SLS results) for employment effects and a negative one for wage effects. Overall, our estimates confirm that STEM workers raise the demand for college-educated and non-college-educated natives, with a smaller effect for the latter group.

A lot of heterogeneity exists among non-college-educated workers. Table 7 explores whether the wage and employment effects of foreign STEM workers are different for natives without a high school diploma (high school dropouts) and those with a high school diploma (high school graduates). The table presents foreign STEM effects for wages (cols. 1 and 2) and employment (cols. 3 and 4). Rows 1–4 present several specifications of the 2SLS regression mirroring those in the corresponding rows of tables 5 and 6. Row 5 reports the coefficients when using total STEM as the explanatory variable.

Table 7
The Effect of Foreign STEM on Non-College-Educated Natives

Explanatory Variable: Growth Rate of Total STEM	Weekly Wage, Native High School Graduates (1)	Weekly Wage, Native High School Dropouts (2)	Employment, Native High School Graduates (3)	Employment, Native High School Dropouts (4)
1. 2SLS; O*NET	5.54**	3.30	3.36	−.03
4% definition	(2.33)	(4.26)	(3.72)	(.55)
2. 2SLS; major-based	4.87**	5.97	5.12	−.50
4% definition	(2.10)	(4.67)	(4.08)	(.66)
3. 2SLS; O*NET	4.10**	2.45	2.48	−.02
8% definition	(1.70)	(3.01)	(2.88)	(.40)
4. Same as row 1	7.05*	6.28	1.38	.50
but dropping	(4.29)	(7.58)	(6.65)	(.96)
top 5 STEM cities				
5. Explanatory variable:	2.73**	1.63	1.65	−.02
total STEM,	(1.15)	(1.99)	(2.12)	(.27)
O*NET 4%				

NOTE.—Each cell in rows 1–4 shows the estimate of the coefficient on the growth in foreign STEM (relative to employment) when the dependent variable is the one described at the top of the column. Row 5 shows the estimate of the coefficient on the growth in total STEM (relative to employment) as the explanatory variable, still instrumented with H-1B imputed growth of foreign STEM. Each regression includes period effects, metropolitan area effects, the Bartik for employment and wage of the relevant group, and the imputed growth of non-college-educated immigrants. Standard errors (in parentheses) are clustered at the metro area level. Units of observations are 219 metro areas over three periods: 1990–2000, 2000–2005, and 2005–10.

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

By separating high school graduates from high school dropouts, we can check whether these two groups exhibit different complementarities with foreign STEM labor. On the one hand, STEM-generated innovation could be skill biased, complementing educational attainment (see Acemoglu 1998, 2002). If so, then foreign STEM would generate the largest positive effects for college-educated workers, followed by high school graduates and, finally, by high school dropouts. On the other hand, it could be polarizing, substituting for intermediate skills but complementing low- and high-end skills (see Autor et al. 2006; Autor 2010). If so, then foreign STEM would generate the largest positive effects at the high and low ends of the educational spectrum at the expense of intermediate-level levels of schooling.

Table 7 shows that STEM effects are significant only for high school graduates, while point estimates for dropouts are smaller but insignificant. Neither group had significant employment effects. The basic specification in row 1 shows that each percentage point increase in foreign STEM employment raised native high school graduate wage growth by 5.54 percentage points. This can be interpreted as evidence that STEM-driven technological progress has been skill (or schooling) biased rather than polarizing.

The difference between the effects on high school graduates and dropouts is not usually significant, however, because of the lack of precision in estimating the effects for dropouts.

VI. Simulated Productivity and Skill Bias Effects

We close our analysis by estimating the long-run effect of STEM on total factor productivity (TFP) and skill-biased productivity (SBP). More specifically, we assume a basic structural model of production and substitute parameter values from our analysis, observed data, and other sources, and then we simulate the TFP and SBP effects that can be explained by growth in foreign STEM workers. The advantage of this approach is that we have an intuitive and standard definition of TFP and SBP based on a city-specific production function. The limitation is its dependence on the assumed nature of productive interactions between different types of labor inherent to the specific production structure.

A full model and derivation are available in the online appendix. Here, we provide just a simple production function and the intuition of the exercise. Suppose that a city (c) produces a homogeneous, tradable, numeraire product (Q_{ct}) in year t. The economy employs three types of labor: non-college-educated (L_{ct}); college-educated, non-STEM (H_{ct}); and STEM workers (ST_c). Production occurs according to the long-run production function in (7):

$$Q_{ct} = (A(ST_{ct})\{\beta(ST_{ct})K_{ct}^{(\sigma_H-1)/\sigma_H} + [1-\beta(ST_{ct})]L_{ct}^{(\sigma_H-1)/\sigma_H}\})^{\sigma_H/(\sigma_H-1)}. \quad (7)$$

Input K is a composite factor combining college-educated and STEM workers such that

$$K_{ct} = [ST_{ct}^{(\sigma_S-1)/\sigma_S} + H_{ct}^{(\sigma_S-1)/\sigma_S}]^{\sigma_S/(\sigma_S-1)}. \quad (8)$$

The parameter $\sigma_H > 1$ captures the elasticity of substitution between non-college- and college-educated labor. Similarly, $\sigma_S > 1$ is the elasticity of substitution between college-educated and STEM workers.

A long literature has recognized STEM workers as the key inputs in developing and adopting new technologies. Equation (7) captures this by allowing the level of TFP, $A(ST_{ct})^{\sigma_H/(\sigma_H-1)} > 0$, to be an increasing function of the number of STEM workers in a city. It also allows for STEM workers to potentially raise SBP, $\beta(ST_{ct}) \in [0, 1]$. Note that our model assumes that STEM workers are uniquely capable of generating ideas, innovation, and externalities that benefit productivity even if STEM and college-educated workers are close substitutes in production itself (i.e., if $\sigma_S \approx \infty$).

We assume that labor is paid its marginal product and then calculate the total logarithmic (percentage) change in wages for each group in response to a change in the supply of STEM workers. After normalizing the resulting demand conditions by the exogenous change of STEM workers expressed as a percentage of total employment, we derive three linear conditions relating the elasticity of each group's wage and employment to STEM (i.e., the $b_{y,X}$ coefficients estimated from eq. [1]). Remaining parameters in the demand functions (including σ_H, σ_s, and wage and employment shares) come from prior studies, our analysis, or census data. By combining them, we can estimate our values of interest: $\phi_A = (\Delta A/A)/(\Delta ST/E)$, the elasticity of TFP to changes in STEM (relative to initial employment), and $\phi_\beta = (\Delta\beta/\beta)/(\Delta ST/E)$, the analogous elasticity of SBP.[17]

Table 8 displays the simulated TFP (col. 1) and SBP (col. 2) changes from 1990 to 2010. We set $\sigma_S = \infty$ since our regression estimates of $1/\sigma_S$ are never significantly different from zero and the elasticity of college-educated wages and STEM wages to STEM supply are always very close to each other (implying high substitutability). Ciccone and Peri's (2005) review of σ_H estimates suggests a value between 1.5 and 2.5. We assume a σ_H value of 2 in our basic simulation and use values of 1.75 and 2.25 in robustness checks. US census data on wages and employment imply a β value equal to 0.57, a share of STEM workers equal to 0.05 in total employment and 0.09 in the total wage bill, and a college-educated share of the wage bill equal to 0.46. Fernald (2009) measures annual TFP growth equal to 0.89%. Our census calculations measure annual SBP growth equal to 1.75%. Foreign STEM increased by 0.04% of total employment each year.

Values for the elasticity of outcome y for group X to STEM workers come from our regression estimates. The first row of table 8 reports the simulated effects when we use coefficients from the basic specification in table 5, row 1. Row 2 uses the estimates from table 5, row 6, in which we control for imputed unskilled immigrants and reduce the attenuation bias by including only large cities in the regression. We label this row conservative estimates because the underlying regression leads to somewhat smaller estimates of the STEM effect on native wages. Row 3 uses the estimates from table 6, row 1, that adopt total STEM as the explanatory variable. These tend to be 40%–50% smaller than those obtained with foreign STEM.[18] Rows 4 and 5 are the same as row 1 but illustrate the robustness of the simulations to changes in values of the parameter σ_H.

Our simulations imply that foreign STEM growth explained only a modest 5%–8% of SBP growth from 1990 to 2010. In contrast, foreign STEM growth explained between one-third and one-half of the average TFP

[17] Note that we can calculate these effects without specifying the labor supply side of the model as long as we have the table 5 and table 6 equilibrium employment elasticity estimates for each factor.

[18] We also use the elasticity of college-educated wages (3.96) for STEM since the model implies that the elasticity of college-educated wages to foreign STEM cannot be smaller than that of native STEM wages.

Table 8
Simulated Foreign STEM Effects on Yearly Average TFP Growth and SBP Change

	Simulated Foreign STEM Effect on TFP Growth (%) (1)	Simulated Foreign STEM Effect on Skill-Biased Growth (%) (2)	Average US TFP Growth 1990–2010 (%) (3)	Average Change in Skill-Biased Productivity 1990–2010 (%) (4)	TFP Growth Explained by Foreign STEM (Col. 3/Col. 1) (5)	Skill-Biased Growth Explained by Foreign STEM (Col. 4/Col. 2) (6)
1. Basic estimates	.47	.13	.89	1.75	.53	.07
2. Conservative estimates	.41	.08	.89	1.75	.47	.05
3. Based on total STEM	.27	.04	.89	1.75	.30	.04
4. $\sigma_H = 1.75$	.54	.13	.89	1.75	.61	.08
5. $\sigma_H = 2.25$	.43	.12	.89	1.75	.48	.7

NOTE.—The table uses the formulas in the online appendix to calculate the implied elasticity ϕ_A and ϕ_B. We then use the growth of US foreign STEM workers as a share of employment to calculate the implied effects on TFP. The average TFP growth 1990–2010 is taken from Fernald (2009) and the average skill-biased growth is calculated using the average US values for the wages and employment (in hours) of college-educated and non-college-educated workers from the 1990 and 2010 censuses. Unless otherwise noted, the elasticity of substitution between college- and non-college-educated workers is $\sigma_H = 2$. The STEM share of employment is 0.05, the STEM share of wages is 0.09, and the college-educated share of wages is 0.46. These values are calculated from the 2000 US census.

growth during the period. While this result might appear to be very high, it is more plausible when assessed in context with two additional figures. First, foreign labor accounted for about-two thirds of the net growth in STEM workers in our data set. Second, STEM workers are the primary source of sustained economic growth. Jones (2002), for example, argued that 50% of long-run US productivity growth in recent decades is attributable to growth in scientists and engineers as a share of employment. The 33% TFP growth implied by combining Jones's figure with our calculation of the foreign contribution to STEM growth aligns with the simulated results presented in table 8.

In income terms, the average annual TFP effect in table 8, column 1, translates to about 0.47 percentage points per year, implying that native income per capita in 2010 was 9.8% larger than it would have been without the growth contributions from foreign STEM. This would be impossible to justify on the basis of the foreign-born increase in skilled labor supply alone; but when considered as a source of technological innovation, foreign STEM workers may credibly generate large productivity and wage increases. Nonetheless, we concede that our simulated results are based on strong assumptions. In particular, we apply parameters that were estimated across cities to simulate national foreign STEM effects. This will overstate productivity effects if the wage coefficients from the underlying regressions are related to the selection of natives. On the other hand, since our regressions capture only within-city productivity effects and ignore spillovers to other cities, we could also be underestimating national productivity gains.

VII. Conclusions

This article uses the inflow of foreign science, technology, engineering, and mathematics (STEM) workers, made possible by the H-1B visa program, to estimate the impact of STEM workers on the productivity of college- and non-college-educated American workers between 1990 and 2010. The uneven distribution of foreign STEM workers across cities in 1980—a decade before the introduction of the H-1B visa—and the high correlation between the preexisting presence of foreign-born workers and subsequent immigration flows allow us to use the variation in foreign STEM as a supply-driven increase in STEM workers across metropolitan areas.

We find that a 1 percentage point increase in the foreign STEM share of a city's total employment increased the wage growth of native college-educated labor by about 7–8 percentage points and the wage growth of non-college-educated natives by 3–4 percentage points. We find insignificant effects on the employment of those two groups. These results indicate that STEM workers spur economic growth by increasing productivity, especially that of college-educated workers.

References

Acemoglu, Daron. 1998. Why do new technologies complement skills? Directed technical change and wage inequality. *Quarterly Journal of Economics* 113, no. 4:1055–89.

———. 2002. Technical change, inequality and the labor market. *Journal of Economic Literature* 40, no. 1:7–72.

Altonji, Joseph G., and David Card. 1991. The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, ed. John M. Abowd and Richard B. Freeman. Chicago: University of Chicago Press.

Autor, David H. 2010. The polarization of job opportunities in the U.S. labor market: Implications for employment and earnings. Hamilton Project, Washington, DC, April.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. The polarization of the U.S. labor market. *American Economic Review* 96, no. 2:189–94.

Aydemir, Abdurrahman, and George J. Borjas. 2011. Attenuation bias in measuring the wage impact of immigration. *Journal of Labor Economics* 29, no. 1:69–113.

Bartik, Timothy J. 1991. *Who benefits from state and local economic development policies?* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Card, David. 2001. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics* 19, no. 1:22–64.

Ciccone, Antonio, and Giovanni Peri. 2005. Long-run substitutability between more and less educated workers: Evidence from U.S. states 1950–1990. *Review of Economics and Statistics* 87, no. 4:652–63.

Depew, Briggs, Peter Norlander, and Todd A. Sorensen. 2013. Flight of the H-1B: Inter-firm mobility and return migration patterns for skilled guest workers. IZA Discussion Paper no. 7456, Institute for the Study of Labor, Bonn.

Ellison, Glenn, and Edward L. Glaeser. 1999. The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review* 89, no. 2:311–16.

Fernald, John. 2009. A quarterly, utilization-adjusted series on total factor productivity. Unpublished manuscript, Federal Reserve Bank of San Francisco.

Glaeser, Edward L. 2011. *The triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier*. New York: Penguin.

Griliches, Zvi. 1992. The search for R&D spillovers. *Scandinavian Journal of Economics* 94:S29–S47.

Iranzo, Susana, and Giovanni Peri. 2009. Schooling externalities, technology, and productivity: Theory and evidence from U.S. states. *Review of Economics and Statistics* 91, no. 2:420–31.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108, no. 3:577–98.

Jones, Charles I. 1995. Time series tests of endogenous growth models. *Quarterly Journal of Economics* 110, no. 2:495–525.

———. 2002. Sources of U.S. economic growth in a world of ideas. *American Economic Review* 92, no. 1:220–39.

Kato, Takao, and Chad Sparber. 2013. Quotas and quality: The effect of H-1B visa restrictions on the pool of prospective undergraduate students from abroad. *Review of Economics and Statistics* 95, no. 1:109–26.

Katz, Lawrence F., and Kevin M. Murphy. 1992. Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107, no. 1:35–78.

Kerr, William, and William F. Lincoln. 2010. The supply side of innovation: H-1B visa reforms and U.S. ethnic invention. *Journal of Labor Economics* 28, no. 3:473–508.

Lowell, Lindsay. 2000. H-1B temporary workers: Estimating the population. Working Paper no. 12, Center for Comparative Immigration Studies, University of California, San Diego.

Moretti, Enrico. 2004a. Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics* 121, no. 1:175–212.

———. 2004b. Workers' education, spillovers and productivity: Evidence from plant-level production functions. *American Economic Review* 94, no. 3:656–90.

———. 2012. *The new geography of jobs*. New York: Houghton Mifflin Harcourt.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2014. Foreign STEM workers and native wages and employment in US cities. Working Paper no. 20093, National Bureau of Economic Research, Cambridge, MA.

Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. Integrated Public Use Microdata Series: Version 5.0 [machine-readable database]. Minneapolis: University of Minnesota.

Saxenian, Anna Lee. 2002. Silicon Valley's new immigrant high growth entrepreneurs. *Economic Development Quarterly* 16, no. 1:20–31.

US Citizenship and Immigration Services. Various years. *Characteristics of specialty occupation workers (H-1B) annual reports*. http://www.uscis.gov/tools/reports-studies/reports-and-studies.

US Department of Labor. National O*NET Consortium. 2012. O*NET 17.0 database. http://www.onetcenter.org/db_releases.html. Accessed July 8, 2013.

US Department of State. 2012. Nonimmigrant visa issuances by visa class and by nationality: FY1997–2011 NIV detail table. http://travel.state.gov/content/visas/english/law-and-policy/statistics/non-immigrant-visas.html. Accessed August 23, 2012.