2017
OKSWP1714

Economics Working Paper Series
Department of Economics

OKLAHOMA STATE UNIVERSITY
http://spears.okstate.edu/ecls/

Do Native STEM Graduates Increase Innovation? Evidence from U.S. Metropolitan Areas

John V. Winters

Oklahoma State University, GLO, and IZA

Department of Economics
Oklahoma State University
Stillwater, Oklahoma

339 BUS, Stillwater, OK 74078, Ph 405-744-5110, Fax 405-744-5180

Do Native STEM Graduates Increase Innovation? Evidence from U.S. Metropolitan Areas

John V. Winters
Oklahoma State University, GLO, and IZA
jvwinte@okstate.edu

August 2017

Abstract

This paper examines the effects of college graduates educated in STEM fields on patenting intensity in U.S. metropolitan areas. Some prior research suggests a positive effect on urban innovation from foreign-born STEM workers, but little is known about the effects of native STEM graduates on innovation. My preferred results use time-differenced 2SLS regressions, and I introduce a novel approach to instrumenting for the growth in native STEM graduates. I find positive effects of foreign STEM on innovation, roughly consistent with previous literature. However, my preferred approach yields a negative coefficient estimate for native STEM graduates on innovation that is not statistically significant but suggests that a meaningfully large positive effect is unlikely during the 2009-2015 time-period. I discuss possible explanations and implications.

JEL Codes: I25; J24; J61; O31; R12
Keywords: STEM; innovation; patents; human capital; higher education

Address mail correspondence to: Oklahoma State University, 331 Business Building, Stillwater, OK 74078-4011.

The author thanks David Neumark, two anonymous reviewers, Rachana Bhatt, Barry Hirsh, Tom Holmes, Tom Mroz, Jordan Rappaport, Dan Rickman, Dave Sjoquist, seminar participants at the Federal Reserve Bank of Kansas City, Georgia State University, and University of Oklahoma, and session participants at the North American Meetings of the Regional Science Association International, the Southern Economic Association meetings, the Southern Regional Science Association meetings, and Western Economic Association International meetings for helpful comments on earlier drafts. Any errors are the responsibility of the author. An earlier version of the paper with different data, methods, and conclusions was previously distributed as IZA Discussion Paper No. 8575 with the title "Foreign and Native-Born STEM Graduates and Innovation Intensity in the United States." The current paper supersedes the prior one.

1. Introduction

Technological innovation is critical for economic growth and development, and many nations and regions are very interested in how they can increase innovation. Skilled labor is widely perceived as an important ingredient in the innovation process (Carlino et al. 2007).[1] College graduates in science, technology, engineering, and math (STEM) fields are believed to be especially important (Atkinson and Mayo 2010). However, many employers and policymakers in advanced economies claim that they are experiencing a shortage of skilled workers, especially in STEM fields (National Academies 2010). Many researchers, policymakers, and businesses, therefore, support various public policies to increase the stock of skilled labor, especially the stock of STEM graduates (Moretti 2013). These polices, however, are not without controversy; some question their benefits. Specifically, there has been relatively little research examining whether and how much various types of skilled labor actually affect innovation. Two previous studies (Hunt and Gauthier-Loiselle 2010; Kerr and Lincoln 2010) suggest a positive effect of foreign STEM workers on innovation in the U.S., but very little is known about the effects of native STEM graduates on innovation.

The current paper fills an important gap in the research literature by examining the effects of native- and foreign-born STEM graduates on patent intensity in U.S. metropolitan areas. Native and foreign STEM graduates differ in important ways, and it is unclear if both increase innovation and by how much. Foreign STEM graduates in the U.S. are likely the best and brightest from their home countries and may have technical skills that exceed those of native STEM graduates. However, natives may possess additional knowledge and skills related to

[1] The local stock of high skilled workers in an area has been shown to increase wages and employment rates for both high and low skilled persons (Moretti 2004; Winters 2013) and increase future employment and population growth (Simon 1998; Simon and Nardinelli 2002; Shapiro 2006). This is consistent with increased innovation but could also result from other possible mechanisms. Furthermore, some are skeptical that the observed effects are causal.

language, culture, and institutions that help facilitate their contributions to urban innovation. It is also unclear if and how college graduates in non-STEM fields affect innovation, especially relative to non-graduates. Individual-level survey data for college graduates reveal that non-STEM graduates do indeed patent innovations at non-trivial rates (Hunt and Gauthier-Loiselle 2010), but they do so at lower rates than STEM graduates.[2]

I examine the effects of various types of college graduates on metropolitan area innovation by examining local changes in patent intensity and human capital levels between 2009 and 2015. My preferred approach uses instrumental variables (IV) methods to estimate causal effects. Specifically, I estimate time-differenced two-stage least squares (2SLS) regressions. I follow previous literature and identify the effects of foreign-born STEM graduates using a "shift-share" instrument that combines historical immigrant location decisions with recent national in-flow rates by country of origin to predict skilled immigrant inflows for each MSA during 2009-2015. However, my primary contribution is to use an instrumental variable strategy to estimate causal effects of native STEM graduates on innovation. I identify the 2009-2015 change in native STEM graduates using a novel instrument based on the age-gender-ethnicity distribution of children ages 0-5 in 1990 in each MSA combined with national STEM education rates by age-gender-ethnicity for those children during early adulthood. To my knowledge, this is the first study to use a time-differenced IV strategy to examine effects of skilled natives on MSA innovation.

[2] Even if they do not directly innovate much themselves, non-STEM graduates may provide important intermediate inputs, e.g. specialized financial, legal, marketing, and managerial services, and thereby increase regional innovation levels. Non-STEM fields account for roughly three-fourths of all college graduates in the U.S. (based on author's estimates using the American Community Survey), so better understanding any effect that they have on innovation is worthwhile.

Previewing the results, I find that foreign STEM graduates increase innovation consistent with expectations based on previous literature, though my use of newer data and specific focus on STEM graduates makes this an important contribution. My results also suggest that non-STEM graduates have a minimal effect on urban innovation, which is a new result in the literature. However, my most important result is for native STEM graduates. I find no positive effect of native STEM graduates on MSA innovation levels. The preferred coefficient estimate is negative, though not statistically significant.

My results suggest that adding young American STEM graduates to a metropolitan area does not increase patenting in the current environment. The 2SLS confidence interval is relatively wide, so I cannot rule out modest positive effects, but the preferred estimates suggest that there is unlikely to be a sizable positive effect of native STEM graduates on metropolitan area patent intensity during the 2009-2015 time period. This is a somewhat surprising result that raises important concerns for researchers and policymakers. Numerous stakeholders have called for more STEM graduates in the hopes of increasing innovation and growing productivity. The current study suggests that simply educating more young natives in STEM fields is not likely to significantly increase innovation. This should be a call for researchers and policymakers to devote more resources to better understanding the causes of innovation and the potential role that STEM education can play. While I can only speculate, I discuss below some possible explanations and interpretations for the lack of a positive effect of native STEM graduates on patent intensity. Additionally, the differing results for foreign and native STEM graduates underscores the importance of skilled immigration for American innovation. America as a whole greatly benefits from the numerous contributions of highly skilled workers who come to the U.S. from other countries.

2. Previous Literature

The stock of college graduates is thought to be very important for regional and national economic growth, and the stock of STEM graduates is often thought to be especially important. The stock of graduates in an area at a given time depends on several factors. One way to increase the stock of skilled labor is by increasing human capital levels of current residents, and there are numerous policy efforts intended to do so. Investments in higher education are costly to both individuals and society, and STEM graduates are among the most expensive to educate (Nelson 2008; Altonji and Zimmerman 2017). There is some evidence that college major decisions can be affected by tuition and financial aid policies (Sjoquist and Winters 2015a, 2015b; Stange 2015; Denning and Turley 2017). However, there is also evidence that many American college students start out pursuing a STEM major but end up switching to less challenging majors because they lack sufficient preparation in math and science skills (Griffith 2010; Arcidiacono et al. 2012; Stinebrickner and Stinebrickner 2014; Arcidiacono et al. 2016). To my knowledge, no previous study has estimated causal effects of native STEM graduates on metropolitan area patenting.[3] This is an important contribution for the current study.

Areas can also increase human capital stocks via in-migration of persons who acquired needed skills elsewhere. This option has led to considerable debate on high-skilled immigration policy in advanced economies such as the United States.[4] However, there is still much that is

[3] An earlier version of the current paper with a different title (Winters 2014b) attempted to do so indirectly using earlier data, but the IV approach in that paper has poor first-stage properties when using the newer more direct measures of STEM shares and the more detailed control variables in the current study. The current study supersedes the prior one.

[4] Much of the debate in the U.S. is centered on the H-1B visa program, which allows employers to apply for temporary visas for skilled foreigners working in specialty occupations. The program began in 1990 and the annual quota has varied considerably since inception. Various stakeholders argue that the quota should be increased, decreased or even reduced to zero. Kerr and Lincoln (2010) provide additional details. See Facchini and Steinhardt (2011) for a broader discussion of U.S. immigration policy.

unknown about how high-skilled immigration affects receiving areas, and many worry that skilled immigrants may adversely affect similarly skilled natives.[5] Surveying the research literature, Kerr (2013) concludes that “the global migration of talented workers … is vastly understudied compared to its economic importance.”

There is only a small literature that directly examines the effects of skilled foreigners on innovation.[6] Hunt and Gauthier-Loiselle (2010) and Kerr and Lincoln (2010) examine the effects of immigrants on innovation by looking at patent data.[7] Hunt and Gauthier-Loiselle (2010) first examine the 2003 National Survey of College Graduates (NSCG) to assess individual determinants of patenting. They find that the average foreign college graduate patents at double the rate of the average native graduate but indicate that this effect largely results

---

[5] Traditional models of supply and demand suggest that increasing the supply of skilled labor through in-migration will lower wages for similar natives, and there is some empirical evidence to support this contention (Borjas 2003, 2006). There is also some evidence that increased immigration partially “crowds out” natives in areas receiving large immigrant inflows by encouraging them to migrate to areas receiving smaller immigrant shocks (Borjas 2006; Ali et al. 2012) and by encouraging them to work in occupations less affected by immigrant labor supply shocks (Levin et al. 2004; Peri and Sparber 2011; Orrenius and Zavodny 2015; Ransom and Winters 2016). However, researchers have also suggested that foreign and native workers may experience considerable complementarities, and some have found wage effects of immigrants on natives to be small, zero, or even positive (Peri and Sparber 2009; Ottaviano and Peri 2012; Peri et al. 2015). Kerr (2013) and Lewis and Peri (2014) review previous literature.

[6] A related literature looks at how foreigners compare to natives in various measures of innovation. Much of this literature has examined differences in academic achievements between native and foreign born faculty and graduate students and found mixed results. Levin and Stephan (1999), Stephan and Levin (2001), Corley and Sabharwal (2007), Chellaraj et al. (2008), and Gaulé and Piacentini (2013) find that foreign born academics outperform their native counterparts. However, Stuen et al. (2012) find that foreign and native doctoral students have statistically comparable effects on academic innovation in science and engineering departments at American universities. Gurmu et al. (2010) find that the relative contributions of natives and foreigners to academic innovation vary between graduate students and postdoctoral scholars and also depend on the temporary or permanent visa status of foreigners.

[7] Akcigit et al. (2017a, b) document an important role for immigrant inventors in the U.S. during 1880-1940, which they call the golden age of U.S. Innovation. Waldinger (2012) and Moser et al. (2014) examine the effects of a specific historical immigration shock, Jewish émigrés from Nazi Germany; both studies find no evidence of knowledge spillovers from émigrés to prior residents. Moser et al. (2014), however, do find a large positive effect on U.S. chemical innovation due to more researchers working in those fields. Related studies have also considered the effects of various types of skilled workers, including skilled immigrants, on regional innovation in other countries, especially in Europe (e.g., Simonen and McCann 2008; Faggian and McCann 2009; Niebuhr 2010; Nathan and Lee 2013; Ozgen et al. 2013; Lee 2014; Maré et al. 2014; Nathan 2014).

because foreigners are more likely to have earned degrees in science and engineering fields.[8] Conditional on earning a degree in science or engineering, foreign and native graduates patent at rates that are statistically similar. However, examining individual self-reported data on patenting has some potential limitations: it ignores potential crowd out effects, knowledge spillovers, complementarities between different types of workers, differences in collaboration patterns between natives and foreigners, and the possibility that foreign graduates disproportionately locate in areas that make them more likely to patent.

Hunt and Gauthier-Loiselle (2010) also examine the effects of skilled foreigners on regional innovation using 1940-2000 state-level panel data on patents per capita and the stocks of foreign college graduates per capita measured at ten year increments. Some of their specifications control for the stock of native college graduates in the state, but their analysis treats this variable as exogenous and does not differentiate based on field of study. Their preferred specifications instrument for decadal growth in the skilled immigrant population share using the "shift-share" predicted growth based on state immigrant shares for various origin countries in 1940 and the national growth in the immigrant population from those countries during the decade in question. They find that foreign graduates increase state patent intensity, and the estimated coefficients imply considerable spillovers relative to the effects predicted by individual-level data.

Kerr and Lincoln (2010) exploit the H-1B visa program to identify large annual changes in skilled foreigner inflows across 281 metropolitan areas for the years 1995-2007. They estimate reduced form regressions of the effect of predicted flows of H-1B visa holders on patent

[8] Hunt (2011) also uses the NSCG to examine the effects of immigrants on innovation by entry visa type. She finds that immigrants who were initially admitted as legal permanent residents (such as through family unification) have similar patenting outcomes as natives.

intensity. Examining annual changes makes their analysis primarily short-run in nature, and they do not examine the effect of native skill levels on patent intensity. They find that increased predicted H-1B immigrant inflows significantly increases local patenting. They also match patents to ethnic surnames and find that much of the increase is attributable to Indian and Chinese surnames. However, they do find some evidence of increased patenting for Anglo-Saxon surnames due to H-1B inflows, which may suggest positive innovation spillovers from foreigners to natives, i.e., natives may be crowded into innovation instead of crowded out.

Carlino et al. (2007) also examine determinants of urban patenting and consider the particular role played by the density of professional and specialty occupation employment. They find generally positive effects but evidence of some non-linearity with a positive linear term and negative quadratic term. Professional and specialty occupations is a relatively broad category that includes scientists and engineers as well as managerial, financial, medical, and educational professionals. Carlino et al. (2007) also separately examine the employment density of scientists and engineers and again find a significant positive linear term and a significant negative quadratic term. Carlino et al. (2007) do not differentiate between native and foreign-born workers.[9]

The current study also differs from previous literature by focusing on the effects of STEM *graduates* on innovation. The limited previous literature has focused on the effects of STEM *occupations*[10] rather than STEM *graduate*s, largely because measures of local stocks of STEM graduates have not been available until recently. STEM graduates and STEM

---

[9] Carlino and Hunt (2009) conduct a follow up study to Carlino et al. (2007) that among other things incorporates patent citations as an additional outcome measure. Adjusting for patent quality via citations yields generally similar results as unweighted patent counts. Both studies find large significant effects of local human capital levels as measured by the share of college graduates, but neither distinguish among the college majors of college graduates or between natives and immigrants.

[10] Researchers have also explored related occupational groupings such as science and engineering, technical, etc.

occupations are closely related, but there are differences. Some individuals who report working in a STEM occupation do not have a STEM degree and some have no college degree at all.[11] Similarly, many STEM graduates end up working in non-STEM occupations such as business, management, healthcare, and education. Furthermore, innovation is not always directly related to what one does at work. Many innovations come from persons developing them at home in their spare time; some of these inventors have day jobs that are not in STEM occupations and others may be unemployed or not in the labor force. While there are some differences between STEM occupations and STEM graduates, trying to separate the effects of the two is not the focus of the current study. Instead, I try to identify exogenous increases in the stocks of native and foreign STEM graduates and examine their effects on innovation.[12]

Measuring human capital based on college major allows one to examine separate effects of STEM graduates and non-STEM graduates on innovation, which more closely aligns with public policies than does looking at occupations. Many researchers and policymakers advocate for increased STEM education based on the expectation that STEM graduates contribute greater benefits to society than non-STEM graduates (National Academies 2010; PCAST 2012). Consistent with that notion, Winters (2014a) considers the magnitude of human capital wage externalities from STEM and non-STEM graduates and finds that STEM graduates have a much larger positive external relationship with the wages of non-college graduates in the same metropolitan area.[13] Thus, the general expectation is that STEM graduates will have a greater

[11] This may in part result from ambiguous titles of occupations. For example, some workers may refer to themselves as engineers but have completed no higher education and perform duties such as operating machinery and equipment that might incline an outside observer to view them as a technician rather than an engineer.
[12] My study also differs from Hunt and Gauthier-Loiselle (2010) in the time period considered and the geographic unit of analysis.
[13] The analysis in Winters (2014a) relies on OLS and, therefore, does not necessarily measure causal effects. Furthermore, in spatial equilibrium wages are jointly determined with housing prices and area amenities. Patents are a more straightforward measure and provide evidence on an important mechanism by which various types of human

effect on regional innovation than non-STEM graduates, but it is ultimately an empirical question. Previous researchers have been unable to offer empirical evidence on this issue because reliable estimates of geographic differences in the densities of STEM and non-STEM graduates across the U.S. have not been available until recently. To my knowledge, this is the first study to exploit recent information in the American Community Survey to assess the importance of STEM and non-STEM major college graduates on sub-national innovation.

Furthermore, the current paper differentiates between effects on innovation from native STEM graduates and foreign STEM graduates. While these effects could be similar, there is plenty of reason to expect native and foreign STEM graduates could have differing effects on innovation. Foreign-born STEM graduates are a highly selected group, and agglomeration economies in innovation-leading countries like the U.S. may attract some of the brightest and most innovative minds in the world. However, skilled immigrants also face a number of barriers that natives do not, and these lower barriers for native STEM graduates may give them an advantage in innovation. Of course, native STEM graduates may also have differing preferences, experiences, and backgrounds that could make them less innovative than skilled foreigners. Ultimately, this is an empirical question, and the lack of understanding about how native graduates affect innovation is an important gap in the literature that I try to fill.

3. Empirical Methods

3.1 Regression Overview

capital might affect regional economic development and well-being. See also Liu (2016) for related evidence on human capital wage externalities by college major.

This paper examines the effects of native and foreign STEM graduates and non-STEM graduates on patent intensity in U.S. metropolitan areas. I begin by estimating descriptive cross-sectional ordinary least squares (OLS) regressions for year 2015 of the form:

$$Patenting_c = \gamma HumanCapital_c + \alpha X_c + \varepsilon_c \qquad (1)$$

, where $HumanCapital_c$ includes measures of human capital for native and foreign STEM graduates and non-STEM graduates in metropolitan area c, X_c includes a number of control variables, and ε_c is an error term. However, some metropolitan areas may have local resources that make them permanently more innovative and cause cross-sectional results to be inaccurate.

I next estimate time-differenced OLS and two-stage least squares (2SLS) regressions of the form:

$$\Delta Patenting_c = \theta \Delta HumanCapital_c + \beta X_c + \mu_c \qquad (2)$$

, where Δ indicates within metropolitan area time differences taken between years 2009 and 2015 for both the dependent variable and the human capital variables. The time-differencing removes time-invariant factors that might influence the dependent variable and also be correlated with contemporaneous human capital variables in equation (1). Notice, however, that the control variables in equation (2) are not time-differenced between 2009 and 2015 because their changes after 2009 are potentially endogenous to the change in patenting. Instead, I use control variables measured in 2009 or earlier years and use the same control variables for equations (1) and (2). Further details are discussed below in sub-section 3.2.

The time-differenced 2SLS results are my preferred estimates, and I discuss my 2SLS identification strategy below in sub-sections 3.3 and 3.4. I also later examine the robustness of the 2SLS results to several alternative specifications, and I explore STEM occupation employment as an alternative outcome and possible mechanism for the estimated effects on patenting. For all regressions in this paper, standard errors are clustered by state to account for possible error correlation across MSAs within the same state.[14]

3.2 Data

Data on native and foreign STEM and non-STEM graduates come from the 2009 and 2015 American Community Survey (ACS) accessed from IPUMS (Ruggles et al. 2015). The ACS began asking college graduates to report their major field of study for their bachelor's degree in 2009. Before 2009 reliable measures of local stocks of STEM and non-STEM graduates were not available. The 2015 ACS was the most recent ACS microdata sample available at the time of the analysis. My primary analysis looks at time differences during 2009-2015, the longest period possible given the data. The relative stock of native (foreign) STEM graduates in each metropolitan area is measured as the number of native-born (foreign) college graduates ages 25 and up with a bachelor's degree in a STEM field divided by the total adult (ages 25 and up) population. The stock of non-STEM graduates is computed as the number of adult college graduates with bachelor's degrees in non-STEM fields relative to the adult population.[15] I define college majors as STEM fields based on definitions used by U.S. Immigration and Customs Enforcement (ICE); a full list of STEM majors and corresponding ACS codes is provided in Appendix Table A.

U.S. Census Bureau confidentiality restrictions for microdata samples from the decennial census and ACS prevent identification of geographic areas with less than 100,000 people. As a

[14] Errors could also be potentially correlated across states within the same Census region or division, but clustered standard errors have poor small sample properties and perform badly when the number of clusters is small (Bertrand et al. 2004), so clustering by region or division is not appropriate.
[15] I do not differentiate between foreign and native non-STEM graduates for various reasons. First, foreigners make up a much smaller percentage of non-STEM graduates than STEM graduates. Second, STEM fields are expected to have a stronger effect on innovation, so I focus attention on STEM graduates.

result, the lowest level of geography in the census microdata, PUMAs, often combine parts of metropolitan areas with parts of other nearby metropolitan areas or non-metropolitan areas. I use the PUMA component files to assign each PUMA to a metropolitan area if the majority of the PUMA population is included in the metropolitan area; other PUMAs, including wholly non-metropolitan ones are excluded from the analysis. I use the year 1999 metropolitan statistical area (MSA) boundaries (with county-based definitions for New England) to measure all of the variables included in this study.[16] PUMA boundaries change over time due to the Census Bureau redrawing boundaries about every 10 years. The 2009 ACS PUMA boundaries are defined based on the 2000 census data, while the 2015 ACS PUMAs are based on the 2010 census. The instrumental variables that I use rely on the 1980 and 1990 decennial census microdata files, which defined PUMAs (county groups in 1980) based on even earlier boundaries. As a result, some small metropolitan areas cannot be linked over time and are by necessity excluded from the analysis.[17] The analytical sample includes 288 metropolitan areas (out of 318) based on 1999 definitions.

Patent intensity is measured as the log of patents per 100,000 population. Patent data are obtained from the U.S. Patent and Trademark Office (PTO), and population data are obtained from the U.S. Bureau of the Census. The PTO reports the origin location for each patent based on the residence of the first-named inventor. I merge county-level patent totals with 1999 MSA definitions to compute patent totals by MSA. I then compute log patents per 100,000 population for each MSA. For the cross-section analysis, the dependent variable is for a single year of data,

[16] I use primary metropolitan statistical area (PMSA) definitions instead of consolidated metropolitan statistical area (CMSA) definitions for large MSAs.
[17] I also exclude these from the OLS analysis to facilitate comparability with the 2SLS results.

2015.[18] For the 2009-2015 time-differenced regressions, the dependent variable is computed as the difference in log patents per 100,000 population between the 2009 and 2015 single-year values.[19] There could be some lumpiness in patent timing, so I also discuss robustness checks below for the time-differenced 2SLS regressions that use patent differences for alternative years. Figure 1 illustrates the bivariate relationship across MSAs between the 2015 STEM graduate share and log patents per 100,000 population. The figure illustrates a strong positive bivariate descriptive relationship between patent intensity and STEM graduates, but more rigorous multivariate regression analyses will be conducted to better understand the relationship.

The spatial distributions of the two patenting dependent variables are illustrated via MSA maps in Figures 2 and 3 with MSAs grouped by patent quartile for each.[20] Figure 2 maps patent intensity for the 2015 cross-section and visually suggests that the most innovative areas are somewhat concentrated among coastal states, big cities, and areas with high levels of human capital; this is consistent with expectations. Figure 3 maps the growth in patent intensity during 2009-2015. Some areas have high or low patenting for both 2015 levels and the 2009-2015 growth illustrated in these two figures. However, many areas fall in different quartiles of these two figures. Furthermore, the correlation coefficient between these two dependent variables is very weak at 0.04. Thus, there are considerable differences between the two dependent variables and patenting growth is not strongly related to initial levels.

[18] One small MSA, Cumberland, MD-WV MSA had zero patents in 2015, so I replace its log patents per 100K as the average over 2014-2015 to retain it in the sample since log of zero is undefined. All other MSAs in the sample had positive patents in 2015.
[19] The patent total in 2009 is zero for Laredo, TX MSA. For both Cumberland MSA and Laredo MSA, I define the 2009-2015 time-difference based on two-year averages for 2009-2010 and 2014-2015. Results are qualitatively robust to excluding these two small MSAs.
[20] Maps were created in Stata using the maptile program written by Michael Stepner. See https://michaelstepner.com/maptile/

The control variables in the regression analysis were included based on consulting previous literature and economic theory.[21] The control variables include several time-varying metropolitan area characteristics measured as of 2009 including the log of the population, the industry employment structure, the unemployment rate, the mean age of the adult labor force, the average firm size, the number of research universities per 100,000 population, and university research expenditures per 100,000 population. I also include eight census division dummies (with the New England division being the omitted group), and the incremental distance to the nearest metropolitan area with a population of at least 250,000, 500,000, and 1,500,000. Finally, I also control for historical innovation levels across MSAs by including three variables for average log patents per 100K population during 2007-2008, 1998-2006, and 1990-1997. I also discuss robustness checks below that include some additional variables as controls.

A large literature following Jaffe et al. (1993) has suggested that knowledge spillovers decline with distance. Larger metropolitan areas are likely to experience greater knowledge spillovers, so log population is included to account for the effects of city size on innovation (Carlino et al. 2007; Carlino and Kerr 2014). However, agglomeration economies might spill across the urban hierarchy as suggested by Partridge et al. (2009, 2010), so I also control for proximity to progressively larger metropolitan areas similarly to Partridge et al. (2009, 2010).[22] Industries differ in innovation activities including patenting, and local industrial structure can affect MSA patenting intensity (Carlino et al. 2007), so I use the 2009 Quarterly Census of

[21] If one uses exogenous instruments in 2SLS regressions, then including control variables is not strictly necessary, and much of the quasi-experimental literature on urban/regional innovation (e.g. Hunt and Gauthier-Loiselle 2010; Kerr and Lincoln 2010; Peri et al. 2015) uses very parsimonious specifications. However, including exogenous controls can improve efficiency by explaining away some of the error variance and can provide a partial check against concerns that the instruments may be endogenous.

[22] Proximity computation requires population centers for MSAs, which can be constructed from county population centers for either 2000 or 2010 decennial census data. I construct all control variables to be observed in 2009 or earlier, so the proximity variables are measured using year 2000 data.

Employment and Wage (QCEW) to control for MSA employment shares in federal government, state and local government, natural resources and mining, construction, manufacturing, transportation and utilities, information, financial activities, professional and business services, and education and health services.

The census division dummies are included to account for broad differences in innovative activity. The unemployment rate measures local labor market conditions and labor utilization, and the mean age of the workforce proxies for worker experience; both are computed using the 2009 ACS. Metropolitan areas with smaller average firm sizes are expected to be more entrepreneurial and experience greater innovation and growth (Glaeser et al. 2010; Chatterji et al. 2014; Glaeser et al. 2015). Average firm size is calculated from the Business Dynamics Statistics (BDS) Data Tables produced by the U.S. Census Bureau. University research is expected to increase local innovative activity (Jaffe 1989; Anselin et al. 1997; Adams 2002; Ponds et al. 2010; Kantor and Whalley 2014) and is also potentially correlated with the primary human capital variables. Therefore, university research expenditures are obtained from the 2009 Integrated Postsecondary Education Data System (IPEDS) and used to compute university research expenditures per 100,000 population in the MSA to be included as a control variable. Given the potential noisiness in university research expenditures, I also control for the number of Carnegie research universities per 100K population in the MSA.[23]

Finally, I control for average MSA patenting levels during 2007-2008, 1998-2006, and 1990-1997 to account for possible omitted variables and in recognition of possible persistence in MSA innovation growth. For example, Combes et al. (2008) suggest that observed urban wage premiums in France partially reflect sorting on both observed skills and on unobserved skills.

[23] Results are also qualitatively robust to including a multi-year average of university research expenditures as a control variable.

Similarly, one might expect more innovative areas to have workers with higher average levels of both observable and unobserved skills. Controlling for prior innovation levels will help account for prior sorting on unobservables.[24] I chose the specific periods for these based on national trends in patent intensity. Figure 4 illustrates mean patents per 100,000 population for the 288 MSAs included in this study for each year from 1990 to 2015. Mean patenting increased steadily but moderately during 1990-1997 and then experienced a large jump in 1998. Mean patenting roughly plateaued during 1998-2003 before dipping in 2004-2005 and then returning to plateau levels in 2006. Mean patenting fell in 2007 and remained low during 2008-2009. Mean patenting increased in 2010 to well above the plateau levels of 1998-2003 and then dipped slightly during 2011. Mean patenting increased during 2012, 2013, and 2014 before a slight dip in 2015. Figure 4 helps in the construction of the past patenting control variables and also indicates that the 2009-2015 period covering the dependent variable was a period of considerable growth in patenting.[25]

Summary statistics for the main variables in the analysis are included in Table 1. A few things are particularly noteworthy. First, MSA averages for patenting, native and foreign STEM occupation employment, native and foreign STEM graduate shares, and non-STEM graduate shares all increased during 2009-2015. However, the relative growth in STEM education and occupations has been more pronounced for foreigners than natives. For example, the average MSA saw its share of adult workers who are foreign STEM graduates increase from 2.0 percent

---

[24] Another possibility might be to include control variables for the presence of star scientists, which have been found to have important impacts on innovation by Zucker and Darby (1996, 2007). However, I do not attempt to control for star scientists because their location decisions are likely endogenous and perhaps in ways correlated with my human capital variables. For example, if some areas become more innovative, entrepreneurial star scientists may be especially drawn to them. Furthermore, exogenous increases in human capital useful for innovation may especially attract star scientists because of the access to complementary inputs that improve their ability to innovate.

[25] I also experimented with numerous alternative specifications including altering the dependent variable to have a different start period and breaking the 1998-2006 patenting control variable into two sub-periods. Results are qualitatively similar for a number of alternatives.

to 2.5 percent during 2009-2015. The increase for native STEM graduates was also 0.5 percent from 5.1 percent to 5.6 percent. Natives still account for the majority of STEM graduates in the U.S., but the overall growth in STEM graduates is heavily fueled by foreign-born STEM graduates. During the same time, the employment shares for STEM occupations grew more modestly, and the average growth was larger for foreign STEM workers than native STEM workers. In particular, the employment of foreign STEM occupation workers increased from 0.86 percent to 1.03 percent of the workforce for the average MSA during 2009-2015. The corresponding change for natives was from 2.51 percent to 2.60 percent of the average MSA workforce. Thus, natives also account for the bulk of STEM occupations, but foreign-born STEM employment is growing faster both relatively and absolutely.

An additional point to note is that most STEM graduates do not work in STEM occupations based on the relatively narrow STEM occupation definition used in this study. Even if we broaden the STEM occupation definition to include medical professionals and a few other STEM-related occupations, there are still a very large number of STEM graduates in non-STEM occupations including management, education, sales, etc. Many STEM graduates choose to work in non-STEM occupations because it matches their current interests and skill sets, and a STEM education might increase productivity in a number of non-STEM occupations. However, there is also some concern among researchers and policymakers that there may not be enough STEM jobs for all the STEM graduates wanting them (Wright et al. 2017). After conducting the primary analysis on patenting, I will also turn toward examining whether an increase in the share of STEM graduates in an MSA increases the share of STEM occupation employment in the MSA.

The summary statistics also indicate that in 2015 non-STEM graduates are generally a substantially larger share of the adult population than are STEM graduates; the mean share for the former is more than three times that of the latter. In fact, no metropolitan area in the U.S. has more STEM graduates than non-STEM graduates in 2015. Metropolitan areas also differ substantially in their concentration of STEM majors and in their relative dependence on foreign STEM majors. Table 2 reports the 2015 population shares for the four college graduate variables for the 25 metropolitan areas with the highest shares of STEM graduates. Not surprisingly, San Jose, CA tops the list with an impressive 24.3 percent of the adult population with a STEM degree. San Jose also has the highest foreign STEM graduate population share in the nation and has nearly twice as many foreign STEM graduates as native STEM graduates.[26] Boulder-Longmont, CO has the second highest total STEM graduate share and the highest native STEM graduate share. Furthermore, Boulder-Longmont has more than five times as many native STEM graduates as foreign ones. Table 1 also reports the minimum and maximum values for the human capital shares in 2015 and their changes during 2009-2015. Tables 1 and 2 together confirm that there is considerable variation in STEM graduate stocks and in relative dependence on domestic and foreign STEM graduates across the country.

Table 3 reports correlations for the human capital variables for 2015 and for the 2009-2015 time differences. The 2015 cross-sectional correlations are positive and relatively strong, indicating that native and foreign STEM graduates and non-STEM graduates are drawn to relatively similar areas. However, the cross-section correlations are not overly strong, so there should be plenty of variation in the data to avoid collinearity problems. The correlations for the

[26] I discuss sensitivity analysis below that excludes outliers such as San Jose.

2009-2015 time-differenced human capital variables are all relatively modest, indicating that the recent MSA growth in each of these variables is largely independent of the others.

OLS estimates may not provide unbiased estimates of causal effects. For example, reverse causality may exist if STEM graduates sort into innovative areas, and omitted variable bias may exist if both innovation and STEM graduate stocks are driven by some unobservable characteristic even after the inclusion of the metropolitan area control variables. Additionally, the ACS human capital variables are measured using a one percent sample of the population, which will lead to some degree of measurement error due to sampling, especially for relatively small areas. Measurement error due to sampling may attenuate coefficients toward zero, and this attenuation may be exacerbated by including multiple related measures and a detailed set of control variables that reduce the signal-to-noise ratio.

3.3 Native STEM Instrument

The preferred estimates in this study utilize instrumental variables to estimate time-differenced 2SLS regressions. While prior research has used instrumental variables to examine effects of skilled foreigners on innovation, to my knowledge no other researchers have used a time-differenced instrumental variables identification strategy to estimate effects of skilled natives on innovation. This is the fundamental contribution of the current paper. I identify the effects of native STEM graduates on innovation using an instrument that combines microdata from the 1990 census 5% sample with microdata from the 2014 ACS. My native STEM instrument is similar in spirit to shift-share instruments for foreign workers used in previous literature and somewhat related to an age-structure based instrument used by Moretti (2004) to estimate external wage effects of local human capital levels. Like these previous studies, my instrument exploits supply-driven increases in local human capital levels. However, my native STEM instrument has several unique attributes and is novel in the literature. In particular, my native STEM instrument is based on STEM education rates of young natives who reach age 25 after 2009.

I use the 1990 census microdata to construct a sample of all children born in the U.S. who were ages 0-5 at the time of the 1990 census. I then define these children into 25 ethnic origin groups based on variables for race, ethnicity, and ancestry; the 25 groups are listed in Appendix Table A2.[27] Combining two genders, six ages, and 25 ethnic groups gives 300 unique groups. For each MSA, c, and each group, g, I compute the population in 1990, $Pop1990_{cg}$. I also compute the population share for each MSA-group in 1990 relative to the total population in the MSA of all natives ages 0-5, i.e.,

$$PopShare1990_{cg} = \frac{Pop1990_{cg}}{\sum_g Pop1990_{cg}} .$$

I next use the 2014 ACS to construct a sample of native-born young adults ages 24-29 in 2014; these young people were ages 0-5 in 1990.[28] I define these young adults into 300 groups by age-sex-ethnicity as done for 1990 in order to link the 1990 and 2014 data. I then use the 2014 college major information to classify each individual as either having a STEM degree or not; persons with no bachelor's degree are included in the category without a STEM degree. For each of the 300 groups, I then compute the 2014 national share of each group with a STEM degree as the ratio of the number of STEM graduates to the total population of the group:

[27] More specifically, I use IPUMS variables racesingd, hispan, and ancestr1. I use the hispan variable to classify Hispanics into four groups, the ancestr1 variable to classify white non-Hispanics into 12 groups, and the racesingd variable to classify all others into nine additional groups. The racesingd variable is an IPUMS constructed variable that classifies all persons, including multi-racial persons, into a time-consistent coding scheme based on their single primary race.

[28] At the time of the analysis in this paper, the racesingd variable was not yet available for 2015, preventing me from using the 2015 ACS for the native STEM instrument.

$$NationalSTEMShare2014_g = \frac{\#of\ STEM\ Graduates\ in\ 2014_g}{Pop2014_g}.$$

The next step is to combine $PopShare1990_{cg}$ with $NationalSTEMShare2014_g$ to form the native STEM instrument. I define the predicted STEM education rate among young natives in each MSA as:

$$Young\widehat{Native}STEM_c = \sum_g PopShare1990_{cg} \times NationalSTEMShare2014_g.$$

This native STEM instrument computes predicted STEM education rates of young people who reach age 25 between 2009 and 2015 for each MSA. However, it does not use the actual location decisions of young adults, which could be influenced by local labor market conditions and other factors. Instead, it is based on the 1990 residential locations of young people, which they themselves did not choose. The instrument exploits the differential geographic distribution by age-gender-ethnic group of these young people in 1990 and the differential national STEM education rates of these young people by 2014. It computes the predicted STEM education rate among young people (ages 25-30 in 2015) in each MSA that would occur if all young people resided in the same location as in 1990 and earned a STEM degree at the national rate for their age-sex-ethnic group.

The intuition for the construction of the native STEM graduate instrument comes from expectations that location decisions are at least somewhat sticky and that STEM education differs at least somewhat by age-gender-ethnicity group. The stickiness of locations means that at least some young adults face various moving costs and are likely to reside in the same MSA as they did as children. Of course, some young people will likely leave their hometown before, during, or after attending college. However, some will never leave and many others will leave for college and come back afterward. The stickiness of locations and importance of moving costs have received considerable attention in the research literature (Bound et al. 2004; Kennan and Walker 2011; Winters 2011; Abel and Deitz 2012; Moretti 2013; Winters 2016, 2017). Moving costs can include financial costs of moving one's self and belongings, opportunity costs of time and effort to search and relocate, utility costs from risk aversion, and psychological and social costs from leaving friends, family, and familiar places behind. Internal migration in the U.S. has declined in recent years, and various moving costs have been posited as partial explanations among others (Molloy et al. 2011; Partridge et al. 2012; Cowen 2017). Thus, there is good reason to expect some stickiness in location decisions that is needed for the native STEM instrument.[29]

Significant differentials in STEM education rates by sex and ethnicity have been documented in previous literature (Bettinger and Long 2005; Carrell et al. 2010; Griffith 2010; Price 2010; Arcidiacono et al. 2012; Bottia et al. 2015; Arcidiacono et al. 2016). In particular, women have lower average STEM education rates than men, blacks and Hispanics have lower STEM education rates than non-Hispanic whites, and some Asian groups have especially high STEM education rates. My native STEM instrument uses relatively detailed ethnic groups, including exploiting heterogeneity among whites based on ancestry. Appendix Table A2 reports national STEM education rates by gender-ethnicity group averaged over ages 24-29 in the 2014 ACS. It also reports the national population share for each group and the non-STEM education shares for additional context. There is considerable variation in STEM education rates across gender-ethnicity consistent with previous literature. However, I also document that there is also considerable heterogeneity within white natives based on ancestry. For example, German

[29] The ACS does not include information for adults on MSA location during childhood, but it does include state of birth. In the 2014 ACS, 55.8 percent of native STEM graduates ages 24-29 reside in their state of birth, confirming a relatively large amount of stickiness.

Americans have meaningfully higher STEM education rates than Irish Americans. The various native ethnic groups are also unevenly distributed across the U.S. as children in 1990, which will facilitate variation in the instrument across MSAs.

To be valid, an instrument should be both relevant and exogenous. Relevant means that the instrument should be a strong predictor of the endogenous variable for which it is being used as an instrument. Exogenous means that the instrument should not be correlated with the regression error term, i.e., it should only be correlated with the dependent variable through its effect on the endogenous explanatory variable. First-stage diagnostic statistics can be used to test the relevance assumption. However, with only one instrument per endogenous variable, one cannot test the exogeneity assumption. Instead, one has to rely on motivating intuition and expectations.

For my specific application, the exogeneity of the native STEM instrument relies on two key assumptions. First is the assumption that the 1990 age-gender-ethnic distribution across MSAs for ages 0-5 is uncorrelated with innovation growth during 2009-2015 except through its effect on native STEM graduate levels. This assumption could be violated if innovative parents locate their children in areas in 1990 that are expected to become more innovative in the future or if innovation levels are diverging greatly across areas over a long time period. These problems could be exacerbated by intergenerational persistence in STEM education decisions and local persistence in innovation growth. However, these violations seem unlikely. In particular, the dependent variable in equation (2) above is measured as the difference in MSA patent intensity between 2009 and 2015, which removes any time-invariant factors affecting permanent innovation differentials. While parents in 1990 might be able to forecast MSA innovation demand over the next few years, they are unlikely to be able to forecast the growth during 2009-2015 that far out. Even if they could forecast that far out, a reasonable intertemporal discount rate and potential for later moves would greatly undermine any incentive to choose a location in 1990 based on innovation demand growth during 2009-2015. Strong long run persistence in innovation growth could be a problem more generally even without very forward-looking location decisions. Therefore, I include average innovation levels during 2007-2008, 1998-2006, and 1990-1997 as control variables. Additionally, I examined raw correlations for the patent variables in this study. The 2009-2015 change in log patents is only weakly correlated with patenting in earlier periods, with correlation coefficients less than 0.2 in absolute value.[30]

The second major exogeneity assumption for my native STEM IV approach is that national native STEM education differentials across groups are not driven by geographic differences in the changing demand for innovation during 2009-2015. For example, the second assumption could be violated if young people of a particular ethnic group are especially concentrated in a few metropolitan areas receiving sizable local shocks to innovation demand and alter their education decisions in response. However, I suggest that this is unlikely to be a significant problem, in part because most young people interested in STEM education choose the field by age 18, so that most young natives ages 24-29 in 2014 would have made their initial college major choices before 2009 and be unaffected by innovation demand shocks during 2009-2015. The control variables in the regression analysis should reduce concerns as well. I also consider several robustness checks below that should address possible concerns about the exogeneity assumption for the native STEM instrument. This includes excluding very high

[30] I also computed the correlation between patenting growth during 2009-2015 with the growth during previous periods and found similarly weak correlations.

STEM education areas from the sample and excluding very high STEM education ethnic groups from the instrument.

Ultimately, my native STEM instrument exploits supply-driven increases in local STEM graduate shares that are due to past demographics. The factors driving the instruments are expected to be unrelated to future demand growth for innovation. This supply-driven instrument is expected to provide exogenous variation useful for estimating causal effects. Policy makers would often like to know what would happen if they implemented policies that increased the stock of native STEM graduates. My IV strategy is designed to provide an answer to this question.

3.4 Foreign STEM Instrument

I identify the effects of foreign-born STEM graduates using a shift-share instrument similar in construction to Card (2001), Hunt and Gauthier-Loiselle (2010) and Peri et al. (2015). Specifically, I compute the 2009-2015 change in the predicted share of foreign STEM workers in each metropolitan area based on metropolitan area immigrant STEM worker shares for various origin countries in 1980 and the U.S. national growth in the immigrant STEM workforce from those source countries between 1980 and 2015.[31]

I first define a set of workers as STEM workers (more details below) and combine foreign origin countries into 14 groups.[32] I next use the 1980 decennial census 5% PUMS to

[31] I follow Peri et al. (2015) and choose 1980 as the base year for foreign employment shares for several reasons. First, census microdata geographic identifiers prior to 1980 greatly reduce the number of identifiable metropolitan areas. Second, 1980 precedes the creation of the H-1B visa program, so that base-year STEM shares are not affected by inflows due to the H-1B visa program. Third, 1980 precedes the information and communications technology (ICT) revolution, so that base-year STEM shares primarily reflect initial worker concentrations in other STEM fields.

[32] I follow Peri et al. (2015) and use the following 14 country groups: Canada, Mexico, Rest of Americas (excluding the U.S.), Western Europe, Eastern Europe, China, Japan, Korea, Philippines, India, Rest of Asia, Africa, Oceania, and Other.

compute the STEM worker share of each foreign nationality group, n, in each metropolitan area, c, relative to the total adult population of the metropolitan area in 1980,[33] i.e.,

$$STEMShare_{cn,1980} = \frac{\#of\ STEM\ Workers_{cn,1980}}{Total\ Adult\ Population_{c,1980}}.$$

I then combine the 1980 5% PUMS with the 2009 and 2015 ACS to compute the U.S. national growth factor in the number of STEM workers from each origin group, i.e.,

$$GrowthFactor_{n,1980-2009} = \frac{\#of\ STEM\ Workers_{n,2009}}{\#of\ STEM\ Workers_{n,1980}},$$

$$GrowthFactor_{n,1980-2015} = \frac{\#of\ STEM\ Workers_{n,2015}}{\#of\ STEM\ Workers_{n,1980}}.$$

I then compute the predicted share of foreign STEM workers in each metropolitan area in 2009 and 2015 by multiplying the 1980 STEM share for each origin group in each metropolitan area by the national growth factor and adding up across origin groups for each metropolitan area, i.e.,

$$STEM\widehat{Share}^{Foreign}_{c,2009} = \textstyle\sum_n [STEMShare_{cn,1980} \times GrowthFactor_{n,1980-2009}],$$

$$STEM\widehat{Share}^{Foreign}_{c,2015} = \textstyle\sum_n [STEMShare_{cn,1980} \times GrowthFactor_{n,1980-2015}].$$

These foreign STEM share variables compute the predicted foreign STEM worker share that each metropolitan area would experience in 2009 and 2015 if its origin-specific foreign STEM workforce grew at the national average since 1980.[34] The foreign STEM instrument that I use is the difference over time in the predicted foreign STEM shares between 2009 and 2015, i.e., the instrument is

[33] Census microdata geographic boundaries changed several times after 1980; as noted above, I use 1999 MSA definitions throughout. However, some MSAs cannot be matched over time and some are matched imperfectly.

[34] This IV approach also holds metropolitan area population at its 1980 level, but multiplying the population for each metropolitan area by a common factor to account for national population growth does not affect the 2SLS second-stage results since the first-stage coefficient on the instrument will adjust accordingly. Using actual metropolitan area population growth in the instrument would create concerns that the population growth is endogenous.

$$\Delta STEM\widehat{Share}^{Foreign}_{c,2009-2015} = STEM\widehat{Share}^{Foreign}_{c,2015} - STEM\widehat{Share}^{Foreign}_{c,2009}.$$

Equivalently, we could define the 2009-2015 national growth factor for each nationality as

$$GrowthFactor_{n,2009-2015} = \frac{\#of\ STEM\ Workers_{n,2015} - \#of\ STEM\ Workers_{n,2009}}{\#of\ STEM\ Workers_{n,1980}},$$

and write the foreign STEM instrument as

$$\Delta STEM\widehat{Share}^{Foreign}_{c,2009-2015} = \sum_n [STEMShare_{cn,1980} \times GrowthFactor_{n,2009-2015}].$$

The motivation for the foreign STEM instrument is based on three stylized facts. First, the foreign STEM workforce has historically been disproportionately concentrated in certain areas relative to the native population. Second, there has been a large influx of skilled foreigners to the U.S. in recent decades, including during 2009-2015. Third, recently arriving foreign-born workers tend to concentrate in areas that already have a relatively high share of foreigners from their country of origin. This clustering in areas with similar countrymen can result from benefits in employment, consumption, and social interactions. For example, recently arriving foreign workers may have limited access to job information, and some employers may be reluctant to hire them because of greater difficulty in understanding their education and work histories. Locating in an area with a relatively large density of other foreign workers from the same country may facilitate networking that increases an individual worker's knowledge about local job opportunities. Previous foreign workers may also act as an intermediary to help employers better screen applicants and may be managers or employers themselves.[35]

[35] Additionally, a larger density of a specific group of foreign-born residents can facilitate greater variety and quality of consumption goods such as ethnic restaurants, stores, and entertainment that foreign-born workers from a given country might find especially desirable. Similarly, national origin similarities may promote desirable social and cultural opportunities not available in areas with few foreigners from a given country.

To be valid, the foreign STEM instrument should be both relevant and exogenous. We will examine instrument relevance below based on first-stage diagnostic tests. With only one foreign STEM instrument, we cannot test the exogeneity assumption. However, intuition and prior literature suggest it is likely an exogenous instrument. Similar variants of my foreign STEM instrument have been used by Card (2001), Hunt and Gauthier-Loiselle (2010), Peri et al. (2015), and many others. In my time-differenced 2SLS analysis, exogeneity requires that the geographic distribution of foreign STEM workers in 1980, $STEMShare_{cn,1980}$, be uncorrelated with the error term in equation (2). The 1980 foreign STEM distribution should only affect patenting through its effect on the 2009-2015 foreign STEM graduate share. Additionally, the shift-share foreign IV strategy assumes that the U.S. national growth factor for each origin group during 2009-2015 is driven by national level forces and not MSA-specific factors correlated with the error term in equation (2). One cannot test this directly, but I experiment with some sensitivity analysis by removing a few metropolitan areas with very high STEM shares. Ultimately, the foreign STEM graduate instrument is intended to exploit exogenous supply-driven variation via past immigrant location decisions and national immigration levels that are exogenous to individual MSAs.

There are a couple of caveats about the foreign STEM share instrument. First, the potentially endogenous explanatory variable is the change in the metropolitan area share of foreign STEM *graduates*, and the instrument is the predicted change in the share of foreign STEM *workers*, which is based primarily on occupation. Occupation and college major are not perfectly collinear, but there is no known better way to instrument for foreign STEM graduates. Fortunately for my analysis, most foreign workers in STEM occupations have college degrees in STEM fields, so STEM graduates and STEM workers are closely related. Because I am ultimately intending to measure the effects of STEM graduates on innovation, the foreign STEM instrument in this study is restricted to persons who have completed at least a bachelor's degree and are working in a STEM occupation; the instrument therefore does not include workers in STEM occupations who have less than a bachelor's degree. A second caveat is that there is no universally agreed upon definition of STEM occupations. My STEM occupation measure is restricted to include only persons employed as engineers, mathematicians, computer scientists, software developers, and natural scientists.[36] Broader definitions of STEM workers may be less connected intuitively and empirically to future flows of foreign STEM graduates.

4. Empirical Results

4.1 OLS Estimates

Table 4 presents cross-sectional OLS regression results treating the human capital explanatory variables as exogenous. The three columns include different combinations of the human capital explanatory variables, but all regressions include the full set of control variables listed in Table 1. The first column includes the STEM graduate population share; OLS regression yields a coefficient estimate of 3.44 that is statistically significant at the one percent level. The coefficient suggests that increasing the share of the adult population with a STEM degree by one percentage point (i.e., increasing the share by 0.01) would increase log patents per 100,000 population by about 0.034. Since the dependent variable is measured in logs, we can interpret this result as a roughly 3.4 percent increase in patent intensity due to a one percentage point increase in the STEM graduate share. Multiplying the coefficient by the STEM variable standard deviation of 0.030, suggests that a one standard deviation increase in the STEM

[36] I use the IPUMS variable occ1990 and define the following occupation codes as STEM: 44, 45, 47, 48, 53, 55, 56, 57, 59, 64, 66, 68, 69, 73, 74, 75, 76, 77, 78, 79, 83, and 229.

graduate share would increase patent intensity by just over 10 percent. This is an economically meaningful magnitude and suggests an important relationship between STEM graduates and innovation as measured by patent intensity.

Column 2 decomposes the STEM graduate share into the native-born and foreign-born components in an attempt to assess their relative contributions to patenting. Both yield positive coefficients statistically significant at the five percent level. The respective coefficients of 3.252 and 3.565 for native STEM and foreign STEM are quite close in magnitude. Column 3 adds the non-STEM graduate population share to the specification in column 2. The foreign STEM coefficient is largely unaffected, but the native STEM coefficient decreases somewhat to 2.515 and is now only significant at the ten percent level. The non-STEM coefficient is relatively small (0.847) and not statistically significant at conventional levels. Thus, though STEM graduates appear to be positively related to innovation, non-STEM graduates do not appear significantly related to increased innovation. However, these cross-section OLS results may not provide causal estimates.

I next examine the relationships between time differences in the dependent variable and the main human capital explanatory variables during the 2009-2015 time period. Time-differenced OLS results are in Table 5, which is structured similarly to Table 4 with three columns. The first column suggests no significant relationship between the 2009-2015 change in log patents and the change in total STEM graduate population share. This is in stark contrast to the significant positive cross-sectional relationship in Table 4. Separating the change in the STEM graduate share into foreign and native components in column 2 present a more nuanced story. The change in the foreign STEM share has a positive coefficient of 4.679 that is significant at the ten percent level; the magnitude is even larger than the corresponding cross-section estimate in Table 4, but the difference is not statistically significant. In contrast, the native STEM graduate coefficient is -2.693, though not statistically significant at conventional levels. Adding the change in the non-STEM graduate share in column 3 gives nearly identical estimates as column 2 for the foreign and native STEM shares, and the non-STEM share coefficient is small and not significant.

The results in Table 5 are only descriptive, but they offer provocative suggestive evidence about recent relationships between local human capital growth and innovation. Foreign STEM graduates appear to increase local innovation, which is consistent with prior work. Non-STEM graduates appear to have little relationship with innovation, which is somewhat to be expected. The startling result is that the growth in the share of the population who are native STEM graduates is not positively related to innovation. Assessing the validity of this result requires further analysis. I next turn to instrumental variables results intended to provide causal estimates.

4.2 Main IV Estimates

Table 6 presents the main results in this paper, effects of native and foreign STEM graduates on MSA patenting intensity using 2009-2015 time-differenced 2SLS regressions. The table columns mostly follow Table 4 and 5 but deviate slightly by first examining effects of native STEM and foreign STEM separately in columns 1 and 2, respectively. Column 3 includes these STEM variables simultaneously, and column 4 adds the change in the non-STEM graduate share as a control variable to the column 3 specification. Panel A reports first-stage regression results, and Panel B provides second-stage results for the explanatory variables of interest. All models again include the control variables listed in Table 1.

First-stage results in column 1 confirm that the predicted young native STEM graduate share is a strong predictor of the change in the native STEM graduate share. In panel A, the instrument has a positive and statistically significant coefficient. The first-stage F-statistic equals 12.08, which comfortably exceeds 10, suggesting that weak instrument concerns are minimal (Stock et al. 2002; Angrist and Pischke 2009). I also report underidentification and weak identification test statistics suggested by Kleibergen and Paap (2006), which will be especially useful for the models with multiple endogenous explanatory variables in columns 3-4. For column 1, the Kleibergen-Paap (KP) test strongly rejects underidentification and weak identification concerns are again minimal.

More important are the second-stage results. Column 1 reports a second-stage coefficient estimate (-5.161) for the effect of the change in the native STEM graduate share on the change in patent intensity that is negative but not statistically significant. The coefficient estimate is not very precise, but the evidence available suggests a non-positive effect. This finding is qualitatively consistent with the negative and insignificant coefficient for native STEM graduates in Table 5. Thus, both time-differenced OLS and 2SLS indicate that increasing the stock of native STEM graduates did not increase patent intensity during the 2009-2015 period.

First-stage results in column 2 confirm that the foreign STEM instrument is a strong predictor of the growth in the foreign STEM graduate share according to both the traditional first-stage F-statistic and the KP statistics. The second-stage results indicate that the growth in the foreign STEM graduate share has a positive effect on log patenting growth with a coefficient of 11.272 that is significant at the ten percent level. This coefficient implies that a one percentage point increase in the MSA population share of foreign STEM graduates increases patents per 100,000 population by roughly 11.3 percent. This is a sizable effect and the magnitude is nearly twice that for foreign STEM graduates in Table 5, though the difference in these coefficients is not statistically significant at conventional levels. It seems possible that the time-differenced OLS estimates in Table 5 might be attenuated toward zero due to measurement error from sampling, which could explain the difference. My preferred estimates are the 2SLS results in Table 6.

Columns 3 and 4 of Table 6 each include two endogenous explanatory variables and two corresponding instruments, so traditional first-stage F-statistics are not the appropriate diagnostic tool to test for weak instruments. Instead, one needs instrument tests that account for the fact that two or more variables are being instrumented. Kleibergen and Paap (2006) provide appropriate tests for underidentification and weak identification useful for this setting. The KP Lagrange Multiplier (LM) statistic provides a test for underidentification and the KP Wald statistic provides a test for weak identification in combination with critical values reported by Stock and Yogo (2005). In both columns 3 and 4, the KP LM statistic allows us to reject the null of underidentification with a p-value of 0.001. For weak identification, the KP Wald statistics (5.543 and 5.648) exceed the critical value to reject 15% max IV size (4.58), but fail to exceed the critical value to reject 10% max IV size (7.03).[37] Thus, one cannot completely rule out the possibility of second-stage test size distortion in columns 3 and 4, but any such distortion is likely to be relatively small.

The second-stage results are nearly identical between columns 3 and 4, and the non-STEM variable in column 4 has a very small coefficient estimate that is essentially zero. The second-stage results in columns 3 and 4 of Table 6 are also qualitatively similar to those in columns 1 and 2. The foreign STEM graduate coefficients are 11.180 and 11.185 in columns 3

[37] These weak identification test statistics are based on a true test size of 5%. Weak instruments can cause the IV size to exceed the true size, and the KP Wald statistics help one rule out that IV size is above some level (the max).

and 4, respectively. The native STEM graduate coefficient estimates are -9.467 and -9.505, respectively, which are larger in magnitude than the estimate in column 1 but not statistically different. Thus, the 2SLS results in Table 6 are qualitatively robust across specifications. The results suggest that foreign STEM graduates have a strong positive effect on metropolitan area innovation but recent growth in native STEM graduates has no meaningfully positive effect on innovation.

The 2SLS estimate for the effect of skilled foreigners on patent intensity is roughly consistent with the most closely related previous study despite the differences in variable measures, time period considered, and econometric specification. Hunt and Gauthier-Loiselle (2010) report a variety of specifications, but their most detailed IV specification indicates that a one percentage point increase in the share of foreign-born college graduates increases patents per capita by 12.3 percent. My Table 6 column 3 coefficient estimate implies an 11.2 percent increase in patent intensity from a one percentage point increase in the share of foreign-born STEM graduates. Both imply that skilled foreigners increase innovation in the U.S. with sizable magnitudes.[38]

4.3 IV Estimates for Effects on STEM Occupation Employment Shares

The stark difference between native and foreign STEM graduate effects on patenting is quite surprising. One might wonder why the results are so different and if this contrast would occur for other related outcomes. I next look at STEM occupation employment shares.

[38] Global effects, however, might be smaller if foreigners would have created innovations in their origin countries had they located there. Of course, there is good reason to believe that moving to an innovative country like the U.S. would make a skilled foreigner more innovative than they would be in a less innovative origin country if they are combined with more and better resources in the U.S. useful for innovation (Kahn and MacGarvie 2016). In particular, concentrating STEM graduates in U.S. metropolitan areas increases a skilled immigrants interactions with other skilled workers and is expected to create agglomeration economies such as learning, knowledge spillovers and cross-fertilization of ideas.

Specifically, I compute two additional outcome variables: 1) the share of all adults (ages 25 and up) in the MSA who are native born and work in a STEM occupation and 2) the share of all adults in the area who are foreign born and work in a STEM occupation. I then use the same time-differenced 2SLS procedure as in Table 6 to examine how supply-driven increases in the local stocks of native and foreign STEM graduates affect these two new STEM occupation measures. Table 7 presents second-stage results for this analysis, with native STEM occupation shares as the dependent variable in Panel A and foreign STEM occupation shares in Panel B. The three columns in Table 7 are parallel to columns 1-3 in Table 6, and the first-stage results in Table 7 are consequently identical to those in Table 6. For conciseness, I no longer include the robustness check with the change in the non-STEM graduate share as a control variable, but including the non-STEM graduate control variable does not meaningfully alter the results.

In Panel A of Table 7, I am first interested in how a supply-driven increase in the MSA stock of native STEM graduates affects the share of the local population who are native-born and employed in a STEM occupation. Columns 1 and 3 both indicate large, positive, and statistically significant effects of native STEM graduate shares on native STEM employment shares, with coefficients of 0.411 and 0.441, respectively. This indicates that a one percentage point increase in the population share of native STEM graduates translates into a 0.4 percentage point increase in the population share of native STEM occupation employment. This effect is statistically different from both zero and one. Thus, there is a clear positive relationship, indicating that increasing the stock of native STEM graduates increases the native STEM workforce. However, the effect is less than proportional. Some of the additional STEM graduates will likely end up in STEM-related occupations (e.g. health care and STEM education) that are not included in my somewhat narrow STEM occupation definition. Others will end up in occupations that are at most weakly related to STEM. Ultimately, many STEM graduates end up in non-STEM occupations, and it is not very surprising that this relationship is less than proportional.

Another concern for Panel A relates to whether supply-driven increases in foreign STEM graduates in an MSA affect native STEM occupation employment shares. A negative effect would indicate that foreigners are crowding out natives, and a positive effect would indicate that strong complementarities and spillovers from skilled foreigners are pulling natives in. The coefficient estimates in columns 2 and 3 are both negative but neither is statistically significant at conventional levels and the magnitudes are relatively small. Thus, we can’t draw strong conclusions on this, but very strong complementary effects from foreign STEM on native STEM employment seem unlikely.

Panel B examines the foreign STEM occupation employment share outcome. The coefficient for the native STEM graduate share in column 1 is positive and significant at the ten percent level, but the coefficient becomes virtually zero in column 3. I put more faith in the column 3 results here and cautiously interpret this to suggest that there is likely no meaningful relationship between supply-driven increases in native STEM graduates and foreign STEM occupation employment shares. However, there is a positive relationship between the foreign STEM graduate share and the foreign STEM occupation share, consistent with expectations. The coefficient of 0.384 is very similar in magnitude to the coefficient for native STEM graduate effects on native STEM employment.

The results in Table 7 are useful for several reasons. First, they address a topic of interest in its own right and indicate that supply-driven increases in the local stock of STEM graduates increases the local share of STEM occupations. Second, they strengthen confidence in the patent intensity results in Table 6 using the same approach, i.e., the non-positive effects of native

STEM graduates on patenting are not due to some extensive failure in the empirical strategy. The same strategy applied to a related outcome gives significant results in the expected direction. Third, it rules out one possible mechanism for the non-positive effect of native STEM graduates on patenting in Table 6. This non-positive effect is likely not due to forces preventing natives from participating in STEM employment. Native STEM graduates significantly increase the STEM workforce at relative rates comparable to foreign STEM graduates, but they have very different effects on patenting.

4.4 IV Sensitivity Analysis

Given the somewhat surprising results for natives in Table 6, I conducted extensive sensitivity analysis. Second-stage results are in Table 8 with the three columns parallel to columns 1-3 of Table 6. Each panel contains a separate sensitivity check.[39] First-stage results for most of these are very similar to Table 6 and are not reported to conserve space. However, first-stage results for a few of these differ somewhat from Table 6, and in such cases I report and briefly discuss the KP weak identification statistic. The qualitative results are largely insensitive to the various checks in Table 8. The coefficients for foreign STEM graduates lose significance under some alternatives, but in no case are the results statistically different from the preferred results in Table 6 at conventional significance levels.

[39] Some additional checks are not reported. For example, I also experimented with instrumenting for the STEM graduate share using the land grant dummy variable used by Moretti (2004) and others. Doing so produced very weak first-stage results indicating that it is not a useful instrument for my setting. I also experimented with testing for nonlinearities in the foreign and native STEM graduate share variables. Adding quadratic terms for the human capital variables doubles the number of required instruments. A potential solution might be to include linear and squared versions of each of the instruments in the first-stage. Unfortunately, the linear and squared versions of the instruments are very highly correlated, and this approach gives a second-stage that is underidentified, meaning that second-stage results are not meaningful. I also experimented with non-linearities for the OLS results, but the estimates were very noisy as one might expect and potential endogeneity would complicate OLS interpretation even if results were precise.

One possible concern is that the results might be affected by the Great Recession that began in December 2007 and ended in June 2009. For example, the recession might have distorted year 2009 patent levels in problematic ways. I first address this by controlling for initial patent levels in Panel A, i.e., by including year 2009 log patents per 100K population as a control variable. This does not qualitatively alter the results. In Panel B, I measure the dependent variable as the change in log patents per 100K over the 2006-2015 period, so that the start period precedes the recession. Results are again qualitatively similar. Panel C measures the dependent variable using two-year averages for both the start period (2009-10) and the end period (2014-15) and also gives qualitatively similar results.

Appendix Table A2 shows that native STEM education rates are especially high among Asian Americans. Asian Americans are a small portion of the population, and it is useful to confirm that this small group is not an outlier driving the results. Panel D of Table 8 excludes Asian Americans from the native STEM graduate instrument; the native STEM graduate share coefficient actually becomes more negative but is again not statistically significant.

Panel E controls for the 2009 STEM occupation share of MSA employment and Panel F controls for the 2009-2015 change in the STEM occupation share. Both of these variables are potentially endogenous and therefore excluded from the preferred specifications in Table 6. Results in Panel E are minimally affected. In column 3 of Panel F, the foreign STEM graduate coefficient is largely unchanged but the p-value drops to 0.109. However, the coefficient for the change in STEM occupation share (not shown) is small and very far from statistically significant. Furthermore, this additional variable weakens the first-stage diagnostics somewhat and can be justifiably excluded from the preferred results given the potential endogeneity.

Panel G excludes from the analysis the five MSAs with the highest STEM graduate shares in 2015. The first-stage diagnostics indicate weaker predictive power for the instruments. The second-stage coefficient for foreign STEM is no longer statistically significant in either columns 2 or 3. However, the magnitudes are fairly similar to the preferred specification and there is no compelling reason to exclude the five largest MSAs. Panel G might cast a little skepticism about the robustness of positive effects of foreign STEM on patenting, but the preferred specification retains these MSAs. Additionally, Panel G does nothing to contradict the non-effect of native STEM graduates on patenting, which is the main contribution of the paper.

Finally, Panel H adds state dummies as control variables, which considerably weakens the instruments especially in column 3. The foreign STEM graduate coefficient is again no longer significant in either column 2 or 3. Similar to the discussion for Panel G, there is no overwhelming reason why state dummies should be included. It might increase confidence in the results if they were completely unaffected by including state dummies, but their inclusion is not clearly warranted. Given that they weaken the first-stage explanatory power of the instruments and have no compelling reason for inclusion, my preferred estimates do not include state dummies.[40] Still, it is at least somewhat reassuring that including state dummies does not give significantly different coefficient estimates.[41]

[40] The preferred results use time-differenced MSA data which remove time-invariant MSA fixed effects. MSA fixed effects absorb state fixed effects (for those MSAs that do not cross state boundaries or when allocating MSAs to the primary state). With a time-differenced dependent variable like this, state dummies are not state fixed effects, they are state-by-time-period effects. They capture time-varying factors that are common to all areas within the state during the time period. Furthermore, they effectively limit the analysis to within-state variation. Some states have relatively few MSAs in them and thus very little within-state variation. This can greatly reduce the statistical power of the analysis, and this can be exacerbated in 2SLS. Since state dummies weaken the explanatory power of the exogenous instrumental variables, I argue that they should be excluded on the grounds of efficiency.

[41] An additional sensitivity check conducted but not reported in Table 8 was to construct MSA human capital variables using the population ages 25-64 rather than 25 and up. The second-stage native STEM graduate coefficient was again negative but not statistically significant in the columns 1 and 3 specifications. The foreign STEM graduate variable was positive and significant at the ten percent level with coefficients of 10.113 and 10.177 in the columns 2 and 3 specifications. My preferred specification continues to use ages 25 and up for the human capital variables because many older persons still work and those who don't may still contribute to urban

5. Further Discussion of Native STEM Results

The lack of a statistically significant positive effect of native STEM graduates on patent intensity for the preferred results in this study is quite surprising and even somewhat alarming. Some additional discussion of the validity and interpretation seems warranted.

5.1 Validity

One might worry that the non-effect is due to unknown problems with the 2SLS method used. However, the coefficient estimate for native STEM graduates is negative and statistically insignificant for both time-differenced OLS and 2SLS. Of course, the OLS estimate could be biased, but most concerns about omitted variables, reverse causality, or sorting would incline one to expect the OLS estimate to have a positive bias instead of a negative one.[42] Taking the time-differenced OLS results at face value rules out a reasonably large positive effect from native STEM graduates on innovation. Additionally, I show that the time-differenced 2SLS approach identifies a significant positive effect of native STEM graduates on native STEM occupation employment shares consistent with expectations. Thus, the IV strategy has the power and ability to find an effect that we likely expect to occur for another related outcome.

One might next worry that patents are an imperfect measure of innovation and might not capture what we really care about for innovation. Furthermore, the positive effect of native STEM graduates on STEM occupation employment is likely a good thing that indicates some benefits to society. In response, I concede that patents are not a perfect measure of innovation, and I agree that the positive effect on STEM occupation employment likely has benefits broader

---

innovation, either directly through developing their own innovations at home or indirectly through knowledge spillovers on others.

[42] Classical measurement error due to sampling should attenuate the OLS coefficient estimate toward zero but it should not bias the sign of the coefficient estimate.

than on innovation. However, patents are likely the single best measure of innovation available for my purposes. For example, incorporating information on patent citations or commercial marketability is difficult for using recent data since these measures and broader benefits of innovation are often not revealed until much later in the future. Patent counts are a generally accepted measure of innovation and have been used widely in previous literature. Even if patent counts reflect only one component of innovation, it is certainly an important component and useful to study. Furthermore, the positive relationship between skilled foreign workers and patent intensity in my study and previous literature might lead us to expect a similar effect for native STEM graduates that I do not find.

One might also worry that the 2009-2015 time period is too short or too influenced by the Great Recession to give "normal" results. The positive effect on STEM occupation shares should partially reduce concerns about the shortness of the time period. Furthermore, I do find a positive effect of foreign STEM graduates on innovation that is significant at the ten percent level. Thus, the time period seems sufficient to reveal an effect of STEM graduates on patent intensity. I also show that the non-effect for native STEM graduates on patenting is robust to controlling for the initial patenting level in year 2009 and to measuring the change in patenting over a longer time period that began prior to the Great Recession. The non-effect is also consistent across a large number of other robustness checks. The non-effect for the 2009-2015 period may not accurately describe the relationship for considerably earlier years, but the non-effect does accurately describe the most recent period available and is the most appropriate reference point for expectations about the near future.

Another possibility is that diminishing marginal benefits might have caused American innovation to have reached a satiation point with respect to American STEM graduates. The IV strategy in this study prevents me from offering precise estimates of non-linear effects. However, if the non-effect of native STEM graduates on patenting was driven by diminishing returns, we might expect to see positive effects if we limit the sample to MSAs with more modest levels of STEM graduates. In results not shown, I experimented with several thresholds for removing the most STEM educated cities from the sample and consistently found non-effects of native STEM graduates on patent intensity. Thus, the non-effect on patenting does not appear due to satiation with respect to American STEM graduates.

Another possible concern is that IV estimates the local average treatment effect (LATE), which could differ from the average treatment effect (ATE) in the presence of parameter heterogeneity. Specifically, the IV for native STEM graduates is based on the inflow of young STEM graduates to the workforce. If marginal increases in young STEM and older STEM graduates have different effects on innovation, then the IV results in this study might not extrapolate to broader settings. Similarly, the results in this study indicate no positive effect of native STEM graduates on innovation in the medium run, but that does not rule out possible positive effects over the longer run. I concede these points and caution that readers should not draw more from the analysis than the data and methods permit. Still, the focus on young native STEM graduates is important and likely most closely aligns with potential policy levers to increase native STEM graduates. In the medium run, having more young native STEM graduates does not appear to increase innovation.

5.2 Interpretation

Once one accepts that young native STEM graduates have not increased MSA patent intensity in recent years, one might next wonder why not. On this, I can only speculate, but it appears consistent with decreased dynamism in the U.S. that has been discussed in a variety of contexts and summarized by Cowen (2017).

As discussed above, Americans have become much less geographically mobile in recent years. This appears at least partially due to reduced willingness to move to a new area to take a job (Partridge et al. 2012). However, Americans have also become less willing to change jobs even in their current location (Hyatt and Spletzer 2013; Molloy et al. 2017). While educated workers are historically more mobile than the less educated, their geographic mobility has declined too, and human capital levels across areas are becoming stickier over time (Winters 2017). Declining mobility could contribute to reduced dynamism in other dimensions if people have less exposure to other areas and experiences, but it might also be a symptom of other broader forces.

America is also experiencing declining self-employment rates, fewer startup firms, and worse startup performance (Decker et al. 2014, 2016; Goldschlag and Tabarrok 2014; Cowen 2017). This may be partially attributable to reduced incentives to start a firm over time, but the exact source of reduced incentives is unclear (Decker et al. 2014). Increased government regulation has been proposed as a partial explanation for declining entrepreneurship, but the empirical evidence does not support this explanation (Goldschlag and Tabarrok 2014; Molloy et al. 2016). Other explanations that have been rigorously tested empirically have at best received weak support. Many discussions end up focusing on cultural trends or changing preferences, which are hard to measure and empirically test. Cowen (2017) argues that on average, Americans, especially young Americans, have become more complacent, more resistant to change, and less willing to take risks.

Cowen (2017) also argues that Americans are becoming more segregated in several ways. Segregation has always been a troublesome aspect of American life, but it is taking new forms. Technology is altering our lives in ways that allow us to interact more with people like us and interact less with people different from us. Information is increasingly accessible to the point of overload. Search engines, social media, and matching algorithms increasingly help us choose which information to consume and which people to interact with. The problem, according to Cowen (2017), is that these algorithms and our own nature push us toward the familiar and comfortable and reduce our exposure to new people, new experiences, and new ideas. At an individual level, the benefits of using information technology outweigh the costs for most people. However, the creation and spread of useful ideas have considerable positive externalities. When individuals increasingly close themselves off to new ideas, society becomes less innovation and eventually worse off.

6. Conclusion

Technological innovation is widely regarded as a key driver of economic growth both for nations and regions, and human capital is thought to play an important role in fueling innovation. However, some types of human capital may have greater effects on innovation than others. In particular, STEM graduates are typically expected to increase local innovation. A few researchers and policymakers have suggested that native and foreign college graduates may have differing effects on innovation. This paper examines differences in patent intensity across U.S. metropolitan areas to assess the importance of different types of human capital on innovation. Consistent with previous literature, my results suggest a positive effect of foreign STEM graduates on innovation in the U.S. Policies that increase the stock of foreign STEM graduates increase innovation and provide considerable economic benefits to regions and nations. I also find suggestive evidence that college graduates educated in non-STEM fields have minimal effect on MSA innovation. Of course, college education in non-STEM fields likely has other significant benefits relative to no higher education.

My primary contribution is to use a novel time-differenced instrumental variable strategy to estimate causal effects of native STEM graduates on metropolitan area innovation. My approach uses 2009-2015 time differences in patent intensity and MSA STEM graduate shares. Time-differencing removes the influence of metropolitan area fixed effects, i.e., time-invariant factors that are fixed over time and might be correlated with both skill levels and innovation. I instrument for the 2009-2015 change in the native STEM graduate share using a shift-share type instrument based on the MSA-specific age-gender-ethnicity distribution of children ages 0-5 in 1990 who would be ages 25-30 in 2015. I combine this 1990 data with national STEM education rates for the age-gender-ethnicity groups in the 2014 ACS to construct my instrument for native STEM graduates. First-stage results confirm that my instrument has a strong significant effect on the growth in MSA native STEM graduate shares during 2009-2015 as these young people enter the labor force. However, the second-stage regression results suggest that this increase in native STEM graduate shares does not increase innovation. The coefficient estimate is negative though not statistically significant. My 2SLS estimates are not sufficiently precise to rule out modest positive effects, but the best evidence available suggests no significant positive effect on MSA patent intensity.

The lack of a significant positive effect of native STEM graduates on innovation is surprising and warrants the attention of researchers and policymakers. Innovation is critical for economic growth. We need to better understand what drives innovation, the role that formal education can play, and what other complementary institutions can strengthen the link between native STEM graduates and innovation. STEM education may provide technical skills that could be useful for innovation in some settings, but these skills appear to be underutilized among young American STEM graduates in the current environment. We need to better understand why and figure out how American STEM graduates can be better utilized to fuel innovation.

To help inform my analysis, I also use my 2SLS approach to test whether increases in native STEM graduates in an MSA increase native STEM occupation employment in the MSA. In this case, I find strong positive effects that are statistically significant and comparable to the effect of foreign STEM graduate increases on foreign STEM employment. This is a notable finding for various reasons. For one, STEM occupations may have social benefits other than the effects on patenting, so there is likely some useful effect of policies that increase native STEM education rates. Perhaps more importantly for understanding the patenting analysis, the significant positive effect of native STEM graduates on native STEM employment rules out a possible mechanism for the non-effect of natives STEM graduates on patent intensity. This non-effect is not due to barriers preventing young native STEM graduates from entering STEM occupations.

I can only speculate at this time, but I discuss that the estimated non-effect of native STEM graduates on innovation may be related to reduced dynamism experienced across a broad range of measures in recent years. There is growing concern that Americans are becoming more complacent, less mobile, and less willing to take risks. My study also suggests that young Americans may be becoming less intensely involved in innovation. Given the enormous importance of innovation for economic growth, prosperity, and security, this is a significant concern in need of additional attention from both researchers and policymakers.

References

Abel, Jaison R. and Richard Deitz. 2012. Do colleges and universities increase their region's human capital? Journal of Economic Geography, 12, 667-691.

Adams, James D. 2002. Comparative localization of academic and industrial spillovers. Journal of Economic Geography, 2(3), 253-278.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas. 2017a. Immigration and the Rise of American Ingenuity. American Economic Review, 107(5), pp.327-31.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas. 2017b. The rise of American ingenuity: Innovation and inventors of the golden age. NBER Working Paper No. 23047.

Ali, Kamar, Mark D. Partridge, and Dan S. Rickman. 2012. International immigration and domestic out-migrants: Are domestic migrants moving to new jobs or away from immigrants? Annals of Regional Science, 49(2): 397-415.

Altonji, Joseph G., and Seth D. Zimmerman. 2017. The costs of and net returns to college major. NBER Working Paper No. 23029.

Angrist, Joshua D. and Jorn-Steffen Pischke. 2009. Mostly Harmless Econometrics. Princeton: Princeton University Press.

Anselin, Luc, Attila Varga, and Zoltan Acs. 1997. Local geographic spillovers between university research and high technology innovations. Journal of Urban Economics, 42(3): 422-448.

Arcidiacono, Peter, Esteban M. Aucejo, and V. Joseph Hotz. 2016. University differences in the graduation of minorities in STEM fields: Evidence from California. American Economic Review, 106(3), 525-562.

Arcidacono, Peter, Esteban M. Aucejo, and Ken Spenner. 2012. What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice. IZA Journal of Labor Economics, 1(5): 1-24.

Atkinson, Robert D. and Merrilea Mayo. 2010. Refueling the U.S. innovation economy: Fresh approaches to science, technology, engineering and mathematics (STEM) education. Washington, DC: Information Technology and Innovation Foundation.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? Quarterly Journal of Economics, 119(1): 249-275.

Bettinger, Eric P. and Bridget Terry Long. 2005. Do faculty serve as role models? The impact of instructor gender on female students. American Economic Review 95(2): 152-157.

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

Borjas, George J. 2006. Native internal migration and the labor market impact of immigration. Journal of Human Resources, 41(2): 221-258.

Bottia, Martha C., Elizabeth Stearns, Roslyn Arlin Mickelson, Stephanie Moller, and Lauren Valentino. 2015. Growing the roots of STEM majors: Female math and science high school faculty and the participation of students in STEM. Economics of Education Review 45, 14-27.

Bound, John, Jeffrey Groen, Gabor Kezdi, and Sarah Turner. 2004. Trade in university training: Cross-state variation in the production and stock of college-educated labor. Journal of Econometrics, 121(1-2), 143-173.

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

Carlino, Gerald A., Satyajit Chatterjee, and Robert M. Hunt. 2007. Urban density and the rate of invention. Journal of Urban Economics, 61(3), 389-419.

Carlino, Gerald A. and Robert M. Hunt. 2009. What explains the quantity and quality of local inventive activity? Brookings-Wharton Papers on Urban Affairs, pp.65-123.

Carlino, Gerald and William R. Kerr. 2014. Agglomeration and innovation. NBER Working Paper No. 20367.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. Sex and science: How professor gender perpetuates the gender gap. Quarterly Journal of Economics 125(3): 1101-1144.

Chatterji, Aaron, Edward Glaeser, and William Kerr. 2014. Clusters of entrepreneurship and innovation. Innovation Policy and the Economy, 14(1), 129-166.

Chellaraj, Gnanaraj, Keith Maskus, and Aaditya Mattoo. 2008. The contribution of skilled immigrations and international graduate students to U.S. innovation. Review of International Economics, 16(3): 444-462.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. Spatial wage disparities: Sorting matters! Journal of Urban Economics, 63(2): 723-742.

Corley, Elizabeth A. and Meghna Sabharwal. 2007. Foreign-born academic scientists and engineers: producing more and getting less than their US-born peers? Research in Higher Education, 48(8): 909-940.

Cowen, Tyler. 2017. The Complacent Class: The Self-Defeating Quest for the American Dream. St. Martin's Press: New York.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. The role of entrepreneurship in US job creation and economic dynamism. Journal of Economic Perspectives, 28(3): 3-24.

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2016. Where has all the skewness gone? The decline in high-growth (young) firms in the US. European Economic Review, 86: 4-23.

Denning, Jeffrey T. and Patrick Turley. 2017. Was that smart? Institutional financial incentives and field of study. Journal of Human Resources, 52(1), 152-186.

Facchini, Giovanni, and Max Friedrich Steinhardt. 2011. What drives U.S. immigration policy? Evidence from congressional roll call votes. Journal of Public Economics 95(7): 734-743.

Faggian, Alessandra, and Philip McCann. 2009. Human capital, graduate migration and innovation in British regions. Cambridge Journal of Economics, 33(2): 317-333.

Gaulé, Patrick and Mario Piacentini. 2013. Chinese graduate students and U.S. Scientific Productivity. Review of Economics and Statistics, 95(2): 698-701.

Glaeser, Edward L., William R. Kerr, and Giacomo Ponzetto, 2010. Clusters of entrepreneurship. Journal of Urban Economics, 67(1), 150-68.

Glaeser, Edward L., Sari Pekkala Kerr, and William R. Kerr. 2015. Entrepreneurship and urban growth: An empirical assessment with historical mines. Review of Economics and Statistics, 97(2), 498-520.

Goldschlag, Nathan, and Alexander T. Tabarrok. 2014. Is Regulation to Blame for the Decline in American Entrepreneurship? George Mason University Working Paper in Economics No. 15-11.

Griffith, Amanda. 2010. Persistence of women and minorities in STEM field majors: Is it the school that matters? Economics of Education Review, 29(6): 911-922.

Gurmu, Shiferaw, Grant Black, and Paula Stephan. 2010. The knowledge production function for university patenting. Economic Inquiry, 48(1): 192-213.

Hunt, Jennifer. 2011. Which immigrants are most innovative and entrepreneurial? Distinctions by entry visa. Journal of Labor Economics, 29(3): 417-57.

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

Hyatt, Henry R., and James R. Spletzer. 2013. The recent decline in employment dynamics. IZA Journal of Labor Economics, 2:5.

Jaffe, Adam. 1989. Real effects of academic research. American Economic Review, 79, 957-70.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics, 108(3): 577-598.

Kahn, Shulamit, and Megan J. MacGarvie. 2016. How important is US location for research in science? Review of Economics and Statistics, 98(2), 397-414.

Kantor, Shawn, and Alexander Whalley. 2014. Knowledge spillovers from research universities: Evidence from endowment value shocks. Review of Economics and Statistics, 96(1): 171-188.

Kennan, John and Walker, James R., 2011. The effect of expected income on individual migration decisions. Econometrica, 79(1): 211-251.

Kerr, William. 2013. U.S. high-skilled immigration, innovation, and entrepreneurship: Empirical approaches and evidence. NBER Working Paper No. 19377.

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(3): 473-508.

Kleibergen, Frank and Richard Paap. 2006. Generalized reduced rank tests using the singular value decomposition. Journal of Econometrics 133(1): 97-126.

Lee, Neil. 2014. Migrant and ethnic diversity, cities and innovation: Firm effects or city effects? Journal of Economic Geography, 15(4), 769-796.

Levin, Sharon G., Grant C. Black, Anne E. Winkler, and Paula E. Stephan. 2004. Differential employment patterns for citizens and non-citizens in science and engineering in the United States: Minting and competitive effects. Growth and Change, 35(4): 456-475.

Levin, Sharon G. and Paula E. Stephan. 1999. Are the foreign born a source of strength for U.S. science? Science, 285: 1213-1214.

Lewis, Ethan and Giovanni Peri. 2014. Immigration and the economy of cities and regions. NBER Working Paper No. 20428.

Liu, Shimeng. 2016. Agglomeration, urban wage premiums, and college majors. Journal of Regional Science, Forthcoming, DOI: 10.1111/jors.12309.

Maré, David C., Richard Fabling, and Steven Stillman. 2014. Innovation and the local workforce. Papers in Regional Science, 93(1): 183-201.

Molloy, Raven, Christopher L. Smith, Riccardo Trezzi, and Abigail Wozniak. 2016. Understanding declining fluidity in the U.S. labor market. Brookings Papers on Economic Activity, pp.183-237.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. Internal migration in the United States. Journal of Economic Perspectives, 25(3): 173-196.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2017. Job changing and the decline in long-distance migration in the United States. Demography, 54(2): 631-653.

Moretti, Enrico. 2004. Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics, 121: 175-212.

Moretti, Enrico. 2013. The New Geography of Jobs. New York: Houghton Mifflin Harcourt.

Moser, Petra, Alessandra Voena and Fabian Waldinger. 2014. German Jewish Émigrés and US Invention. American Economic Review, 104(10): 3222-55.

Nathan, Max. 2014. Same difference? Minority ethnic inventors, diversity and innovation in the UK. Journal of Economic Geography, 15(1), 129-168.

Nathan, Max and Neil Lee. 2013. Cultural diversity, innovation and entrepreneurship: Firm-level evidence from London. Economic Geography, 89: 367-394.

National Academies. 2010. Rising above the gathering storm revised: Rapidly approaching category 5. Washington, DC: National Academies Press.

Nelson, Glen R. 2008. Differential tuition by undergraduate major: Its use, amount, and impact at public research universities. Dissertation, University of Nebraska-Lincoln.

Niebuhr, Annekatrin. 2010. Migration and innovation: Does cultural diversity matter for regional R&D activity? Papers in Regional Science, 89(3): 563-585.

Orrenius, P.M. and Zavodny, M., 2015. Does immigration affect whether US natives major in science and engineering? Journal of Labor Economics, 33(S1), S79-S108.

Ottaviano Gianmarco I.P. and Giovanni Peri. 2012. Rethinking the effects of immigration on wages. Journal of the European Economic Association, 10(1): 152-197.

Ozgen, Ceren, Peter Nijkamp, and Jacques Poot. 2013. The impact of cultural diversity on firm innovation: Evidence from Dutch micro-data. IZA Journal of Migration, 2(1): 1-24.

Partridge, Mark D., Dan S. Rickman, Kamar Ali and M. Rose Olfert. 2009. Agglomeration spillovers and wage and housing cost gradients across the urban hierarchy. Journal of International Economics, 78(1): 126-140.

Partridge, Mark D., Dan S. Rickman, Kamar Ali and M. Rose Olfert. 2010. Recent spatial growth dynamics in wages and housing costs: Proximity to urban production externalities and consumer amenities. Regional Science and Urban Economics, 40(6): 440-452.

Partridge, Mark D., Dan S. Rickman, M. Rose Olfert, and Kamar Ali. 2012. Dwindling US internal migration: Evidence of spatial equilibrium or structural shifts in local labor markets? Regional Science and Urban Economics, 42(1): 375-388.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. STEM workers, H-1B visas, and productivity in US cities. Journal of Labor Economics, 33(S1), S225-S255.

Peri, Giovanni and Chad Sparber. 2009. Task specialization, immigration, and wages. American Economic Journal: Applied Economics, 1(3): 135-169.

Peri, Giovanni and Chad Sparber. 2011. Highly-educated immigrants and native occupational choice. Industrial Relations, 50(3): 385-411.

Ponds, Roderik, Frank van Oort, and Koen Frenken. 2010. Innovation, spillovers and university-industry collaboration: An extended knowledge production function approach. Journal of Economic Geography, 10(2): 231-255.

President's Council of Advisors on Science and Technology (PCAST). 2012. Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Washington, DC.

Price, Joshua. 2010. The effect of instructor race and gender on student persistence in STEM fields. Economics of Education Review 29(6): 901-910.

Ransom, Tyler, and John V. Winters. 2016. Do foreigners crowd natives out of STEM degrees and occupations? Evidence from the US Immigration Act of 1990. IZA Discussion Paper No. 9920.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. Integrated Public Use Microdata Series: Version 6.0 [dataset]. Minneapolis: University of Minnesota. http://doi.org/10.18128/D010.V6.0.

Shapiro, Jesse M. 2006. Smart cities: Quality of life, productivity, and the growth effects of human capital. Review of Economics and Statistics, 88: 324-335.

Simon, Curtis J. 1998. Human capital and metropolitan employment growth. Journal of Urban Economics, 43: 223-243.

Simon, Curtis J. and Clark Nardinelli. 2002. Human capital and the rise of American cities, 1900–1990. Regional Science and Urban Economics, 32: 59-96.

Simonen, Jaakko and Philip McCann. 2008. Firm innovation: The influence of R&D cooperation and the geography of human capital inputs. Journal of Urban Economics, 64(1): 146-154.

Sjoquist, David L. and John V. Winters. 2015a. State merit-aid programs and college major: A focus on STEM. Journal of Labor Economics, 33(4): 973-1006.

Sjoquist, David L. and John V. Winters. 2015b. The effect of Georgia's HOPE Scholarship on college major: A focus on STEM. IZA Journal of Labor Economics, 4(15): 1-29.

Stange, Kevin M. 2015. Differential pricing in undergraduate education: Effects on degree production by field. Journal of Policy Analysis and Management, 34(1): 107-135.

Stephan, Paula and Sharon Levin. 2001. Exceptional contributions to US science by the foreign-born and foreign-educated. Population Research and Policy Review, 20(1): 59-79.

Stinebrickner, Todd R. and Ralph Stinebrickner. 2014. A major in science? Initial beliefs and final outcomes for college major and dropout. Review of Economic Studies, 81: 426-472.

Stock, James, Jonathan Wright, and Motohiro Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics, 20(4): 518-29.

Stock, James H., and Motohiro Yogo. 2005. Testing for weak instruments in linear IV regression. In D. W. K. Andrews, & J. H. Stock (Eds.), Identification and inference for econometric models: A Festschrift in Honor of Thomas Rothenberg (pp. 80–108). Cambridge: Cambridge University Press.

Stuen, Eric, Ahmed Mobarak, and Keith Maskus. 2012. Skilled immigration and innovation: Evidence from enrollment fluctuations in U.S. doctoral programs. Economic Journal, 122: 1143-1176.

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

Winters, John V., 2011. Why are smart cities growing? Who moves and who stays. Journal of Regional Science, 51(2): 253-270.

Winters, John V. 2013. Human capital externalities and employment differences across metropolitan areas of the USA. Journal of Economic Geography, 13(5): 799-822.

Winters, John V. 2014a. STEM graduates, human capital externalities, and wages in the U.S. Regional Science and Urban Economics, 48: 190-198.

Winters, John V. 2014b. Foreign and native-born STEM graduates and innovation intensity in the United States. IZA Discussion Paper No. 8575.

Winters, John V, 2016. In-state college enrollment and later life location decisions. Oklahoma State University Working Paper No. 1702.

Winters, John V, 2017. Do higher college graduation rates increase local education levels? Papers in Regional Science, Forthcoming, DOI: 10.1111/pirs.12258.

Wright, Richard, Mark Ellis, and Matthew Townley. 2017. The matching of STEM degree holders with STEM occupations in large metropolitan labor markets in the United States. Economic Geography, 93(2): 185-201.

Zucker, Lynne G., and Michael R. Darby. 1996. Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. Proceedings of the National Academy of Sciences, 93(23), 12709-12716.

Zucker, Lynne G., and Michael R. Darby. 2007. Star scientists, innovation and regional and national immigration. NBER Working Paper No. 13547.

Figure 1: Bivariate Relationship between Patent Intensity and STEM Graduate Share for 2015

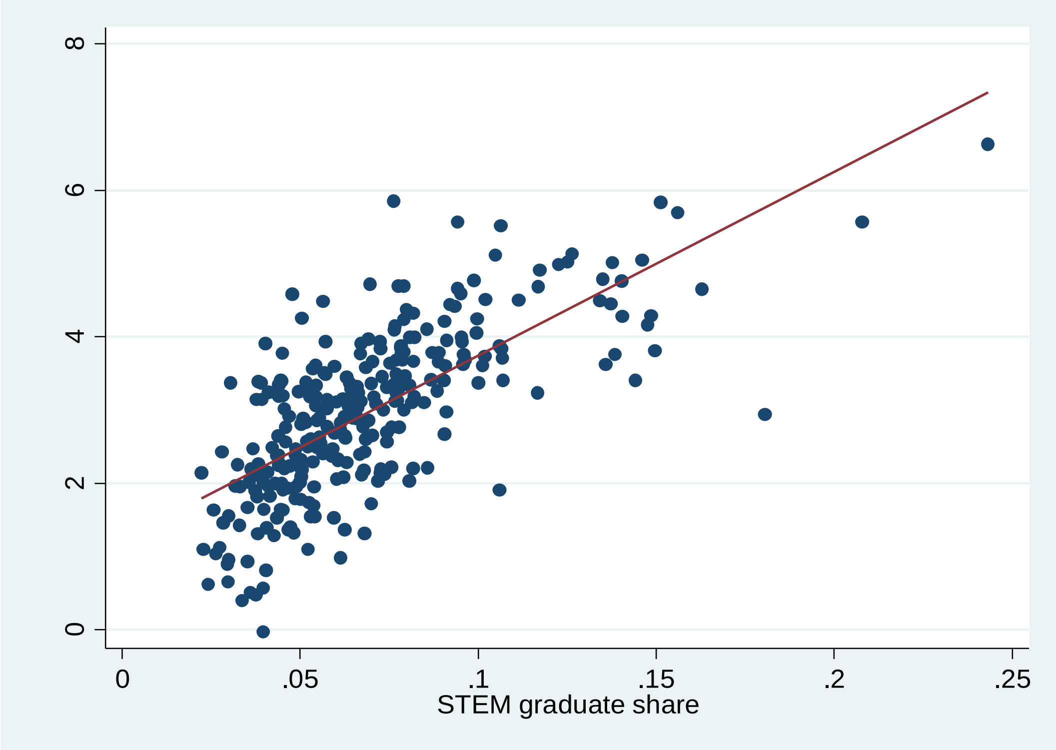

Figure 2: Map Illustrating MSA Log Patents per 100K Population in 2015

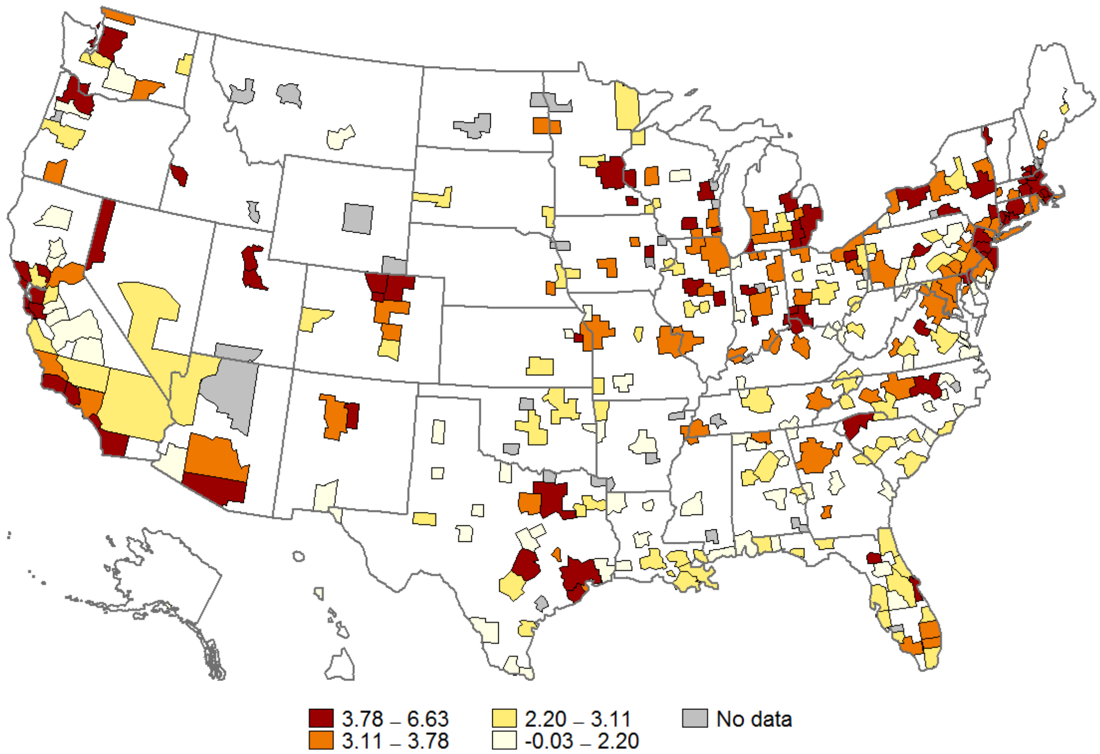

Figure 3: Map Illustrating MSA Change in Log Patents per 100K Population, 2009-2015

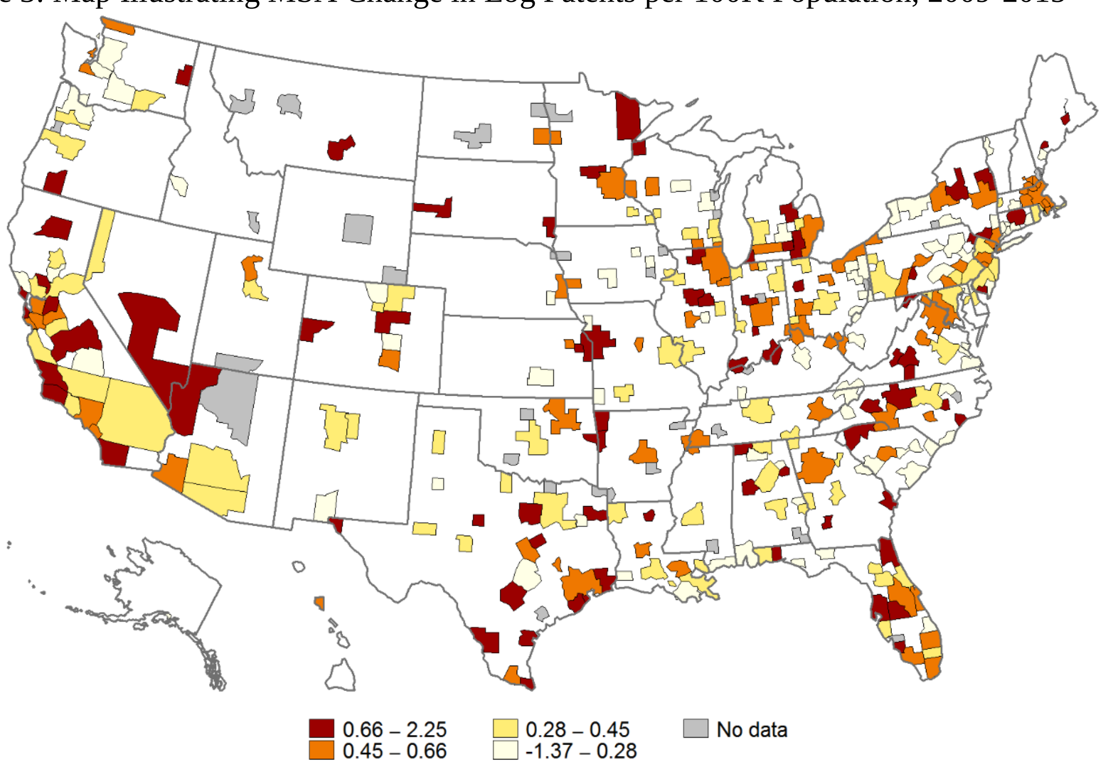

Figure 4: Mean MSA Patents per 100K Population by Year, 1990-2015

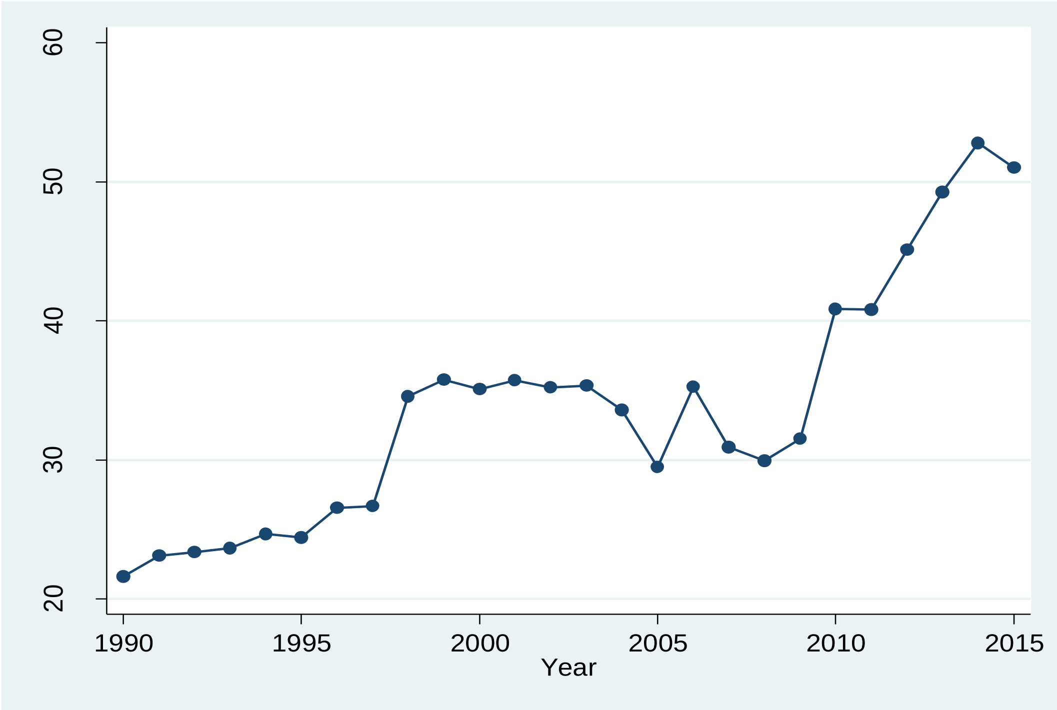

Table 1: Summary Statistics

	Mean	St. Dev.	Min	Max
Dependent variables				
Log of patents per 100K population, 2015	3.469	0.943	-0.030	6.633
Δ log patents per 100K population, 2009-2015	0.489	0.253	-1.375	2.254
Δ native STEM occupation employment share, 2009-2015	0.001	0.004	-0.020	0.022
Δ foreign STEM occupation employment share, 2009-2015	0.002	0.003	-0.020	0.017
Explanatory variables of interest				
STEM graduate population share, 2015	0.081	0.030	0.022	0.243
Native STEM graduate share, 2015	0.056	0.018	0.014	0.175
Foreign STEM graduate share, 2015	0.025	0.020	0.001	0.160
Non-STEM graduate share, 2015	0.248	0.051	0.112	0.396
Δ STEM graduate population share, 2009-2015	0.010	0.008	-0.039	0.068
Δ Native STEM graduate share, 2009-2015	0.005	0.006	-0.025	0.054
Δ Foreign STEM graduate share, 2009-2015	0.005	0.005	-0.014	0.036
Δ Non-STEM graduate share, 2009-2015	0.017	0.012	-0.042	0.078
Instruments				
Predicted young native STEM graduate share, 2009-2015	0.072	0.010	0.028	0.118
Predicted Δ in foreign STEM occupation share, 2009-2015	0.004	0.003	0.000	0.023
Control variables				
Division dummy - Middle Atlantic	0.152	0.359	0	1
Division dummy - East North Central	0.145	0.353	0	1
Division dummy - West North Central	0.053	0.225	0	1
Division dummy - South Atlantic	0.193	0.396	0	1
Division dummy - East South Central	0.043	0.203	0	1
Division dummy - West South Central	0.114	0.318	0	1
Division dummy - Mountain	0.066	0.248	0	1
Division dummy - Pacific	0.185	0.389	0	1
Log metropolitan area population, 2009	14.338	1.153	11.511	16.097
Share of employment in federal government, 2009	0.022	0.021	0.003	0.189
Share of employment in state and local government, 2009	0.134	0.042	0.000	0.385
Share of employment in natural resources and mining, 2009	0.011	0.027	0.000	0.261
Share of employment in construction, 2009	0.046	0.013	0.015	0.119
Share of employment in manufacturing, 2009	0.084	0.040	0.010	0.413
Share of employment in transportation and utilities, 2009	0.190	0.025	0.125	0.305
Share of employment in information, 2009	0.023	0.011	0.000	0.061
Share of employment in financial activities, 2009	0.062	0.020	0.022	0.173
Share of employment in professional and business services, 2009	0.135	0.033	0.037	0.226
Share of employment in education and health services, 2009	0.146	0.034	0.068	0.443
Unemployment rate, 2009	0.083	0.020	0.025	0.146
Mean age of adult labor force, 2009	43.986	0.979	40.757	48.741
Log mean firm size (# of Employees), 2009	8.511	0.219	7.875	9.548
Research universities per 100K population, 2009	0.099	0.102	0.000	0.909
Log university research expenditure per 100K population, 2009	15.176	3.514	0.000	19.932
Distance to metro area with population > 250K (in year 2000)	6.911	28.154	0.000	384.431
Incremental distance to metro area with pop > 500K (in 2000)	12.053	57.832	0.000	1430.779
Incremental distance to metro area with pop > 1500K (in 2000)	48.611	165.187	0.000	2393.967
Log patents per 100K average during 2007-2008	2.975	0.921	-0.476	6.028
Log patents per 100K average during 1998-2006	3.187	0.828	-0.417	5.945
Log patents per 100K average during 1990-1997	2.950	0.725	-0.308	4.915

Note: analytical sample includes 288 MSA/PMSA observations.

Table 2: College Graduate Shares in 2015 for the Top 25 STEM Graduate Share Metropolitan Areas

MSA/PMSA Name	STEM graduate population share	Native STEM graduate share	Foreign STEM graduate share	Non-STEM graduate share
San Jose, CA PMSA	0.243	0.083	0.160	0.251
Boulder-Longmont, CO PMSA	0.208	0.175	0.033	0.345
Columbia, MO MSA	0.181	0.136	0.045	0.325
Middlesex-Somerset-Hunterdon, NJ PMSA	0.163	0.062	0.101	0.295
San Francisco, CA PMSA	0.156	0.089	0.068	0.374
Rochester, MN MSA	0.151	0.113	0.039	0.305
State College, PA MSA	0.150	0.100	0.050	0.265
Champaign-Urbana, IL MSA	0.149	0.080	0.069	0.320
Yolo, CA PMSA	0.148	0.096	0.052	0.267
Seattle-Bellevue-Everett, WA PMSA	0.146	0.088	0.059	0.306
Bryan-College Station, TX MSA	0.144	0.091	0.053	0.262
Gainesville, FL MSA	0.141	0.094	0.046	0.299
Fort Collins-Loveland, CO MSA	0.140	0.122	0.018	0.321
Huntsville, AL MSA	0.138	0.124	0.015	0.240
Oakland, CA PMSA	0.138	0.062	0.076	0.292
Madison, WI MSA	0.137	0.109	0.028	0.358
Washington, DC-MD-VA-WV PMSA	0.136	0.084	0.052	0.353
Raleigh-Durham-Chapel Hill, NC MSA	0.135	0.098	0.037	0.299
Lafayette, IN MSA	0.134	0.087	0.048	0.212
Burlington, VT NECMA	0.126	0.113	0.013	0.320
Ann Arbor, MI PMSA	0.125	0.096	0.029	0.278
Trenton, NJ PMSA	0.123	0.065	0.057	0.263
Austin-San Marcos, TX MSA	0.117	0.084	0.033	0.301
Boston-Worcester-Lawr., MA-NH NECMA	0.117	0.080	0.037	0.308
Jersey City, NJ PMSA	0.117	0.041	0.076	0.275

Table 3: Pearson Correlation Coefficients for STEM and Non-STEM Human Capital Variables

	Native STEM Graduate Share	Foreign STEM Graduate Share
2015 Cross-Section		
Foreign STEM graduate share	0.249	
Non-STEM graduate share	0.589	0.484
2009 2015 Time Differences		
Foreign STEM graduate share	0.056	
Non-STEM graduate share	0.037	-0.055

Table 4: OLS 2015 Cross-Section Patent Intensity Results

	(1)	(2)	(3)
STEM graduate population share	3.443		
	(1.192)***		
Native STEM graduate share		3.252	2.515
		(1.286)**	(1.394)*
Foreign STEM graduate share		3.565	3.433
		(1.368)**	(1.351)**
Non-STEM graduate share			0.847
			(0.596)

Notes: Dependent variable is the log of patents per 100K population in 2015. All regressions include control variables listed in Table 1. Standard errors in parentheses are robust to heteroskedasticity and clustered by state. *Statistically significantly different from zero at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 5: 2009-2015 Time-Differenced OLS Patent Intensity Results

	(1)	(2)	(3)
Δ STEM graduate population share	-0.252		
	(1.863)		
Δ Native STEM graduate share		-2.693	-2.687
		(2.435)	(2.441)
Δ Foreign STEM graduate share		4.679	4.652
		(2.637)*	(2.651)*
Δ Non-STEM graduate share			-0.180
			(0.920)

Notes: Dependent variable is the change in the log of patents per 100K population, 2009-2015. All regressions include control variables listed in Table 1. Standard errors in parentheses are robust to heteroskedasticity and clustered by state.

*Statistically significantly different from zero at 10% level.

Table 6: 2009-2015 Time-Differenced 2SLS Patent Intensity Results

	(1)	(2)	(3)	(4)
A. First-Stage Results				
Endogenous Variable: Δ Native STEM graduate share				
Predicted young native STEM graduate share	0.234		0.236	0.238
	(0.067)***		(0.067)***	(0.069)***
Predicted Δ in foreign STEM occupation share			0.054	0.053
			(0.175)	(0.175)
F-statistic	12.08		6.22	5.95
Endogenous Variable: Δ Foreign STEM graduate share				
Predicted young native STEM graduate share			0.124	0.131
			(0.032)***	(0.032)***
Predicted Δ in foreign STEM occupation share		0.920	0.954	0.951
		(0.190)***	(0.189)***	(0.187)***
F-statistic		23.35	16.58	18.14
Kleibergen-Paap Underidentification Statistic	13.638	5.907	10.315	10.376
Underidentification p-value	0.0002	0.015	0.001	0.001
Kleibergen-Paap Weak Identification Statistic	15.258	20.986	5.543	5.648
Weak Identification Critical Value for 10% max size	16.38	16.38	7.03	7.03
Weak Identification Critical Value for 15% max size	8.96	8.96	4.58	4.58
Weak Identification Critical Value for 20% max size	6.66	6.66	3.95	3.95
Weak Identification Critical Value for 25% max size	5.53	5.53	3.63	3.63
B. Second-Stage Results				
Δ Native STEM graduate share	-5.161		-9.467	-9.505
	(10.733)		(11.471)	(11.683)
Δ Foreign STEM graduate share		11.272	11.180	11.185
		(5.981)*	(5.550)**	(5.624)**
Δ Non-STEM graduate share				0.028
				(0.988)

Notes: The second-stage dependent variable is the change in log of patents per 100K population, 2009-2015. All regressions include control variables listed in Table 1. Standard errors in parentheses are robust to heteroskedasticity and clustered by state.

*Statistically significantly different from zero at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 7: 2009-2015 Time-Differenced 2SLS Effects on STEM Occupation Employment Shares

	(1)	(2)	(3)
A. Native STEM Occupation Second-Stage Results (First-Stage is identical to Table 6)			
Δ Native STEM graduate share	0.411		0.441
	(0.121)***		(0.133)***
Δ Foreign STEM graduate share		-0.082	-0.078
		(0.084)	(0.099)
B. Foreign STEM Occupation Second-Stage Results (First-Stage is identical to Table 6)			
Δ Native STEM graduate share	0.167		0.018
	(0.093)*		(0.077)
Δ Foreign STEM graduate share		0.384	0.384
		(0.073)***	(0.073)***

Notes: The second-stage dependent variables are the changes in the STEM occupation employment share during 2009-2015 for natives in Panel A and foreigners in Panel B. All regressions include control variables listed in Table 1. Standard errors in parentheses are robust to heteroskedasticity and clustered by state.

*Statistically significantly different from zero at 10% level; ***Significant at 1% level.

Table 8: Time-Differenced 2SLS Patenting Sensitivity Analysis

	(1)	(2)	(3)
A. Controlling for 2009 Log Patents per 100K			
Δ Native STEM graduate share	-7.663		-12.615
	(9.549)		(10.318)
Δ Foreign STEM graduate share		12.691	12.503
		(6.519)*	(5.983)**
B. 2006-2015 Dependent Variable			
Δ Native STEM graduate share	0.663		-5.952
	(10.327)		(9.204)
Δ Foreign STEM graduate share		15.561	15.579
		(8.970)*	(8.650)*
C. 2009/10-2014/15 Dependent Variable			
Δ Native STEM graduate share	-12.175		-17.043
	(11.297)		(13.063)
Δ Foreign STEM graduate share		12.805	12.639
		(6.703)*	(6.435)**
D. Native STEM Instrument Excluding Asian Americans			
Δ Native STEM graduate share	-20.392		-18.578
	(15.339)		(14.802)
Δ Foreign STEM graduate share		11.272	11.091
		(5.981)*	(5.734)*
E. Controlling for 2009 STEM Occupation Share of MSA Employment			
Δ Native STEM graduate share	-7.811		-9.420
	(11.561)		(11.641)
Δ Foreign STEM graduate share		11.585	11.278
		(6.333)*	(5.643)**
F. Controlling for 2009-2015 Δ STEM Occupation Share of MSA Employment			
Kleibergen-Paap Weak Identification Statistic	9.539	26.707	4.345
Δ Native STEM graduate share	-6.906		-10.155
	(15.573)		(14.920)
Δ Foreign STEM graduate share		12.795	10.722
		(6.383)**	(6.681)
G. Excluding Top 5 STEM Graduate MSAs in 2015			
Kleibergen-Paap Weak Identification Statistic	13.030	12.029	3.791
Δ Native STEM graduate share	-9.810		-12.921
	(13.660)		(13.309)
Δ Foreign STEM graduate share		11.586	7.801
		(7.890)	(8.260)
H. Including State Dummies			
Kleibergen-Paap Weak Identification Statistic	7.109	14.340	1.970
Δ Native STEM graduate share	-17.717		-24.644
	(17.456)		(25.351)
Δ Foreign STEM graduate share		9.152	17.710
		(7.834)	(16.619)

Notes: All regressions include control variables listed in Table 1. Standard errors in parentheses are robust to heteroskedasticity and clustered by state. First-stage statistics for Panels A-E are very similar to those in Table 6 and are not reported to conserve space.

*Statistically significantly different from zero at 10% level; **Significant at 5% level.

Appendix Table A1: List of STEM Majors and ACS Codes

ACS Code and Description			
1103	Animal Sciences	2504	Mechanical Engineering Related Technologies
1104	Food Science	2599	Miscellaneous Engineering Technologies
1105	Plant Science and Agronomy	3600	Biology
1106	Soil Science	3601	Biochemical Sciences
1301	Environmental Science	3602	Botany
1302	Forestry	3603	Molecular Biology
2001	Communication Technologies	3604	Ecology
2100	Computer and Information Systems	3605	Genetics
2101	Computer Programming and Data Processing	3606	Microbiology
2102	Computer Science	3607	Pharmacology
2105	Information Sciences	3608	Physiology
2106	Computer Information Mgmt. & Security	3609	Zoology
2107	Computer Networking & Telecommunications	3611	Neuroscience
2400	General Engineering	3699	Miscellaneous Biology
2401	Aerospace Engineering	3700	Mathematics
2402	Biological Engineering	3701	Applied Mathematics
2403	Architectural Engineering	3702	Statistics and Decision Science
2404	Biomedical Engineering	3801	Military Technologies
2405	Chemical Engineering	4002	Nutrition Sciences
2406	Civil Engineering	4003	Neuroscience
2407	Computer Engineering	4005	Mathematics and Computer Science
2408	Electrical Engineering	4006	Cognitive Science and Biopsychology
2409	Engineering Mechanics, Physics, & Science	5000	Physical Sciences
2410	Environmental Engineering	5001	Astronomy and Astrophysics
2411	Geological and Geophysical Engineering	5002	Atmospheric Sciences and Meteorology
2412	Industrial and Manufacturing Engineering	5003	Chemistry
2413	Materials Engineering and Materials Science	5004	Geology and Earth Science
2414	Mechanical Engineering	5005	Geosciences
2415	Metallurgical Engineering	5006	Oceanography
2416	Mining and Mineral Engineering	5007	Physics
2417	Naval Architecture and Marine Engineering	5008	Materials Science
2418	Nuclear Engineering	5098	Multi-disciplinary or General Science
2419	Petroleum Engineering	5102	Nuclear, Industrial Radiology, & Biol. Tech.
2499	Miscellaneous Engineering	5901	Transportation Sciences and Technologies
2500	Engineering Technologies	6106	Health and Medical Preparatory Programs
2501	Engineering and Industrial Management	6108	Pharmacy, Pharmaceutical Sciences, & Admin.
2502	Electrical Engineering Technology	6202	Actuarial Science
2503	Industrial Production Technologies	6212	Management Information Systems & Statistics

Appendix Table A2: Mean STEM and Non-STEM Degree Rates for Young Natives by Ethnic Group and Sex

Race Group	National Population Share	Male STEM Share	Female STEM Share	Male Non-STEM Share	Female Non-STEM Share
White Dutch	0.011	0.120	0.066	0.268	0.437
White English	0.060	0.107	0.071	0.258	0.411
White French	0.021	0.098	0.054	0.224	0.350
White German	0.179	0.112	0.073	0.237	0.386
White Irish	0.079	0.084	0.061	0.249	0.388
White Italian	0.041	0.102	0.064	0.318	0.465
White Nordic	0.024	0.117	0.084	0.269	0.441
White Scottish	0.019	0.109	0.071	0.259	0.413
White Other West Europe	0.015	0.117	0.062	0.278	0.446
White Polish	0.022	0.119	0.081	0.319	0.472
White Other East Europe	0.023	0.155	0.086	0.341	0.503
White Other	0.204	0.085	0.049	0.174	0.274
Black/African American	0.147	0.031	0.031	0.107	0.184
Mexican Hispanic	0.081	0.031	0.019	0.097	0.153
Puerto Rican Hispanic	0.014	0.031	0.024	0.118	0.178
Cuban Hispanic	0.003	0.070	0.055	0.199	0.366
Other Hispanic	0.021	0.059	0.039	0.170	0.278
American Indian/Alaska Native	0.010	0.034	0.024	0.063	0.091
Asian Indian	0.003	0.352	0.329	0.389	0.510
Chinese	0.005	0.304	0.295	0.415	0.498
Filipino	0.005	0.131	0.117	0.277	0.416
Japanese	0.002	0.218	0.183	0.301	0.495
Korean	0.002	0.236	0.219	0.418	0.481
Other Asian/Pacific Islander	0.008	0.145	0.119	0.197	0.322
All Other	0.002	0.103	0.058	0.227	0.428

Note: All variables are based on people born in the U.S. during years 1985-1990. These are the people used to construct the instrument for native STEM. National population share is the share of each group relative to the total population of persons born in the U.S. during 1985-1990. STEM and Non-STEM shares are unconditional on education level and are measured based on persons ages 24-29 in the 2014 ACS consistent with the native STEM instrument.

Appendix Table A3: Control Variable Second-Stage Results for Main 2SLS Specification (Table 6 Column 3)

	Coefficient	St. Error
Division dummy - Middle Atlantic	-0.073	(0.076)
Division dummy - East North Central	-0.040	(0.076)
Division dummy - West North Central	-0.047	(0.088)
Division dummy - South Atlantic	-0.150	(0.085)*
Division dummy - East South Central	-0.144	(0.109)
Division dummy - West South Central	0.074	(0.083)
Division dummy - Mountain	0.013	(0.122)
Division dummy - Pacific	0.048	(0.081)
Log metropolitan area population, 2009	-0.027	(0.029)
Share of employment in federal government, 2009	0.842	(1.061)
Share of employment in state and local government, 2009	-0.334	(0.548)
Share of employment in natural resources and mining, 2009	-1.146	(0.731)
Share of employment in construction, 2009	-0.538	(1.574)
Share of employment in manufacturing, 2009	0.038	(0.582)
Share of employment in transportation and utilities, 2009	0.130	(0.809)
Share of employment in information, 2009	-1.568	(2.285)
Share of employment in financial activities, 2009	2.232	(1.222)*
Share of employment in professional and business services, 2009	1.042	(0.965)
Share of employment in education and health services, 2009	-0.518	(0.864)
Unemployment rate, 2009	2.306	(0.869)***
Mean age of adult labor force, 2009	0.018	(0.024)
Log mean firm size (# of Employees), 2009	0.143	(0.091)
Research universities per 100K population, 2009	0.388	(0.179)**
Log university research expenditure per 100K population, 2009	0.002	(0.006)
Distance to metro area with population > 250K (in year 2000)	0.001	(0.001)
Incremental distance to metro area with pop > 500K (in year 2000)	-0.0004	(0.0002)***
Incremental distance to metro area with pop > 1500K (in year 2000)	-0.00004	(0.0001)
Log patents per 100K average during 2007-2008	0.044	(0.100)
Log patents per 100K average during 1998-2006	-0.232	(0.179)
Log patents per 100K average during 1990-1997	0.064	(0.103)
P-value for joint significance of division dummies	0.095	
P-value for joint significance of employment shares	0.003	
P-value for joint significance of lagged patent variables	0.007	

Notes: The second-stage dependent variable is the change in log of patents per 100K population. See Tables 6 Column 3 for more details. Standard errors in parentheses are robust to heteroskedasticity and clustered by state. *Statistically significantly different from zero at 10% level; **Significant at 5% level; ***Significant at 1% level.