NBER WORKING PAPER SERIES

IMMIGRATION, INNOVATION, AND GROWTH

Konrad B. Burchardi
Thomas Chaney
Tarek Alexander Hassan
Lisa Tarquinio
Stephen J. Terry

Working Paper 27075
http://www.nber.org/papers/w27075

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2020

We are grateful to Ran Abramitzky, Daron Acemoglu, David Autor, David Atkin, Murat Celik, William Kerr, Klaus Desmet, Ray Fisman, Lisa Kahn, Kevin Lang, Ethan Lewis, Ömer Özak, Michael Peters, Chris Tonetti, and Pascual Restrepo for helpful comments. We also thank seminar participants at AEA Annual Meeting, the University of California San Diego, University of Toronto, Imperial College London, the Federal Reserve banks of Boston, Philadelphia, and Richmond, Southern Methodist University, University of Toronto, Goethe University Frankfurt, Duke, Princeton, the University of Chicago, London Business School, the Wharton Conference on Migration, and the NBER EFG, EFEG, and PRMP group meetings. Chaney is grateful for financial support from ERC grant N 337272–FiNet. All remaining mistakes are our own. This paper is dedicated to the memory of Dr. Sherif A. Hassan. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Immigration, Innovation, and Growth
Konrad B. Burchardi, Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J. Terry
NBER Working Paper No. 27075
May 2020
JEL No. J61,O31,O40

ABSTRACT

We show a causal impact of immigration on innovation and dynamism in US counties. To identify the causal impact of immigration, we use 130 years of detailed data on migrations from foreign countries to US counties to isolate quasi-random variation in the ancestry composition of US counties that results purely from the interaction of two historical forces: (i) changes over time in the relative attractiveness of different destinations within the US to the average migrant arriving at the time and (ii) the staggered timing of the arrival of migrants from different origin countries. We then use this plausibly exogenous variation in ancestry composition to predict the total number of migrants flowing into each US county in recent decades. We show four main results. First, immigration has a positive impact on innovation, measured by the patenting of local firms. Second, immigration has a positive impact on measures of local economic dynamism. Third, the positive impact of immigration on innovation percolates over space, but spatial spillovers quickly die out with distance. Fourth, the impact of immigration on innovation is stronger for more educated migrants.

Konrad B. Burchardi
Institute for International Economic Studies
Stockholm University
SE-106 91 Stockholm
Sweden
konrad.burchardi@iies.su.se

Thomas Chaney
Sciences Po
28 rue des Saints Peres
75005 Paris
France
thomas.chaney@gmail.com

Tarek Alexander Hassan
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
thassan@bu.edu

Lisa Tarquinio
Department of Economics
Boston University
270 Bay State Road
Boston, MA
ltarq@bu.edu

Stephen J. Terry
Department of Economics
Boston University
Boston, MA 02215
stephenjamesterry@gmail.com

1 Introduction

Does immigration cause more or less innovation and economic dynamism? In this paper, we answer this question in the context of international migration to the US over the last three decades. We find a positive causal impact of immigration on both innovation and economic dynamism at the county level.

Theories crafted to study endogenous growth, economic geography, and dynamism suggest a role for immigrants in driving local economic outcomes. When immigrants bring ideas, skills, and effort into the research process, or when they stimulate demand for new inventions, they will also stimulate growth according to the logic of a range of endogenous growth models (Romer, 1990; Jones, 1995). In the presence of frictions on mobility, trade, or idea flows, a range of models of regional growth suggest immigrants should have local, not just aggregate, effects on innovation and wages (Desmet et al., 2018; Peters, 2019). Innovations or growth caused by immigrants but embodied in new firms or creative destruction should also affect local measures of economic dynamism (Karahan et al., 2016; Hopenhayn et al., 2018).

In contrast to these predictions of canonical theory, fierce political controversies surround the economic contribution of migrants: Are the new arrivals draining resources of their host communities and stifling innovation and economic dynamism?

A rigorous quantification of the causal impact of immigration on innovation and dynamism has proven difficult. The reason is that migrants do not allocate randomly across space, but instead are likely to choose destinations that offer the best prospects for them and their families. In particular, migrants arriving in the US might select into regions that are more innovative, economically dynamic, and fast-growing, creating a spurious correlation between local immigration, local innovation, and local economic dynamism.

In this paper, we propose a formal identification strategy that allows us to identify the causal impact of migration on local innovation and dynamism. To do so, we use 130 years of detailed data on immigration from foreign countries to US counties. Our identification strategy combines a set of instruments for the pre-existing ethnic composition of US counties (Burchardi et al., 2019) with a version of the canonical shift-share approach (Bartik, 1991; Katz and Murphy, 1992; Card, 2001) to construct a valid instrument for immigration into each US county in the last 30 years. In a first step, we isolate plausibly exogenous variation in the number of residents of a US county with ancestry from each foreign country, following Burchardi et al. (2019). In a second step, we use these exogenous components of pre-existing ancestry shares to predict where recent migrants will settle within the US, using a shift-share instrument. Doing so, we guard against the potential critique that where migrants settle within the US, both in recent decades (the distribution of immigrants) and in the more distant past (the distribution of ancestry), may be correlated with unobserved factors that also affect local innovation and dynamism.

In our first step, we use the interaction of time-series variation in the relative attractiveness of different destinations within the US with the staggered timing of arrival of migrants from different origins to isolate quasi-random variation in the ancestry composition of US counties. Implicitly, we assume historical migration patterns are in part driven by (*i*) a push factor, causing emigration from a given foreign country to the entire US, and (*ii*) an economic pull factor, attracting migrants from all origins to a given US county at a given point in time. To further ensure our predicted historical migration is not contaminated by endogenous unobserved factors, we carefully leave out large population groups when predicting ancestry. In particular, because our focus is on immigration to the US after 1975, primarily originating from non-European countries, we use the historical location choices of European migrants to predict where non-European migrants settled prior to 1975. In other words, we predict that US counties that were attractive to migrants from Europe in a period when a large number of migrants from a given non-European origin country were arriving in the US will receive a large number of migrants from that (non-European) origin country. Iterating this procedure over 100 years, we isolate quasi-random variation in the distribution of ancestry across US counties in 1975.

In our second step, we then use this predicted pre-existing distribution of ancestry to predict where new migrants arriving in the US after 1975 will settle. Implicitly, we assume migrants' choice of destination is also driven by a social pull factor, such that new migrants will tend to settle in locations with a large pre-existing community from the same ethnic background. So, if a large community with ancestry from origin country o already resides in destination county d, and many migrants from o arrive in the US, we predict a large inflow of migrants from o to d. Summing over all possible origin countries, we are then able to predict the total number of migrants flowing into different US counties at each point in time post 1975. This predicted immigration shock is plausibly orthogonal to any origin-destination-specific factor that may make a destination US county more innovative and dynamic after 1975.

Finally, to further guard against any lingering concerns about identification, we estimate the impact of plausibly exogenous variations in immigration on *changes* in local innovation, dynamism, and growth, not on levels. In many specifications, we are even able to include county fixed effects, thus controlling for any county-specific trend in innovation.

This formal identification strategy allows us to reach four main conclusions.

First, we find a strong and significant causal impact of immigration on the number of patents filed per person: on average, the arrival of 10,000 additional immigrants increases the flow of patents over a five-year period by one patent per 100,000 people. Put differently, a one standard deviation increase in the number of migrants (about 12,000 migrants) increases the flow of patents by 27% relative to its mean.

Second, we find a strong and significant causal impact of immigration on measures of economic dynamism and growth at the local level. For our measures of economic dynamism, or creative destruction, we use several variables, each shedding light on different aspects of economic dynamism: a one standard deviation increase in local immigration increases the job creation rate by 7%, the job destruction rate by 11%, the job growth skewness by 3%, and local wages by 3%, all expressed as changes relative to their mean. The significant increase in local wages suggests immigration not only affects innovation and creative destruction, but also the overall level of economic growth.

Third, we find evidence that the positive effect of immigration on innovation and growth diffuses over space, but this spatial diffusion dies out quickly with distance. For instance, if more migrants settle in counties near d, innovation in d increases significantly. However, this spillover effect of immigration to nearby counties decays rapidly with distance: compared to the direct effect of immigration in a county, the indirect effect is 30% smaller for immigration 100km away (60 miles), 80% smaller at 250km (150 miles), and statistically indistinguishable from zero beyond 500km (300 miles).

Fourth, we find the positive effect of immigration on innovation and growth is significantly stronger for more educated migrants. We are able to reach this conclusion because our identification strategy allows us to construct separate instruments for migrations from each origin to each destination at each point in time. This versatility is one of the strengths of our identification strategy and makes it potentially applicable in a range of other contexts. To separately identify the impact of the total number of incoming migrants from that of their education level, we leverage the fact that the level of education of migrants varies dramatically across countries of origin and over time. For example, Japanese immigrants, on average, have about twice the number of years of schooling as those from Guatemala, whereas the education levels of Mexican arrivals increased by about 30% during our sample period. We find large heterogeneity in the impact of immigration on innovation as we exogenously vary the education level of migrants. For instance, an inflow of relatively uneducated migrants (in the bottom third of the distribution of years of schooling among incoming migrants) has almost no effect on local innovation, whereas the increase in innovation induced by highly educated migrants (in the top third) is an order of magnitude larger than for the average migrant.

Related Literature. Our paper contributes to several strands of the literature.

First, standard theories of endogenous growth predict strong positive impacts of overall population growth on economic growth and innovation (Romer, 1990), with the nature of these effects depending upon the details of the technology for producing ideas and the horizon of analysis (Jones, 1995, 1999; Peretto, 1998; Young, 1998; Laincz and Peretto, 2006; Bloom et al., 2017). Our empirical work can be thought of broadly as a reduced-form test of these predictions, in the sense that immigration constitutes a large part of regional population growth in the US. (We show an illustrative example of such a model in Appendix B.) However, we also differ from this literature because of our focus on the local, rather than the aggregate, effects of immigration on innovation and growth. In this sense, our evidence relates more closely to a burgeoning set of theories of the spatial distribution of economic growth, which also tend to predict a positive impact of immigration on *local* innovation and economic dynamism. The exact nature of the local effects of immigration differ in such models based on the type and extent of frictions on mobility, trade, and idea diffusion. Some theories emphasize migration restrictions and local market size (Desmet et al., 2018), others tease out subtle distinctions between the long-run and short-run effects of migration (Peters, 2019), and still others link migration frictions to traditional gravity relationships across regions (Monte et al., 2018). Rather than sharply distinguishing between these approaches, our empirical evidence offers a reduced-form test of common predictions of this class of models, as well as a useful empirical methodology to identify and discipline key parameters in those models.

Second, a long tradition studies economic dynamism or creative destruction using a combination of theory and empirics. If innovations appear through creative destruction, Schumpeterian growth models link immigrants' impact on innovation to churn and gross flows (Aghion and Howitt, 1992; Grossman and Helpman, 1991; Klette and Kortum, 2004), a prediction we test at the local level. In the US, dynamism has declined recently, a pattern emphasized by multiple recent papers using rich empirical evidence (Decker et al., 2014; Hathaway and Litan, 2014; Alon et al., 2018). We bring new causal evidence to bear to this line of work. Theoretical explanations for declining dynamism related to knowledge diffusion, IT developments, markups, or population growth have been proposed by a burgeoning set of papers (Akcigit and Ates, 2019; Aghion et al., 2019; Gordon, 2012; Karahan et al., 2016; Hopenhayn et al., 2018; Walsh, 2019). Our paper suggests immigration may also be a local driver of dynamism.

Finally, we contribute to a growing empirical literature on the link between immigration, innovation, and technology adoption. Different branches of this literature have documented large contributions of high-skilled immigrants to innovation in the US (Kerr and Lincoln, 2010; Hunt and Gauthier-Loiselle, 2010; Akcigit et al., 2017), spillovers from the arrival of high-skilled scientists and inventors on the productivity of their American peers (Borjas and Doran, 2012; Moser et al., 2014; Bernstein et al., 2018), and the contribution of migrants to the diffusion of knowledge across borders (Kerr, 2008).[1] Lewis (2011) and Lafortune et al. (2019) study the effect of immigration on local technology adoption, whereas Tabellini (2018) shows positive effects of immigration on output and employment. Khanna and Lee (2018) show a positive association between high-skilled migration and firm-level measures of dynamism. Many of these studies use variants of the canonical shift-share instrument (Card, 2001) that takes pre-existing ancestry shares as given (exogenous). Consistent with our intuition linking local innovation to the endogenous presence of particular ethnicities in the local population, this approach has recently been shown to lead to bias and over-rejection in a number of different contexts (Borusyak et al., 2018; Goldsmith-Pinkham et al., 2018; Adão et al., 2019). We contribute to this literature by isolating exogenous variation in the pre-existing spatial distribution of ancestry and using this variation to construct plausibly exogenous immigration shocks to US counties in recent decades.

Closely related to our own work, Sequeira et al. (2020) also develop an alternative to the canonical shift-share approach using the gradual expansion of railways across the US for identification. Consistent with our finding that a positive immigration shock increases local economic dynamism in the short-run (over 5 or 10 years), they document positive long-term effects of European immigration to the US 1850-1920 on local economic development that persists to the present day.

The remainder of this paper is structured as follows. Section 2 introduces our data. Section 3 lays out our strategy for identification and isolates quasi-random immigration shocks to US counties. Section 4 estimates the causal effect of immigration on innovation and economic dynamism. Section 5 tests for geographic spillovers in the effect of immigration on innovation and disentangles the impact of high-skilled from that of low-skilled migration. Section 6 concludes.

[1] Hanson (2009, 2010) and Lewis (2013) provide early surveys. Lewis and Peri (2015) and Abramitzky and Boustan (2017) give an overview of the broader literature on the effect of immigration on regional economies.

2 Data

We collect detailed data on migration, ancestry, the education level of migrants, patents issued, and measures of dynamism of local firms and local labor markets. Below is a description of our data sources and the construction of our main variables. Further details on the construction and sources of the data are given in Appendix A.

Immigration and Ancestry. Following Burchardi et al. (2019), our immigration and ancestry data are constructed from the individual files of the Integrated Public Use Microdata Series (IPUMS) samples of the 1880, 1900, 1910, 1920, 1930, 1970, 1980, 1990, and 2000 waves of the US census, and the 2006-2010 five-year sample of the American Community Survey.[2] We weigh observations using the personal weights provided by these data sources. Appendix A.1 gives details on specific samples and weights used.

Throughout the paper, we use $t-1$ and t to denote the end years of consecutive five-year periods,[3] o for the foreign country of origin, and d for the US destination county. We construct the number of migrants from origin o to destination d at time t, $I^t_{o,d}$, as the number of respondents born in o who live in d in a given census year and emigrated to the US between $t-1$ and t. The exception to this rule is the 1880 census (the first in our sample), which did not record the year of immigration. The variable $I^{1880}_{o,d}$ instead measures the number of residents who were either born in o or whose parents were born in o, thus covering the two generations of immigrants arriving prior to 1880.[4] Since 1980, respondents have also been asked about their primary ancestry in both the US Census and the American Community Survey, with the option to provide multiple answers. Ancestry $A^t_{o,d}$ corresponds to the number of individuals residing in d at time t who report o as their first ancestry. Note this measure captures self-reported ancestry and is thus subject to (endogenous) recall.[5]

The respondents' residence is recorded at the level of historic counties, and at the level of historic county groups or PUMAs from 1970 onwards. Whenever necessary, we use contemporaneous population weights to transition data from the historic county group or PUMA level to the historic county, and then use area weights to transition data from the historic county

[2]We cannot use data from the 1940, 1950, and 1960 censuses, because these censuses did not collect information on the year of immigration. The original 1890 census files were lost in a fire.

[3]Due to variation in the question regarding the year of arrival for migrants across the 1970, 1980, and 1990 censuses, the period length may vary slightly. For more details, see Appendix A.1.

[4]If the own birthplace is in the US, imprecisely specific (e.g., a continent), or missing, we instead use the parents' birthplace, assigning equal weights to each parent's birthplace.

[5]See Duncan and Trejo (2017) for recent evidence on recalled versus factual ancestry in CPS data.

level to the 1990 US county level. The respondents' stated ancestry (birthplace) often, but not always, directly corresponds to foreign countries in their 1990 borders (e.g., "Spanish" or "Denmark"). When no direct mapping exists (e.g., "Basque" or "Lapland"), we construct transition matrices that map data from the answer level to the 1990 foreign-country level, using approximate population weights where possible and approximate area weights otherwise. In the few cases when answers are imprecisely specific or such a mapping cannot be constructed (for example, "European" or "born at sea"), we omit the data.[6] The resulting dyadic dataset covers 3,141 US counties, 195 foreign countries, and 10 census waves.

Innovation. We use patent data to measure innovation. Starting from the universe of patent microdata provided by the US Patent and Trademark Office (USPTO) from 1975 until 2010, we study corporate utility patents with US assignees, around 4.7 million observations. We convert assignee locations provided by the USPTO in coordinate form to 2010 US counties, tabulating the number of corporate utility patents granted to assignees in each county in each year of the sample, and then use area weights to transition to 1990 US counties. In earlier periods, the location of inventors was a more natural choice for the location of innovations (Akcigit et al., 2017); however, in recent years, the overwhelming majority of patents are assigned to corporations, making assignees the natural baseline location measure for our purposes. In addition to this baseline choice, we also explore alternative means of locating patents by inventors, and we also conduct various quality or citation weighting checks following Hall et al. (2001). We sum patent flows over five-year periods, with the measure in t corresponding to the sum of patents in a given county d over the five years between $t-1$ and t. We then scale this measure by the 1970 population of county d from the Census microdata to yield a five-year patents-per-capita variable. The change in the flow of patents per capita from period $t-1$ to t is our primary outcome of interest.[7] Appendix A.3 gives additional details.

Dynamism. A growing empirical literature emphasizes that measures of dynamism and creative destruction in the US have declined in recent decades (Decker et al., 2014). Our dynamism measures come from two sources, motivated by the prior work on this subject. The first dataset – the US Business Dynamism Statistics (BDS) database from the US Census – contains measures

[6]Appendix A.1 provides a detailed description of the data transformation.

[7]We manually check the patents-per-capita measure for outliers likely due to errors in location coding by the USPTO, finding a few instances in which manual correction was possible. However, to guard against the possibility that any miscoding remains, we winsorize the resulting distribution of the change in patents-per-capita outcome variable at the 1st and 99th percentile.

computed from the underlying Longitudinal Business Data (LBD) microdata on the employment levels of the universe of US business establishments. The BDS data include job creation and job destruction rates at the yearly level and spanning 1977–2015.[8] We apportion the native MSA geography to 1990 US counties by population.[9] Our main dynamism outcomes of interest from the BDS data in county d in period t correspond to the change in either job creation or job destruction rates from $t-1$ to t.

In addition to measures of gross employment flows, the dynamism literature also emphasizes a decline over time in the skewness of employment growth rates, that is, a decline in the relative importance of "superstar" growth performance in driving US employment dynamism (Decker et al., 2014). In this spirit, we construct growth rate skewness measures starting from the US Census County Business Patterns (CBP) dataset. The raw data contain county-by-year-by-4-digit industry employment levels from 1985 to 2010. For each county and year, we compute the Kelley Skewness of employment growth rates across 4-digit sectors. This measure gives a sense of whether certain strongly performing industries drive overall employment growth in that period and location. The final measure of interest for county d in period t is the change in the growth rate skewness measure over the five years from $t-1$ to t.

Other Data. We compute local average annual wages from the Quarterly Census of Wages (QCEW) dataset provided by the US Bureau of Labor Statistics. The data stem from state-level unemployment insurance records. The QCEW records employment and wages at the county-by-industry-by-year level starting in 1975. We compute the total wages per capita in a given county-year combination, and deflate using the Personal Consumption Expenditure price index from the same source. The outcomes of interest in specifications studying income growth is the change in wages per capita in county d over the five-year period ending in t. We also construct data on the change in average annual wages for US-born working individuals (natives) and the subset of US-born working individuals who have lived in their county of residence for the past five years at the time of the Census (native non-movers) using data from IPUMS USA; for these outcomes, we consider the change in average CPI-deflated wages for natives (or native non-movers) in county d over the 10-year period ending in t.

[8]Job creation and destruction rates are gross flows representing the ratio of the number of jobs created or destroyed as a fraction of total average employment in the current and past year. See Davis and Haltiwanger (1999) for an overview of gross labor market flows.

[9]This apportionment is necessary because the BDS statistics are available only at the local level for MSAs rather than for the universe of individual counties. Creating equivalent county measures would require use of the underlying confidential Census LBD microdata.

Summary Statistics. Table 1 reports summary statistics on the outcomes described above, as well as various other instruments and derived variables studied below. The series are observed at the county by five-year-window level. The table reveals sensible patterns. Counties, on average, received around 1,400 non-European immigrants in each five-year period between 1975 and 2010, a meaningful contribution to overall average population growth of around 4,000. Innovation (as measured by per-capita patenting) increased on average over the period, with substantial heterogeneity across counties. As emphasized by the dynamism literature, measures of creative destruction including job creation rates, job destruction rates, and growth rate skewness declined on average during our sample period, although the average obscures wide differences in experience: some counties became substantially more dynamic over the period we study. Wages per capita grew on average, as expected. The statistics on the remaining variables, reflecting the variation in subsets of our data and in several constructed instruments, will become useful in our discussion below.

3 Constructing a Valid Instrument for Immigration

Our aim is to estimate the causal impact of immigration on innovation and local economic dynamsim. To do so, we estimate the following equation:

$$\Delta Y_d^t = \delta_t + \delta_s + \beta \cdot \text{Immigration}_d^t + \epsilon_d^t, \tag{1}$$

where Immigration_d^t measures the number of migrants flowing into destination county d between $t-1$ and t, ΔY_d^t is a change from $t-1$ to t in the outcome of interest, and δ_t and δ_s are time and state fixed effects, respectively. Our most conservative specifications also include a county fixed effect, δ_d, which controls for any county-specific trend in Y_d^t, so that we exploit only variation over time within a given county.

The main concern with a simple OLS estimate of (1) is that unobserved factors may affect both immigration and innovation or dynamism. For instance, it is likely that migrants are disproportionately drawn to more innovative destinations within the US. We estimate (1) in differences, so that any systematic differences in the level of innovation are controlled for. Nevertheless, migrants may be disproportionately drawn to counties within the US that are temporarily on an upward innovation trend.

To address this concern, one possibility would be to construct a "shift-share" instrument in the spirit of Card (2001), predicting immigration flows using the interaction of pre-existing foreign ancestry shares in a given destination county with the total number of migrants arriving in the US from that origin country. However, omitted factors that make a set of US counties more innovative may also have attracted disproportionately many migrants from specific sets of origin countries in the past, rendering pre-existing ancestry shares endogenous. For example, Indian engineers may be particularly good programmers and may have historically migrated to Silicon Valley (and to other information technology hubs) because those destinations provided attractive employment opportunities for programmers; for the same reason, Indian engineers would systematically migrate to Silicon Valley (and other information technology hubs) whenever the information technology industry experiences a boom. In this case, the canonical shift-share approach would falsely identify a causal effect of immigration on innovation: innovations in software are both the reason why some destinations have high pre-existing Indian ancestry shares, and why those destinations experience more Indian immigration. Ancestry shares that are themselves endogenous – potentially correlated with unobserved factors affecting innovation – thus pose a challenge to the canonical shift-share approach.

To overcome this challenge, we augment the canonical shift-share approach with a set of instruments that isolate quasi-random variation in the pre-existing ancestry composition of US counties. This variation results only from the coincidental timing of two forces driving historical migration patterns to the US: (i) time-series variation in the relative attractiveness of different destinations within the US to the average migrant arriving at the time (e.g., end of 19th century Midwest vs early 20th century West) and (ii) the staggered arrival of migrants from different origins (e.g., end of 19th century China vs early 20th century Japan). We argue the interaction of these two forces can be used to construct valid instruments for the distribution of ancestries across US counties that are orthogonal to origin-destination specific confounding factors, such as the affinity of Silicon Valley and Indian engineers for software development mentioned above. We then use only the exogenous component of the pre-existing distribution of ancestries to predict migration into each US county post 1975. Doing so, we eliminate a wide range of concerns relating to the endogeneity of pre-existing ancestry composition. We discuss this procedure, its merits, and also limits, in detail below.

3.1 Constructing an Instrument for Immigration

To construct our instrument for the number of migrants flowing into a given destination county at a given point in time, we build upon Burchardi et al. (2019), and start from a simple reduced-form model of migration. Migrants from origin country o settle in destination county d at time

t according to

$$I_{o,d}^{t} = \delta^{t} + \delta_{o}^{t} + \delta_{d}^{t} + X'_{o,d}\beta + I_{o}^{t}\left(a_{t}\frac{I_{d}^{t}}{I^{t}} + b_{t}\frac{A_{o,d}^{t-1}}{A_{o}^{t-1}}\right) + u_{o,d}^{t}, \tag{2}$$

where ancestry evolves recursively as cohorts of migrants accumulate,

$$A_{o,d}^{t} = \delta^{t} + \delta_{o}^{t} + \delta_{d}^{t} + X'_{o,d}\gamma + c_{t}I_{o,d}^{t} + d_{t}A_{o,d}^{t-1} + v_{o,d}^{t}. \tag{3}$$

In both equations, the δ terms are fixed effects, and the vector $X'_{o,d}$ controls for observables.

Our key assumption on the forces driving migration, upon which our identification is built, corresponds to the interaction terms, $I_{o}^{t}\left(a_{t}\left(I_{d}^{t}/I^{t}\right) + b_{t}\left(A_{o,d}^{t-1}/A_{o}^{t-1}\right)\right)$. We model the choices of migrants as driven by two distinct forces, which we label "economic push-pull" and "social push-pull". The economic push-pull force is captured by the term $a_{t}I_{o}^{t}\left(I_{d}^{t}/I^{t}\right)$: in time periods when many migrants arrive from country o to the US (a large I_{o}^{t} push factor), and when a destination county d is particularly attractive to the average migrant arriving at the time (a large economic pull I_{d}^{t}/I^{t} factor), we expect many migrants from o to settle in d. This force corresponds to an economic motive for migration: upon arriving in the US, migrants tend to flock to destination counties that are attractive to the average migrant arriving at the time. The social push-pull force is captured by the term $b_{t}I_{o}^{t}\left(A_{o,d}^{t-1}/A_{o}^{t-1}\right)$: migrants arriving from o (the push factor I_{o}^{t}) tend to locate in destinations d with a pre-existing community from their home country (the social pull factor $A_{o,d}^{t-1}/A_{o}^{t-1}$). This force corresponds to a social motive for migration: all else equal, migrants tend to settle near others of their own ethnicity.

The seminal Card (2001) shift-share approach is simply a special case of our model (2)-(3), ignoring the economic push-pull in (2) by setting $a_{t} = 0$, and using pre-existing ancestry shares as an instrument for contemporaneous migrations, ignoring the recursive nature of ancestry in (3). As we discuss above, this shift-share instrument is invalid if omitted factors affect both innovation and past immigration, inducing endogenous ancestry shares.

Our full model (2)-(3) offers a solution: we use the recursive set of equations to construct a valid instrument for immigration. Iterating (2)-(3) over several periods prior to 1975, we isolate plausibly exogenous variation in pre-existing ancestry inherited exclusively from the cumulated series of economic push-pull forces. We can then use this predicted stock of ancestry in the social push-pull term of (2) post 1975 to isolate plausibly exogenous immigration shocks.

To fix ideas more concretely, Figures 1 and 2 directly examine the underlying variation in economic push and pull factors.

Figure 1 shows time variation in the economic push factor. It plots the share of non-European immigration to the US from the five non-European origin nations with the largest cumulative immigration to the US. This push-factor variation within countries is generally clustered in time in bursts of immigration to the US, often driven by historic events in the home countries or by changes in origin-specific rules for migration to the US. For example, Mexican migration to the US experiences a spike during the period of the Mexican Revolution from 1910-20. Cuban immigration flows increase during the 1960s and 1970s in the decades after the Cuban revolution. Immigration from Vietnam reaches substantial numbers only from the mid 1970s onwards in the wake of US involvement in the Vietnam War. Chinese and Japanese migration to the US fell from relatively higher levels early in the sample to low levels before increasing over time, in this case as various US immigration exclusion acts were repealed.

Figure 2 shows time variation in the economic pull factor. It plots color-coded maps of European migration to the US over Census waves from 1880 to 2010, with darker shades representing a higher intensity of migration to a given county. The location of relatively attractive destination counties – our source of variation for the economic pull factor – changed substantially over time. Early on in the sample during the late 19th century, northeastern locations were particularly attractive destinations. By the early 1900s, the average European immigrants' favored destinations shifted to the midwestern and western regions, before shifting again to the coastal and southeastern regions.

So, to summarize, the rich variation in Figures 1 and 2 allows us to isolate variation in pre-existing ancestry attributable to the coincidence of historical push and economic pull factors operating on the average immigrant arriving in the US at different times from different origins. Instead of, say, considering the stock of individuals with Mexican ancestry in each US county in 1975, our eventual set of quasi-random variation instead exploits, for example, the fact that certain southeastern and midwestern regions happened to be popular destinations (economically "pulling" people into destination counties $d \in \{Southeast, Midwest\}$) during the period of heavy Mexican immigration around the Mexican Revolution ("pushing" people out of that origin nation $o = Mexico$).

Having isolated plausibly exogenous variations in pre-existing ancestry, we can then confidently use the social push-pull force to predict contemporaneous immigration shocks post 1975.

In practice, we construct our instrument for the number of migrants flowing into a given destination county at a given point in time in three steps, each of which is easy to implement and follows the simple logic of the above model of migration.

Step 1: Isolating quasi-random variation in ancestry. We predict the number of residents of destination county d with ancestry from origin country o in baseline year t (in thousands), $A^t_{o,d}$, by using the economic push-pull force in (2), and by cumulating successive migration waves using (3). We complement this model by adding "leave-outs," to ensure our instruments are not polluted by any, potentially confounding, origin-destination-specific factors. Formally, we estimate

$$A^t_{o,d} = \delta^t_{o,r(d)} + \delta^t_{c(o),d} + X'_{o,d}\zeta + \sum_{\tau=1880}^{t} a^\tau_{r(d)} I^\tau_{o,leaveout} \frac{I^\tau_{leaveout,d}}{I^\tau_{leaveout}} + v^t_{o,d}. \tag{4}$$

The leave-outs ensure we do not use the endogenous choice of migrants from o to settle in d to predict ancestry from o in d. In our baseline, we use as our push factor $I^\tau_{o,leaveout} = I^\tau_{o,-r(d)}$, the total number of migrants arriving from o who settle in locations *outside* of the region where d is located over the five-year period ending in τ (our regions are defined as Census divisions, a grouping of several adjacent US states, as shown in Appendix Table 1); and we use for our pull factor $I^\tau_{leaveout,d}/I^\tau_{leaveout} = I^\tau_{Europe,d}/I^\tau_{Europe}$, the fraction of all incoming European migrants who settle in d.[10] Our results are robust to using various alternative leave-out strategies, as we show below. $\delta^t_{o,r(d)}$ and $\delta^t_{c(o),d}$ are a series of origin country $\times$ destination region and continent of origin $\times$ destination county interacted fixed effects, whereas $X_{o,d}$ contains a series of time-invariant controls for $\{o, d\}$ characteristics (including distance and absolute latitude difference). We estimate (4) separately for each $t = 1980, 1985, 1990, 1995, 2000, 2005, 2010$ using all non-European countries in our sample.

From this estimation, we derive predicted ancestry in county d from origin o at time t as

$$\hat{A}^t_{o,d} = \sum_{\tau=1880}^{t} \hat{a}^\tau_{r(d)} \left(I^\tau_{o,leaveout} \frac{I^\tau_{leaveout,d}}{I^\tau_{leaveout}} \right)^{\perp}, \tag{5}$$

where $\left(I^\tau_{o,leaveout}{}^{I^\tau_{leaveout,d}}/{}_{I^\tau_{leaveout}}\right)^{\perp}$ are residuals of a regression of $I^\tau_{o,leaveout}{}^{I^\tau_{leaveout,d}}/{}_{I^\tau_{leaveout}}$ on $\delta^t_{o,r(d)}$, $\delta^t_{c(o),d}$ and $X_{o,d}$, and $\hat{a}^\tau_{r(d)}$ are the coefficients estimated from (4). This additional orthogonization ensures our constructed immigration shocks rely only on the exogenous component of pre-existing ancestry, after removing all variation that could also be accounted for by fixed effects or other observables. Again, our baseline specification uses region and continental leave-outs, $I^\tau_{o,leaveout} = I^\tau_{o,-r(d)}$ and $I^\tau_{leaveout,d}/I^\tau_{leaveout} = I^\tau_{Europe,d}/I^\tau_{Europe}$.

[10] The focus of our main regression of interest is on non-European migrants who arrived in the US in recent decades, a period during which most migrants were *not* coming from Europe. Using the historical migrations of Europeans to predict the settlement patterns of non-Europeans ensures our results are not driven by other origin countries with similar characteristics and settlement patterns. Note this leave-out imposes a stricter requirement than simply removing o's migrants from I^τ_d/I^τ.

Step 2: Predicting migration from individual countries. Having isolated plausibly exogenous variation in the stock of ancestry at the $\{o, d\}$ level for all periods after 1975, we use the social push-pull force from (2) to predict contemporaneous immigration. This method is similar to Card (2001), except we address the concern that ancestry itself is an endogenous variable. We predict immigration from o to d in period t by estimating

$$I^t_{o,d} = \delta_{o,r(d)} + \delta_{c(o),d} + \delta_t + X'_{o,d}\theta + b_t \cdot [\hat{A}^{t-1}_{o,d} \times \tilde{I}^t_{o,leave-out}] + u^t_{o,d}, \tag{6}$$

where again the δ's are time, country×region, and continent×county fixed effects, $X'_{o,d}$ observable controls, $\hat{A}^{t-1}_{o,d}$ is predicted ancestry from (5), and $\tilde{I}^t_{o,leave-out} = I^t_{o,-r(d)} \left(I^t_{Europe,r(d)} / I^t_{Europe,-r(d)} \right)$. Because we leave out from $I_{o,-r(d)}$ all migrants from o who settle in d's region, $\tilde{I}^t_{o,leave-out}$ includes a scaling factor at the regional level, $I^t_{Europe,r(d)} / I^t_{Europe,-r(d)}$, to correct for contemporaneous differences in region sizes.

Step 3: Generating immigration shocks. We are finally able to generate our main instrument for the total number of migrants settling in county d in period t, Immigration^t_d in equation (1),

$$\hat{I}^t_d = \sum_o \hat{b}_t \cdot [\hat{A}^{t-1}_{o,d} \times \tilde{I}^t_{o,leave-out}]. \tag{7}$$

Identifying assumption. A sufficient condition for the validity of this instrument is that instrumented ancestry, $\hat{A}^{t-1}_{o,d}$, is truly exogenous in equation (1). With our baseline regional and continental leave-outs, we can write this condition as

$$I^\tau_{o,-r(d)} \frac{I^\tau_{Europe,d}}{I^\tau_{Europe}} \perp \epsilon^t_d \ \forall o, d, \tau \leq t. \tag{8}$$

It requires that any confounding factors that drive temporary increases in given US county's' innovation post-1975 (ϵ^t_d) do not systematically correlate with vintages of the historical interaction of pre-1975 immigration from that origin to other regions within the US ($I^\tau_{o,-r(d)}$) and past instances of the simultaneous settlement of European migrants in that US destination ($I^\tau_{Europe,r(d)} / I^\tau_{Europe,-r(d)}$). If this condition is satisfied, the ancestry shares used to predict immigration in Step 2 are exogenous, as is the variation in total immigration calculated in (7).[11]

We believe this assumption is plausible: consider again a productivity shock to software development in Silicon Valley ($\epsilon^t_{Santa\,Clara}$) that attracts Indian software engineers ($I^t_{India,Santa\,Clara}$).

[11] Exogeneity of ancestry shares is a sufficient, but generally not a necessary, condition for the validity of the canonical shift-share approach (Goldsmith-Pinkham et al., 2018). For work identifying necessary and sufficient conditions for the validity of the shift-share instrument as proposed by Card (2001) and Bartik (1991), see Borusyak, Hull, and Jaravel (2018).

This confounding shock, and any other origin-destination specific factor that drives migration and might affect the destination's capacity for future innovation generally has no effect on $\hat{A}^{t-1}_{India,Santa\,Clara}$: the fact that Indians excel at programming has no effect on how the historical destination choices of Europeans ($I^{\tau}_{Europe,r(Santa\,Clara)}/I^{\tau}_{Europe,-r(Santa\,Clara)}$) coincide with the number of Indians arriving in US destinations outside the West Coast ($I^{\tau}_{India,-r(Santa\,Clara)}$). To violate (8), the confounding shock ($\epsilon^{t}_{Santa\,Clara}$) would instead have to systematically affect the destination choices of a large number of Europeans, say French software engineers (a shock to $I^{\tau}_{France,Santa\,Clara}$ large enough to affect $I^{\tau}_{Europe,r(Santa\,Clara)}/I^{\tau}_{Europe,-r(Santa\,Clara)}$), while also attracting a large number of Indians to US counties outside the West Coast, say to Route 128 in Massachusetts (a shock to $I^{\tau}_{India,Middlesex}$ large enough to affect $I^{\tau}_{India,-r(Santa\,Clara)}$). We address this remaining (if unlikely-sounding) concern below by varying how we construct the leave-out categories in our estimation.[12]

3.2 The Construction and Performance of the Instrument

We now review the estimation results of each of the steps toward the construction of our instrument, including the performance of the resulting instrument for county-level immigration in the relevant first-stage regression.

In Step 1 of our instrument construction, we isolate quasi-random variation in ancestry ($\hat{A}^{t}_{o,d}$) using historical interactions between the push and economic pull factors in (5). Figure 3 reports the estimated coefficients on these interactions when predicting 2010 ancestries (assuming for presentational purposes only that $a^{\tau}_{r(d)} = a^{\tau}\ \forall r(d)$). The results indicate we identify variation in current ancestry levels across the full range of time periods in our sample, with statistically precise contributions from periods as far back as the pre-1900s census waves. These coefficients are positive and mostly significant. The negative coefficient in the late 1920s is consistent with large return migrations during the Great Depression, when arriving migrants swiftly returned home and possibly attracted earlier migrants to follow suit (Abramitzky and Boustan, 2017). Figure 4 presents a bin scatter plot of $\hat{A}^{2010}_{o,d}$ against realized ancestry in 2010. The two variables are tightly aligned along the 45-degree line, suggesting that the interaction of historical push and economic pull factors is a powerful predictor of the present-day composition of ancestry.

In Step 2, we interact lagged predicted ancestry with contemporaneous scaled push factors ($\hat{A}^{t-1}_{o,d} \times \tilde{I}^{t}_{o,-r(d)}$) for each five-year period post 1975 to predict plausibly exogenous variation in

[12] Note this scenario implies a series of non-zero correlations, pairwise between $\epsilon^{t}_{Santa\,Clara}, \epsilon^{t}_{Middlesex}, I^{t}_{India,Santa\,Clara}, I^{t}_{India,Middlesex}$ and $I^{\tau}_{France,Santa\,Clara}$. We explicitly explore this possibility in Section 4.2.

immigration $I_{o,d}^t$ at the $\{o, d\}$ level in (6). We allow the coefficient b_t to vary by time period t, and Table 2 reports the resulting estimates. Our ancestry and push factors positively and significantly predict immigration at the country-county level in all seven time periods post 1975, with an R^2 value of 65.6% in column 1 indicating high explanatory power with no other predictors included. The following columns add controls for distance and latitude difference, as well as a full set of origin-country, destination-county, and time fixed effects. Column 3 adds a total of 12,564 interactions of origin-country $\times$ destination-census-division and destination-county $\times$ continent-of-origin fixed effects. Throughout these variations, the coefficients on our (instrumented) push-social pull terms remain virtually unchanged. Remarkably, they even remain unchanged in column 5, where we control directly for contemporaneous economic forces shaping migration by including economic push-pull interactions for each period post 1975,[13] and even when we include the (endogenous) total flow of European migration to the same county as an additional control in column 4. We conclude our instruments for origin-destination-specific migration are orthogonal to a wide range of observables, and Step 2 successfully predicts plausibly exogenous variation in immigration at the $\{o, d\}$ level.

In Step 3, we sum across origin countries to compute our instrument $\hat{I}_d^t$ for total non-European immigration to county d at time t (7). Figure 5 presents a series of maps displaying this exogenous immigration shock for each five-year period from 1975 to 2010. To highlight the variation in this immigration shock, we remove county and year fixed effects. The maps in Figure 5 show our "immigration shock" instrument picks up substantial variation both over time and between counties.

4 The Impact of Immigration on Innovation and Growth

In this section, we exploit our quasi-random immigration shocks above to test and quantify the causal link between immigration and innovation explicitly.

4.1 Immigration and Innovation

We first test the hypothesis that immigration *causes* an increase in innovation at the county level. Table 3 panel A shows estimates of (1), where we instrument for the number of immigrants

[13]Because this specification is saturated with controls, the incremental increase in R^2 from adding this variable appears small. However, if we allow for flexible coefficients at the census region level (as in (4)), the R^2 increases by about six percentage points.

arriving in the county during the five-year period using (7). The dependent variable is the change in patenting *per capita* relative to the previous five-year period. Column 2 shows our standard specification, which includes state and time fixed effects, thus controlling for differential trends in innovation growth at the state level. The estimated effect is positive and statistically highly significant (0.101, s.e.=0.031). We interpret it as the local average treatment effect of immigration (particularly, immigration induced by the social push-pull factor in (2)) on county-level innovation. It implies the arrival of 10,000 additional immigrants in a given county on average increases the flow of patents filed over a five-year period by one patent per 100,000 people. Comparing these magnitudes to the summary statistics above, an increase in immigration flows of one standard deviation - 12,000 immigrants - causes around 1.2 more patents per 100,000 people, an increase of 27% relative to the mean (4.45 patents per 100,000 people).[14]

Panel B shows the immigration shock we construct through the procedure outlined before, $\hat{I}_d^t$, is highly predictive of (non-European) immigration. In our standard specification shown in column 2, a regression of a county's five-year immigration flow on $\hat{I}_d^t$ yields a coefficient of 2.119 (s.e.= 0.070), suggesting a one standard deviation increase in our immigration shock (4.99) is associated with a 0.86 standard deviation increase in the county's number of newly arrived immigrants. The F-statistic on the excluded instrument in the first stage of our standard specification in (7) is 911, far above conventional critical values.

Standard errors in this and all subsequent specifications are clustered by state. This choice is conservative because it ensures our standard errors are robust to correlations in the error structure both within counties over time, across counties within a state, and even across counties in different time periods within a state.[15]

Column 1 shows the OLS estimate of (1) for comparison. As expected, it is larger than our preferred estimate (by about two standard errors), consistent with the view that, all else equal, immigrants select into innovative counties in equilibrium, resulting in an upward bias in OLS estimates of the effect of immigration on innovation.[16]

Columns 3 and 4 show our results are robust and the estimated impact of immigration on innovation varies little if we include interacted time and state fixed effects (0.100, s.e.=0.032 in

[14] To first focus on issues of identification and the sign of the local average treatment effect, we defer a detailed characterization of functional forms, particularly those commonly used in endogenous growth models, to Section 5.3.

[15] Appendix Table 2 reports results from a number of permutation tests of this choice. When reassigning the instrument, right-hand-side rejection rates in the first stage vary between 0.00% and 2.70% where a nominal rejection rate of 2.50% would be expected. For the reduced form, they vary between 0.70% and 4.40%.

[16] Notice that with clustering, the OLS standard errors are not necessarily smaller than the IV standard errors.

column 3), or even county fixed effects (0.108, s.e.=0.033 in column 4).[17] Panel B of Table 3 shows the immigration shock we exploit continues to be highly predictive of immigration also in these alternative specifications.

We find similar results and magnitudes for the impact of population growth on innovation, instrumenting population growth with immigration shocks; see Appendix Table 3. Consistent with the positive effect of immigration on innovation, we also find immigration shocks generate agglomerative effects. That is, an exogenous increase in the number of immigrants to a county also attracts more native-born Americans to that same county; see Appendix Table 4.

4.2 Robustness

Below, we show our results are robust to a large array of alternative specifications.

Alternative Instruments. Recall that a confounding factor violating our assumption (8) would have to systematically generate economic dynamism in a given US destination (e.g., Silicon Valley), attract immigration from a non-European origin country (e.g., India) to that US destination as well as to other US regions (e.g. Route 128), and would need to be correlated with past instances of the simultaneous settlement of European migrants (e.g. France) in that US destination and a large immigration from that non-European origin to other US regions. An instance of such a concern would be that pre-1975 periods of IT innovation attracted Indian *as well as* French software engineers to Silicon Valley *as well as* to Route 128, and post-1975 periods of IT innovation have the same effect.

Table 4 shows how variants of our instrumentation affect our estimates of (1). Column 2 uses a different leave-out strategy for the push factor in step 1 of the construction of our instrument: instead of leaving out migrants from country o who settle in the same census division as county d when predicting migrations from o to d, we leave out migrants from o who settle in counties with migrations that are serially correlated with those toward d. Because this instrument removes immigrants to counties with a similar pattern of migration as county d form the push factor, it is immune to the concern, highlighted above, that immigration from a country (e.g., Indian software engineers) is driven by innovation in a set of correlated counties (e.g., IT hubs in Silicon Valley and Route 128). Columns 1 and 3 present further variants of our instrument. Column 1 freezes predicted ancestry at its 1975 level, instead of updating

[17] The fact that the estimates with and without county fixed effects are almost identical (0.100, s.e.=0.032 vs. 0.108, s.e.=0.033) strongly suggests that $\hat{I}_d^t$ is not spuriously correlated with highly persistent responses to prior shocks, a common problem with traditional implementations of the shift-share approach emphasized in recent work by Jaeger et al. (2018).

predicted ancestry each period. Column 3 uses a different leave-out strategy for the pull factor in step 1: instead of using the European migrants to county d as a measure of the pull toward d when predicting migrations from o to d, we use all migrants to d originating from countries outside o's continent. All of these variations yield estimates that are almost identical to the one in our standard specification (0.101). If confounding factors as described above were indeed at work, we would expect our estimates to change dramatically when we change the leave-out categories in the construction of our instruments to exclude the number of migrants arriving in destinations that tend to receive inflows of migrants at the same time (column 2) or when we use shares of migrants from other continents (instead of Europeans) to measure historical pull factors (column 3). Instead, our estimates remain virtually unchanged (0.098, s.e.=0.033 and 0.094, s.e.=0.027, respectively).

Construction of the Baseline Instrument. The construction of our baseline instrument $\hat{I}_d^t = \sum_o \hat{\gamma}_t \cdot [\hat{A}_{o,d}^{t-1} \times \tilde{I}_{o,leave-out}^t]$ differs from canonical applications of the shift-share approach (Card, 2001) in three respects. First, it instruments for pre-existing ancestry; second, it leaves out all migrants from o who migrate to the same census region as d when calculating the push ($\tilde{I}_{o,leave-out}^t$); and third, it uses a different functional form, where migrants are assumed to respond to the number of individuals of their own ancestry in d rather than their share in the local population. Table 5 retraces each of these steps to make clear how each modification affects our estimates and how they help address econometric shortcomings of canonical applications of the shift-share approach highlighted in the recent literature (Adão et al., 2019).

Column 1 implements our baseline instrument but replaces ancestry in levels with ancestry shares, so that $\hat{I}_d^t = \sum_o \hat{\gamma}_t \cdot [\tilde{A}_{o,d}^{t-1} / \tilde{A}_o^{t-1} \times \tilde{I}_{o,leave-out}^t]$. This procedure has the added complication that, in some instances, predicted ancestry shares lie outside of the $[0, 1]$ interval, because predicted ancestry from the linear model in (5) is sometimes negative. We remedy this issue by performing a simple translation of predicted ancestries that avoids negative shares, $\tilde{A}_{o,d}^{t-1} = \hat{A}_{o,d}^{t-1} - \min[0, \min_\delta[\hat{A}_{o,\delta}^{t-1}]], \forall \{t-1, o\}$. Using this translation, we find a larger positive and significant effect of immigration on innovation (0.195, s.e.=0.090), though the larger standard error makes it statistically indistinguishable from our standard specification.

Though less statistically precise, this formulation of our instrument has the advantage of allowing us to test whether our instrumentation approach successfully addresses an over-rejection problem that has been shown to arise in conventional applications of the shift-share approach, which take pre-existing ancestry shares as given (Adão et al., 2019). This over-rejection problem arises because two US counties with similar pre-existing ancestry composition may also have similar exposure to other (unobservable) economic forces, which may lead to a dependency across regression residuals that is not accounted for by conventional clustered standard errors. To test for this issue, we implement the statistical placebo test for shift-share instruments pioneered by Adão et al. (2019).[18] Following their procedure, we randomly generate immigration shocks (for each $\{o, r, t\}$ country-region-time triplet), and construct placebo instruments by interacting these random shocks with our predicted ancestry shares. We then run 1,000 placebo regressions of actual immigration on our randomly generated instrument, and report the fraction for which we reject the null hypothesis of no effect at the 5% statistical significance threshold. Comforting for our inference, we find a false rejection rate of 4.5%, almost exactly equal to the theoretical asymptotic 5% level.[19]

For comparison, column 2 repeats the same estimation as that of column 1 but utilizing realized rather than predicted ancestry shares, so that $\hat{I}_d^t = \sum_o \hat{\gamma}_t \cdot [A_{o,d}^{t-1}/A_o^{t-1} \times \tilde{I}_{o,leave-out}^t]$. Consistent with the findings in Adão et al. (2019), the false rejection rate is now close to 28%, far above the expected 5%, pointing to a significant tendency to over-reject the null (and a correspondingly much narrower standard error, although the point estimate remains similar). Finally, in column 3, we fully converge to the conventional shift-share approach by also dropping our leave-out adjustment (so that $\hat{I}_d^t = \sum_o \hat{\gamma}_t \cdot [A_{o,d}^{t-1}/A_o^{t-1} \times \tilde{I}_o^t]$). The coefficient of interest is again close to our standard specification (0.132, s.e.=0.055) but continues to suffer from dramatic over-rejection in the placebo test, close to 28%. The fact that our instrument circumvents this statistical problem is intuitive: although the spatial correlation of realized ancestry shares is related to the spatial correlation of the second-stage error term, our instrument only exploits a small share of the variation in ancestry shares.

We conclude our instrumentation strategy is sufficiently powerful to isolate quasi-random variation in ancestry levels, or shares, and that it effectively removes spurious correlations with the error term, bolstering our confidence in a causal interpretation of the results. Finally, note that all approaches to identification, including the simpler ones, unanimously find a positive effect of innovation on county-level innovation.

Robustness: Additional Controls. In Table 6, we go one step further and control parametrically for a number of initial conditions that could be considered drivers of long-term economic

[18] To clarify the comparison, the shifts are industry shocks in Adão et al. (2019) versus immigration shocks in our case, whereas the shares are employment shares in Adão et al. (2019) versus ancestry shares in our case; the variation is at the sector-commuting zone level in Adão et al. (2019), versus country-county in our case.

[19] We report additional details of this placebo test in Appendix Table 5.

growth: population density in 1970, the number of patents generated in 1975 per 1,000 inhabitants (1975 is the first year for which our patent data is available), and the share of the 1970 population that is high-school and college educated, respectively. All of these covariates could be considered "bad controls" (Angrist and Pischke, 2009), in the sense that they may be outcomes of migration and should thus more appropriately be thought of as various channels through which historical migrations and ancestry affects innovation. Nevertheless, the finding that controlling for these initial conditions has only modest effects on our result is comforting for our identification assumption. The largest change in the coefficient of interest occurs when we include the share of the population with a college education in 1970, lowering it from 0.101 to 0.082, less than one standard error (s.e.=0.031). Column 6 imposes an even stronger identification restriction, by including a county fixed effect to control for county-specific trends in innovation (as already shown in Table 3). Throughout these variations, the estimated effect of immigration varies little, and it remains positive and statistically highly significant.

Robustness: Alternative Samples. Table 7 further probes the robustness of these results by excluding important origin countries (panel A), or using *only* important origin countries (panel B). In panel A, we exclude migrations from the five largest sending countries post 1975 (Mexico, China, India, Philippines, and Vietnam), one at a time, from the sum in (7), thus treating migration from these countries as endogenous. Although dropping Mexican immigration from our instrument lowers the F-statistic in the first stage by about half (to 666), the estimated coefficients vary little across these alternative samples, showing that no single large sending country drives our results. In panel B, we include only migration from those individual origin countries. Again, although our F-statistics decrease as we move to smaller origin countries, the coefficients vary surprisingly little across specifications.

Robustness: Different Time Horizons. Table 8 serves two purposes. First and reassuringly, we show in column 1 that contemporaneous migrations have no effect on past innovation. This finding further strengthens our confidence in our identification strategy. Second, in column 2-4, we explore the dynamic impact of exogenous immigration shocks on innovation. Column 2 replicates our baseline results, the contemporaneous effect of immigration on innovation over a five-year period, as in Table 3, Panel A, column 4. Columns 3 and 4 consider the impact of immigration on patenting over a 10- and 15-year period. We find the effect of immigration on innovation gradually increases, and stabilizes after about 10 years. In other words, the effect more than doubles from 5 to 10 years, and remains constant beyond. This speed of adjustment is plausible, and consistent with endogenous growth models, as the population shock induced by immigration gradually percolates through the local labor market and firms are able to innovate.

Robustness: Alternative Measures of Innovation. We show in Appendix Table 6 that our results are robust to alternative measures of innovation. Our measure for innovation comes from USPTO patent microdata. In our main specification, we assign patents to a specific county d according to the firm or assignee owning the patents, and we treat all patents as equally important, simply counting the total number of patents. We consider two variations on each dimension. First, we assign patents according to the place of residence of the inventor of the patent, not the firm owning the patent. Second, we weight each patent according to their relative citation counts following Hall et al. (2001) in order to distinguish between high-impact/high-citation patents and low-impact/low-citation patents. Across all four possible measures of patents, our results are similar, with a positive and significant effect of immigration on innovation, and a similar estimated size for this effect. The differences in the point estimates simply reflect differences in the scale of the various measures of innovation.

4.3 Immigration, Economic Dynamism, and Income Growth

In Table 9, we supplement our analysis of the impact of immigration on innovation with a range of additional economic dynamism and income growth outcomes, which endogenous growth theory suggests should link positively to innovation.

Immigration causes an increase in creative destruction or gross flows in jobs, as reported in columns 1 and 2. Both the job creation rate and job destruction rate increase with immigration, implying the overall churning or reallocation in the labor market also responds positively. Recall that dynamism measures decline on average over this period, as emphasized by the wide literature on declining creative destruction in the US. Our positive estimated responses to immigration indicate immigration may help dampen such declines. Turning to magnitudes, a one standard deviation increase in immigration in a county - around 12,000 more people - causes an increase in the job creation rate of 2.1 percentage points (around 7% relative to the mean decline) and an increase in the job destruction rate of 1.8 percentage points (around 11% relative to the mean decline).[20]

Exploring an alternative measure of dynamism, higher immigration causes an increase in

[20]Note that although most endogenous growth theories link higher dynamism to innovation, higher income growth, and higher welfare, the impact of dynamism on the subjective well-being of individuals exposed to such creative destruction is more ambiguous (Aghion et al., 2016).

the skewness of employment growth (column 3). Intuitively, when more immigrants arrive, the best-performing "superstar" sectors outpace the broader growth of the local economy (Decker et al., 2014). A one standard deviation increase in immigration causes about a 3% increase in skewness relative to the mean decline in this measure over the sample.

At their core, endogenous growth models link innovation to income growth, and column 4 of Table 9 confirms that more immigration causes an increase in wages per person. Immigration of around 12,000 more people to a county, on average, increases wages per capita by around 3% relative to the mean observed growth.

Because the QCEW wage data do not allow us to distinguish between wages of natives and non-natives, columns 5 and 6 repeat this estimation using 10-year changes in average wages measured from the US census, aggregating separately across all natives (individuals born in the US) and natives who report having lived in the same county five years prior to the census. We find a positive and statistically significant effect of immigration on the average wage of both groups, with estimates of 0.049 (s.e.=0.106) and 0.056 (s.e.=0.020), respectively.[21] Because of the different time horizon and various differences in how wages are measured in the two data sources, these latter effects appear smaller than the one shown in column 4. Nevertheless, they imply similar relative effects: the arrival of 12,000 additional migrants (half of a standard deviation in 10-year immigration) leads to an increase in the average wage of natives and native non-movers of 5.5% and 4%, respectively.[22]

To summarize the estimates in this section, immigration causes moderately large increases in creative destruction and income growth at the local level, validating traditional endogenous growth theories, and potentially serving as a potent counterweight to trend decline in dynamism and growth in the US in recent decades.

[21]Note this positive average effect of immigration to the US on the average wage of native non-movers does not preclude the possibility that the influx of migrants may lower the wages of *some* native workers, or even the average wage of all natives in some historical circumstances. For example, average wages may increase while wages of a sub-group of native workers who are in direct competition with immigrants may fall. For an overview of an active literature on this subject see, Borjas (2003), Cortes (2008), Ottaviano and Peri (2012), Foged and Peri (2016), Dustmann et al. (2017), Monras (forthcoming), Jaeger et al. (2018), and Bratsberg et al. (2019). We show some direct evidence of such heterogeneity in section 5.

[22]Interesting for the interpretation of our results in the context of structural spatial growth models, we find this effect of immigration on wages is higher in services (non-traded sector), with a coefficient of 0.239 (s.e. 0.082), than in manufacturing (traded sector), with a coefficient of 0.147 (s.e. 0.053).

5 Spillovers and Education

The local positive impact of immigration on innovation that we document above validates long-standing theoretical mechanisms linking innovation to population growth. However, two natural questions remain. First, if ideas and goods flow across regions, to what extent do the impacts of immigration spill over across counties in our data? Second, because most theoretical models predict more highly skilled workers bring more effective input to bear for innovation or production, to what extent do the impacts of immigration on innovation vary with the education level of migrants? We tackle both issues directly in this section, and find that positive spillovers appear meaningful and that the impact of immigration on innovation increases with average schooling levels.

5.1 Spatial Spillovers

To explore the impact of cross-county spillovers, we consider three geographic spillover concepts in Table 10. First, we consider within-state spillovers, constructing for each destination county d at each time t a measure of immigration to all counties other than d in the same state. This measure, labeled $Immigration^t_{State}$, varies at the same level as the county-specific baseline immigration flow $Immigration^t_d$. To construct a separate instrument for state-level immigration flows, we simply add up the immigration shocks for all other counties within the same state as d. In a second approach, we consider a specification allowing spillovers from neighboring counties to vary smoothly by distance. For county d at time t, we construct the sum of all immigration to other counties, inversely weighted by the distance to the reference county d. The distance measures reflect a matrix of great circle distances computed from county centroids using the Census mapping files for county geographies. The resulting distance-weighted measure of immigration to other counties, labeled $Neighbor's\, Immigration^t_d$, varies at the county d by time t level. Finally, we also consider a non-parametric estimate for the diffusion of the effect of immigration, with separate instruments for immigration within 100km (60 miles), excluding county d itself, immigration between 100km and 250km (150 miles), and between 250km and 500km (300 miles).

We explore the spatial spillovers of immigration on both innovation (panel A of Table 10), and on local wages (panel B of Table 10).

Innovation Spillovers. In column 1 of Table 10, we first report an instrumental-variable (IV) estimate of the effect of own-county immigration on innovation using census division instead of state fixed effects. The coefficient of interest is similar to those in Table 3 (0.130, s.e.=0.039). Column 2 adds a second endogenous variable, the state-level sum across other counties' immigration. The first-stage F statistics reveal strong power for both the own-county and state-level immigration flows.[23] The impact of own-county migration on immigration remains strongly positive with a similar magnitude. In addition to this direct effect of immigration, more immigration to other counties within the same state also increases local innovation. The magnitudes implied by column 2 are sensible. A one standard deviation increase in immigration to a county (12,000 people) on average increases patenting per capita by 29% relative to mean, holding state-level immigration to other counties constant. Similarly, a one standard deviation increase in immigration to all other counties in the state (1.4 million more immigrants), holding the county's own immigration flow constant, increases patenting per capita by around 31% relative to the mean. In other words, both local immigration and immigration to the surrounding state positively affect local innovation. Although migrants to other counties matter less individually for a county's innovation, the larger scale of those flows means such immigrants bring similarly sized innovation impacts to the local economy.

Columns 3 and 4 explore the spatial diffusion of the effect of immigration on innovation, doing away with the somewhat arbitrary notion of state boundaries. Column 3 shows immigration to nearby counties, where we discount distant counties inversely with distance.[24] We show immigration to nearby counties (defined as geographically proximate counties) has a strong positive effect on innovation. Column 4 quantifies this spatial diffusion in a non-parametric way. It shows immigration to nearby counties has a positive effect on innovation, but this effect dies out with distance. A one standard deviation increase in immigration within 100km increases innovation by 80% relative to the mean; a one standard deviation increase in immigration between 100km and 250km increases innovation by 42%; but immigration beyond 250km no longer has a statistically detectable effect on innovation.

Appendix Table 7 displays the first-stage regression results, showcasing the strength of our identification strategy. We are able to successfully identify independent variations for each of the separate endogenous variables. For instance, columns 6-9 present the first-stage results corresponding to each of the separate instruments for the non-parametric spatial diffusion model of column 4 in Table 10. Each instrument, for local immigration shocks ($\hat{I}_d^t$), immigration shocks within 100km ($\hat{I}_{100km}^t$), between 100km and 250km ($\hat{I}_{100km}^t$), between 250km and 500km ($\hat{I}_{500km}^t$),

[23]For all specifications involving multiple endogenous variables, we use the Angrist and Pischke (2009, p. 217-218) first-stage F-statistic, separately testing for each regressor the null of weak identification.

[24]This measure is akin to a measure of market access in international trade.

correctly predicts actual local immigration, actual immigration within 100km, between 100km and 250km, and between 250km and 500km. This ability to simultaneously identify exogenous variations in immigration for different units is one of the strengths of our identification method.

Wage Growth Spillovers. The spatial spillovers of the effect of immigration on wage growth, shown in panel B of Table 10, are similar to those on innovation, although they seem more local than the innovation spillovers. Immigration to other counties within the state (column 2) do not have a significant impact on wage growth. Immigration to nearby counties, using an inverse-distance-weighted sum of immigration to other counties, does have a strong and significant impact on local wage growth (column 3), though it appears smaller than that for innovation. Immigration within 100km positively affects local wage growth, with a one standard deviation increase leading to around a 2% increase relative to mean. However, the effect is statistically indistinguishable from zero beyond 100km (column 4).

5.2 Education of Immigrants

We now explore whether more educated immigrants have different impacts on local innovation and wages. First, to measure educational attainment for individuals who might reasonably have had the time to complete their schooling, we limit ourselves to the analysis of immigrants age 25 or older, constructing the endogenous measure of immigration at the county level within this subset of immigrants. We then interact this overall adult immigration flow with the average schooling levels of migrants arriving in a given county at a given time (which we again construct from IPUMS), adding a second endogenous variable to our baseline specification.[25]

To successfully instrument for both the total immigration flow and the interaction of immigration and education, we exploit the fact that different origin countries send migrants with different levels of education to the US at different times. (For example, Japanese immigrants, on average, have about twice the number of years of schooling as those from Guatemala, whereas the education levels of Mexican arrivals increased by about 30% during our sample period.) Our identification strategy allows us to construct a separate instrument for each origin-destination pair and each time period, $\hat{I}^t_{o,d} = \hat{A}^{t-1}_{o,d} \times \tilde{I}^t_{o,-r(d)}$. We disaggregate our baseline instrument to this level, using the predicted immigration shocks for each of the the top 20 origin countries as a joint set of instruments for both total immigration and immigration interacted with the average education level of the new arrivals, so that the first stage for this additional endogenous

[25]IPUMS lists information on the number of years of schooling and the number of years of college education each respondent has received. See Appendix A.1 for details.

variable takes the form

$$Average\ Years\ Education_d^t \times Immigration_d^t = \delta_s + \delta_t + \sum_{o=1}^{20} \kappa_o \hat{I}_{o,d}^t + \nu_d^t.$$

Because migrants arriving from different countries at different times have different schooling levels and emigrate to different counties, we are able to isolate exogenous variations in the level of education of migrants arriving in a given destination at a given point in time. For example, other things equal, an exogenous increase of Japanese arrivals to a given destination implies, on average, a relative increase in the average education level of local immigrants at that point in time.

Innovation and Education Table 11 reports the results of our analysis. The top panel examines heterogeneity in the impact of immigrants on innovation by education level. Column 1 replicates our standard specification for the age 25+ immigration sample, with a positive - now slightly stronger than baseline - impact of immigration on the growth of patenting per capita.[26] Column 2 adds the interaction of immigration with (demeaned) average years of education for immigrants to the same county. The estimates indicate more highly educated immigrants cause a larger increase in innovation. To inspect the magnitude of the heterogeneity at work here, consider two counties, both receiving 10,000 more migrants. A county receiving migrants of average education (about 11 years) would see innovation increase by around $10 \times 0.200 = 2$ more patents per 100,000 people. However, a county receiving the same number of immigrants with one standard deviation extra years of education per person, about 3.7 years, would see $10 \times (0.200 + 3.7 \times 0.221) \approx 10$ more patents per 100,000 people. Column 4 reports a similar analysis, but measuring educational attainment by average years of college completed rather than average years of total schooling. Unsurprisingly, the impact of immigration on innovation increases even more strongly with college attainment than with overall educational attainment. Note columns 2-4 rely on a linear interaction of immigration and education, imposing functional-form restrictions on the link between education and innovation responses. Column 5 instead conducts a nonparametric analysis, separately instrumenting for the immigration into counties receiving migrants with low, medium, and high levels of average education per person by terciles. The more flexible analysis in this column reveals that counties receiving the most highly educated

[26] In column 1 of Table 11 (both panels), we consider a specification with a single endogenous regressor and multiple instruments, and therefore report the first-stage F-statistic developed in Montiel Olea and Pflueger (2013). The remaining columns in this table report results for specifications with multiple endogenous variables and multiple instruments and, to our knowledge, there is no comparable effective F-statistic to report in this case.

immigrants see an order-of-magnitude higher impact on patenting relative to counties receiving medium-education migrants, whereas the impact of the lowest-education immigrants is too noisy to determine.

Wage Growth and Education The bottom panel of Table 11 examines heterogeneity in the impact of immigration on overall wages per capita by education levels, with a structure identical to the top panel. The average immigrant in our sample, with approximately 11 years of education, increases average wages in column 1. Column 2 reveals a higher impact in counties that receive immigrants with a higher average education level. To evaluate magnitudes, once again consider two counties, both receiving 10,000 more migrants. A county receiving migrants of average education would see wages increase by around $10 \times 0.367 \times \$100 = \$367$ more per person over five years (measured at 2010 prices). A county receiving the same number of immigrants with one standard deviation or about 3.7 extra years of education per person would see more wage growth by around $10 \times (0.367 + 3.7 \times 0.352) \times \$100 \approx \$1{,}669$ per person over five years. Column 4 reveals similar - and unsurprisingly stronger - patterns for college education rather than total years of education. And in column 5, a nonparametric analysis splitting education levels into terciles reveals that, just as in the case of patenting, the most highly skilled immigrants have an order-of-magnitude higher impact on local wages per person than moderately educated immigrants, with only noisily estimated impacts from the lowest-educated migrants.

5.3 Growth Models and Population Change

Endogenous growth models often link changes in macro growth to shifts in total population. By contrast, all of our results are local, rather than aggregate, in nature. In our baseline results, we also focus on immigration in particular, rather than on total population change. To provide empirical results in a form more readily digestible by the theoretical literature, Table 12 explores a range of variations on our baseline specification loosely inspired by endogenous growth models.

First, column 1 replicates our baseline results. We then note that if population growth rates determine innovation growth, counties with larger absolute immigration flows should see smaller marginal increments in innovation. Column 2 tests for such concavity by adding the squared immigration flow, instrumenting for this higher-order term with the square of the baseline predicted immigration instrument. The negative coefficients on squared innovation suggests such nonlinearities or concavities are present.

To this point, we have found it convenient to conservatively analyze the impact of immigration on the *change* in patenting in order to flexibly account for any permanent destination-county-specific variation in the levels of patent flows. However, baseline growth models typically relate to the *flow* of patenting rather than its difference. So we explore the implications of a switch to this flow measure, first duplicating our baseline specification and revealing a positive impact of immigration on patent flows (column 3) and evidence of some concavity again (column 4).

The final four columns switch to an alternative unit-free outcome measure, the inverse hyperbolic sine (IHS) of patent flows. The inverse hyperbolic sine function, approximately equivalent to the natural logarithm for non-negative values, allows us to examine the semi-elasticity or elasticity of innovation to various changes. Column 5 reveals a positive semi-elasticity of innovation to immigration. Column 6 reveals a positive semi-elasticity to total population changes, although our instrument, designed to predict immigration rather than overall population change, has less power, as measured by the first-stage F-statistic. Columns 7 and 8 repeat the exercise, instrumenting for the IHS of immigration or population change. The resulting coefficients are interpretable as elasticities, with a 1% increase in immigration inducing 1.7% higher innovation in column 7. Column 8, for which we have the least first-stage power and more noisy estimates, reveals an increase in innovation of around 2.5% after a 1% higher population change.

6 Conclusion

The economic, social, political, and cultural changes immigrants bring to their host communities are often the subject of fierce political controversy. Is immigration an asset or a liability for the receiving communities? Informing this debate with data has often proven difficult, not only because different migrants may affect their host societies in many different and highly heterogenous ways, but also due to an identification problem: immigrants do not randomly allocate across space, but likely select into host communities that offer the best prospects for them and their families. This selection generates endogenous correlation between past and present immigration, ancestry composition, and local economic outcomes, making isolating the causal effects of immigration on these outcomes difficult.

In this paper, we introduced a novel approach to this identification problem that allows the construction of local immigration shocks – instruments for the total number of migrants arriving in each US county for each five-year period since 1975. Importantly, these immigration shocks remain valid even if migrations prior to 1975, and thus the county's pre-existing ancestry composition, are endogenous to local economic activity, and can be flexibly disaggregated to obtain different instruments for migrations from each of the origin countries to each destination county at each point in time.

We use these instruments to document three substantive facts. First, we show that, on average, immigration to the US between 1975 and 2010 had a positive causal effect on local innovation, local economic dynamism, and average wages of natives, where, for example, a 1% increase in immigration to a given county on average increased the number of patents filed by local residents by 1.7% over a five-year period. These positive local economic effects of immigration are consistent with the predictions of a large theoretical literature on endogenous growth but should also be interpreted with caution because they describe the average effect of the average migrant on the economy of an average US county. Second, we find positive spillovers of these positive local effects, where, for example, an increase in a given US county's immigration significantly increases the patenting rate in surrounding areas up to a distance of 250km (150 miles). Third, we find the effect of immigration on innovation and local wages is far from uniform. Instead, we show the arrival of highly educated migrants has a much larger positive effect on local innovation and wages than the arrival of migrants with little or no education.

Although interesting in their own right, we hope these findings, and the exogenous immigration shocks used to generate them, will prove useful to discipline an emerging class of quantitative spatial models of economic growth and to aid future empirical work quantifying the highly complex and heterogenous effects of immigration on a broad range of social and economic outcomes at the local level.

Beyond our application to immigration, we believe our approach linking pre-existing (ancestry) shares to the interaction of historical push and economic pull factors may prove useful in other applications of the canonical shift-share instrument. For example, the cumulative forces that lead to the establishment of migrants of a given ethnicity in a given location over time may be quite similar to the historical forces that generate variation in pre-existing shares of industries, occupations, and other specializations in a given location. Our procedure for isolating quasi-random variation in pre-existing shares may thus prove useful in other settings that have studied the local effects of import competition, the local fiscal multiplier, local supply elasticities, and other important subjects.

References

ABRAMITZKY, R. AND L. BOUSTAN (2017): "Immigration in American economic history," *Journal of economic literature*, 55, 1311–45.

ADÃO, R., M. KOLESÁR, AND E. MORALES (2019): "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 134, 1949–2010.

AGHION, P., U. AKCIGIT, A. DEATON, AND A. ROULET (2016): "Creative destruction and subjective well-being," *American Economic Review*, 106, 3869–97.

AGHION, P., A. BERGEAUD, T. BOPPART, P. J. KLENOW, AND H. LI (2019): "A theory of falling growth and rising rents," National Bureau of Economic Research working paper.

AGHION, P. AND P. HOWITT (1992): "A model of growth through creative destruction," *Econometrica (1986-1998)*, 60, 323.

AKCIGIT, U. AND S. T. ATES (2019): "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," Working paper.

AKCIGIT, U., J. GRIGSBY, AND T. NICHOLAS (2017): "The rise of American ingenuity: Innovation and inventors of the Golden Age," National Bureau of Economic Research working paper.

ALON, T., D. BERGER, R. DENT, AND B. PUGSLEY (2018): "Older and slower: The startup deficit's lasting effects on aggregate productivity growth," *Journal of Monetary Economics*, 93, 68–85.

ANGRIST, J. D. AND J.-S. PISCHKE (2009): *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

BARTIK, T. (1991): *Who Benefits from State and Local Economic Development Policies?*, W.E. Upjohn Institute.

BERNSTEIN, S., R. DIAMOND, T. MCQUADE, AND B. POUSADA (2018): "The Contribution of High-Skilled Immigrants to Innovation in the United States," Working paper.

BLOOM, N., C. I. JONES, J. VAN REENEN, AND M. WEBB (2017): "Are ideas getting harder to find?" National Bureau of Economic Research working paper.

Borjas, G. J. (2003): "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market," *The quarterly journal of economics*, 118, 1335–1374.

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

Borusyak, K., P. Hull, and X. Jaravel (2018): "Quasi-Experimental Shift-Share Research Designs," Working Paper.

Bratsberg, B., A. Moxnes, O. Raaum, K.-H. Ulltveit-Moe, et al. (2019): "Opening the Floodgates: Industry and Occupation Adjustments to Labor Immigration," Tech. rep., CEPR Discussion Papers.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019): "Migrants, Ancestors, and Foreign Investments," *The Review of Economic Studies*, 86, 1448–86.

Card, D. (2001): "Immigrant Inflows, Native outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19, 22–64.

Cortes, P. (2008): "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data," *Journal of Political Economy*, 116, 381–422.

Davis, S. J. and J. Haltiwanger (1999): "Gross job flows," *Handbook of labor economics*, 3, 2711–2805.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014): "The secular decline in business dynamism in the US," Working paper.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018): "The geography of development," *Journal of Political Economy*, 126, 903–983.

Duncan, B. and S. J. Trejo (2017): "The complexity of immigrant generations: Implications for assessing the socioeconomic integration of Hispanics and Asians," *ILR Review*, 70, 1146–1175.

Dustmann, C., U. Schönberg, and J. Stuhler (2017): "Labor supply shocks, native wages, and the adjustment of local employment," *The Quarterly Journal of Economics*, 132, 435–483.

Foged, M. and G. Peri (2016): "Immigrants' effect on native workers: New analysis on longitudinal data," *American Economic Journal: Applied Economics*, 8, 1–34.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018): "Bartik Instruments: What, When, Why and How," Working Paper.

Gordon, R. J. (2012): "Is US economic growth over? Faltering innovation confronts the six headwinds," National Bureau of Economic Research working paper.

Grossman, G. M. and E. Helpman (1991): "Quality ladders in the theory of growth," *The Review of Economic Studies*, 58, 43–61.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001): "The NBER Patent Citaitons Data File: Lessons, Insights, and Methodological Tools," NBER Working Paper 8498.

Hanson, G. H. (2009): "The economic consequences of the international migration of labor," *Annual Review of Economics*, 1, 179–208.

——— (2010): "Future Directions for Research on Immigration," *American Economic Association, Ten Years and Beyond: Economists Answer NSF's Call for Long-Term Research Agendas.*

Hathaway, I. and R. E. Litan (2014): "Declining business dynamism in the United States: A look at states and metros," *Brookings Institution*, 2.

Hopenhayn, H., J. Neira, and R. Singhania (2018): "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," National Bureau of Economic Research working paper.

Hunt, J. and M. Gauthier-Loiselle (2010): "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics*, 2, 31–56.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-share instruments and the impact of immigration," Tech. rep., National Bureau of Economic Research.

Jones, C. I. (1995): "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103.

——— (1999): "Growth: with or without scale effects?" *American Economic Review*, 89, 139–144.

Karahan, F., B. Pugsley, and A. Sahin (2016): "Demographic origins of the startup deficit," New York Fed working paper.

Katz, L. F. and K. M. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107, 35–78.

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

Kerr, W. R. and W. F. Lincoln (2010): "The Supply Side of Innovation: H1B Visa Reforms and U.S. Ethnic Invention," *Journal of Labor Economics*, 28, 473–508.

Khanna, G. and M. Lee (2018): "High-skill Immigration, Innovation, and Creative Destruction," NBER Working Paper 24824.

Klette, T. J. and S. Kortum (2004): "Innovating firms and aggregate innovation," *Journal of Political Economy*, 112, 986–1018.

Lafortune, J., E. Lewis, and T. José (2019): "People and Machines: A Look at the Evolving Relationships Between Capital and Skill in Manufacturing 1860-1930 Using Immigration Shocks," *Review of Economics and Statistics*, 101, 30–43.

Laincz, C. A. and P. F. Peretto (2006): "Scale effects in endogenous growth theory: An error of aggregation not specification," *Journal of Economic Growth*, 11, 263–288.

Lewis, E. (2011): "Immigration, Skill Mix, and Capital Skill Complementarity," *The Quarterly Journal of Economics*, 126, 1029–1069.

——— (2013): "Immigration and production technology," *Annual Review of Economics*, 5, 165–191.

Lewis, E. and G. Peri (2015): "Immigration and the Economy of Cities and Regions," in *Handbook of regional and urban economics*, Elsevier, vol. 5, 625–685.

Monras, J. (forthcoming): "Immigration and wage dynamics: Evidence from the mexican peso crisis," *Journal of Political Economy*.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018): "Commuting, migration, and local employment elasticities," *American Economic Review*, 108, 3855–90.

Montiel Olea, J. L. and C. Pflueger (2013): "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 31, 358–369.

Moser, P., A. Voena, and F. Waldinger (2014): "German Jewish émigrés and US invention," *American Economic Review*, 104, 3222–55.

Ottaviano, G. I. and G. Peri (2012): "Rethinking the effect of immigration on wages," *Journal of the European economic association*, 10, 152–197.

Peretto, P. F. (1998): "Technological change and population growth," *Journal of Economic Growth*, 3, 283–311.

Peters, M. (2019): "Market Size and Spatial Growth – Evidence from Germanys Post-War Population Expulsions," Working paper.

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

Sequeira, S., N. Nunn, and N. Quian (2020): "Immigrants and the making of America," *Review of Economic Studies*, 382–419.

Tabellini, M. (2018): "Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration," Working paper.

Walsh, C. (2019): "Firm Creation and Local Growth," Working paper.

Young, A. (1998): "Growth without scale effects," *Journal of political economy*, 106, 41–63.

Table 1: Summary Statistics by County-Year

	N	mean	sd	iqr
Immigration Flows and Population Change				
Immigration$_d^t$	21,987	1.42	12.21	0.22
Δ Population$_d^t$	21,986	4.03	19.56	2.56
Immigration Shock, $\hat{I}_d^t$	21,987	0.00	4.99	0.24
Patents				
Patenting per 100,000 People	21,987	32.60	94.73	20.66
5-Year Diff. in Patenting per 100,000 People	18,846	4.45	47.83	6.45
5-Year Diff. in Patenting per 100,000 People (Inventors)	18,846	18.85	90.45	41.13
5-Year Diff. in Patenting per 100,000 People (Citation Weighted)	18,846	5.23	63.91	6.04
Dynamism and Wages				
5-Year Diff. in Job Creation Rate	6,600	-32.47	209.90	50.00
5-Year Diff. in Job Destruction Rate	6,600	-17.37	199.58	38.46
5-Year Diff. in Job Growth Rate Skewness	12,564	-6.82	48.91	51.87
5-Year Diff. in Average Annual Wage	21,978	46.07	25.41	25.54
10-Year Diff. in Average Annual Wage of Natives	12,546	10.75	25.80	32.20
10-Year Diff. in Average Annual Wage of Native Non-Movers	6,274	16.85	27.19	33.08
Immigration and Education				
Immigration$_d^t$ (Age 25+)	21,987	0.80	6.91	0.11
Average Years College$_d^t$ (Age 25+)	21,987	1.50	1.41	1.82
Average Years Education$_d^t$ (Age 25+)	21,987	10.88	3.65	4.59
Spillovers				
Immigration$_{State}^t$	21,987	808.00	1,438.90	557.47
Neighbors' Immigration$_d^t$ (Inverse Distance Weight)	21,987	1.15	0.78	0.65
Immigration^t within 100km	21,987	18.58	64.65	9.21
Immigration^t within 250km	21,987	74.96	133.50	67.60
Immigration^t within 500km	21,987	123.10	149.52	143.69

Notes: This table displays the number of observations, mean, standard deviation, and interquartile range for all outcome variables considered, as well as the variables for immigration and the immigration instrument. The first section of the table contains summary statistics for immigration (here we focus only on non-European migration) and population growth in 1000s of people. The second section lists summary statistics for patenting and differences in patenting per 100,000 people. The third section reports summary statistics for dynamism and wages ($100). Finally, the fourth and fifth section provide summary statistics on the immigration variables used in the education and spillovers analyses, respectively. Variables for immigration, population growth, and education are all for five-year periods, as are the differenced outcomes except in the case of differences in average annual wage for natives and native non-movers, which are over 10-year periods.

Table 2: Regressions of Immigration on Push-Pull Instruments at the Country-County Level

	$Immigration^{t}_{o,d}$				
	(1)	(2)	(3)	(4)	(5)
$\hat{A}^{1975}_{o,d} \times \tilde{I}^{1980}_{o,-r(d)}$	0.0036***	0.0036***	0.0035***	0.0035***	0.0035***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\hat{A}^{1980}_{o,d} \times \tilde{I}^{1985}_{o,-r(d)}$	0.0016***	0.0016***	0.0016***	0.0016***	0.0016***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\hat{A}^{1985}_{o,d} \times \tilde{I}^{1990}_{o,-r(d)}$	0.0018***	0.0018***	0.0018***	0.0018***	0.0018***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\hat{A}^{1990}_{o,d} \times \tilde{I}^{1995}_{o,-r(d)}$	0.0005***	0.0005***	0.0005***	0.0005***	0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\hat{A}^{1995}_{o,d} \times \tilde{I}^{2000}_{o,-r(d)}$	0.0004***	0.0004***	0.0004***	0.0004***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\hat{A}^{2000}_{o,d} \times \tilde{I}^{2005}_{o,-r(d)}$	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\hat{A}^{2005}_{o,d} \times \tilde{I}^{2010}_{o,-r(d)}$	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$I^{t}_{Euro,d}$				0.0109***	
				(0.0031)	
$I^{t}_{o,-r(d)} \frac{I^{t}_{Euro,d}}{I^{t}_{Euro}}$					0.3913**
					(0.1558)
N	3,583,881	3,583,881	3,583,881	3,583,881	3,583,881
R^2	0.656	0.657	0.709	0.709	0.709
Distance	No	Yes	Yes	Yes	Yes
Latitude Dis.	No	Yes	Yes	Yes	Yes
Region-Country FE	No	No	Yes	Yes	Yes
County-Continent FE	No	No	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	Yes
Concurrent European Immigration	No	No	No	Yes	No
Contemporaneous Push-Pull	No	No	No	No	Yes

Notes: This table reports coefficient estimates for step 2 of our instrument construction, shown in equation (6), at the country-county level. Moving from column 1 to column 3 we introduce controls for distance and latitude distance and then fixed effects into the regression specification. Column 4 adds contemporaneous European migration as a control while column 5 instead introduces the contemporaneous push-economic pull factor for non-European migration. Standard errors are clustered by country for all specifications and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3: County-Level Panel Regressions of Difference in Patenting on Immigration

	(1)	(2)	(3)	(4)
Panel A: Second Stage	*5-Year Difference in Patenting per 100,000 People Post 1980*			
	(OLS)	(IV)	(IV)	(IV)
Immigration_d^t	0.167**	0.101***	0.100***	0.108***
	(0.080)	(0.031)	(0.032)	(0.033)
N	18,846	18,846	18,840	18,846
Panel B: First Stage	$\textit{Immigration}_d^t$			
Immigration Shock ($\hat{I}_d^t$)		2.119***	2.124***	1.610***
		(0.070)	(0.075)	(0.175)
F-Stat		911	807	85
R^2		0.762	0.766	0.956
Geography FE	State	State	State	County
Time FE	Yes	Yes	Yes	Yes
State-Time FE	No	No	Yes	No

Notes: Panel A of this table reports the results of our IV specification, described in equation (1), where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) in county d in the five-year period ending in t and the endogenous variable is non-European immigration (1,000s) in d and period t. Panel B reports the results for step 3 of instrument construction, or the coefficient estimates for the first-stage specification for non-European immigration (1,000s) for the instrument described in equation (7). Column 1 provides the results of the OLS estimation of equation (1), whereas columns 2-4 provide an IV estimate of the second stage (panel A) and first stage (panel B). The table includes the first-stage F-statistic on the excluded instrument for each of the IV specifications. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Robustness - Alternative Instruments

Specification:	*5-Year Difference in Patenting per 100,000* *Ancestry in 1975 Only*	 *Leave-Out Correlated Counties*	 *Leave-Out Own Continent*
	(1)	(2)	(3)
Immigration$_d^t$	0.093***	0.098***	0.094***
	(0.027)	(0.033)	(0.027)
N	18,846	18,846	18,846
First Stage F-Stat	1,171	127	830
Geography FE	State	State	State
Time FE	Yes	Yes	Yes

Notes: This table displays the results of estimating equation (1), where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) and the endogenous variable is non-European immigration (1,000s) to d in t. In this table, each column utilizes the same approach for instrument construction as the main instrument but with one adjustment. Column 1 replaces predicted ancestry in $t-1$ with predicted ancestry in 1975 for all periods. Column 2 uses an alternative leave-out strategy in Step 1: the push factor excludes all destination counties whose overall time series of immigration flows are correlated with those of d (as opposed to excluding counties in the same census division ($r(d)$) as d). Column 3 replaces the economic pull factor in Step 1 with the share of all migrants who settle in d but excluding migrants from the same continent as o (instead of using only European migrants). We report the first-stage F-statistic on the excluded instrument for each specification. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Robustness - Construction of the Baseline Instrument and Shares Instrument

	5-Year Difference in Patenting per 100,000 People Post 1980		
Specification:	*Predicted Ancestry Shares*	*Realized Ancestry Shares*	*Realized Ancestry No Leave-Out*
	(1)	(2)	(3)
Immigration$_d^t$	0.195**	0.106***	0.132**
	(0.090)	(0.035)	(0.055)
N	18,846	18,846	18,846
First Stage F-Stat	656	265	361
Adão et al (2019) First Stage False Rejection Rate:	4.5	28.2	28.2
Instrument Functional Form:			
Instrumented Ancestry	Yes	No	No
Push Factor Leave-Out	Yes	Yes	No
Controls:			
Geography FE	State	State	State
Time FE	Yes	Yes	Yes

Notes: This table displays the results of estimating equation (1), where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) and the endogenous variable is non-European immigration (1,000s) to d in t. Column 1 is a Card-style instrument but replaces realized ancestry shares with predicted ancestry shares. Column 2 utilizes the same instrument as column 1 but with realized ancestry shares. Finally, column 3 takes the instrument in column 2 but removes the leave-out (as well as the regional adjustment) in the push factor as in the traditional Card-style shift-share instrument. We report the first-stage F-statistic on the excluded instrument for each specification. For each instrument, we report the false rejection rate in the first-stage regression for a robustness test that follows the method proposed by Adão et al. (2019). See Appendix Tabel 5 for details. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Robustness - Additional Controls from Baseline Year (1970)

	5-Year Difference in Patents per 100,000 People for 1980 to 2010					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigration_d^t	0.101***	0.102***	0.100***	0.092***	0.082***	0.108***
	(0.031)	(0.032)	(0.031)	(0.029)	(0.027)	(0.033)
Population Density (1970)		-0.001				
		(0.004)				
Patents per 1,000 People (1975)			0.089**			
			(0.042)			
Share High School Education (1970)				27.821**		
				(11.059)		
Share 4+ Years College (1970)					103.990***	
					(29.961)	
N	18,846	18,846	18,846	18,846	18,846	18,846
First Stage F-Stat	911	1,658	911	945	1,017	85
Geography FE	State	State	State	State	State	County
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) and the endogenous variable is non-European immigration (1,000s) to d in t. Column 1 repeats our main specification, whereas columns 2-5 add as a control county d's population density in 1970, patents per 1,000 people in 1975 (1970 population is used to match the dependent variable), share of high school educated, and share of the population with 4+ years of college, respectively. Finally, column 6 then adds a county fixed effect. We report the first-stage F-statistic on the excluded instrument for each specification. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Robustness - Alternative Samples

	Difference in Patenting per 100,000 People Post 1980				
	Mexico	*China*	*India*	*Philippines*	*Vietnam*
	(1)	(2)	(3)	(4)	(5)
Panel A: Excluding Given Country					
Immigration_d^t	0.080***	0.102***	0.101***	0.100***	0.101***
	(0.025)	(0.032)	(0.031)	(0.031)	(0.031)
N	18,846	18,846	18,846	18,846	18,846
First Stage F-Stat	666	1,576	1,267	1,261	1,179
Panel B: Including Only Given Country					
Immigration_d^t	0.103***	0.068**	0.129***	0.133**	0.123**
	(0.032)	(0.032)	(0.032)	(0.051)	(0.060)
N	18,846	18,846	18,846	18,846	18,846
First Stage F-Stat	2,094	535	318	22	2
Geography FE	State	State	State	State	State
Time FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), run on alternative samples where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) and the endogenous variable is non-European immigration (1,000s) to d in t. In instrument construction, each column either drops migrants from the given country (Panel A) or drops all other migrants except those from the specified country (Panel B) from the sum in equation (7) for each of the five largest sending countries post 1975 (Mexico, China, India, Philippines, and Vietnam). We report the first-stage F-statistic on the excluded instrument for each specification, and note the instrument constructed using only migrants from Vietnam does not significantly predict non-European immigration. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8: Robustness - The Effect of Immigration on Long Differences in Innovation

	Difference in Patenting per 100,000 People			
	ΔPat_{t-2}^{t-1}	ΔPat_{t-1}^{t}	ΔPat_{t-1}^{t+1}	ΔPat_{t-1}^{t+2}
	(1)	(2)	(3)	(4)
$Immigration_d^t$	-0.099	0.108***	0.369***	0.332**
	(0.069)	(0.033)	(0.098)	(0.137)
N	15,705	18,846	15,705	12,564
First Stage F-Stat	80	85	11	7
Geography FE	County	County	County	County
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), for changes in patenting per 100,000 people with non-European immigration to d in t as the endogenous variable. Column 1 uses the one-period lag of the dependent variable. Column 2 repeats the standard specification (5-year change in patenting). Columns 3 and 4 then utilize the two-period (10-year) and three-period (15-year) change in patenting as the dependent variable, respectively. We report the first-stage F-statistic on the excluded instrument for each specification. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 9: Immigration and Economic Dynamism

	5-Year Difference in:				*10-Year Difference in Avg. Annual Wage:*	
	Job Creation Rate	*Job Destruction Rate*	*Job Growth Rate Skewness*	*Average Annual Wage*	*Native*	*Native Non-Mover*
	(1)	(2)	(3)	(4)	(5)	(6)
Immigration_d^t	0.176***	0.152***	0.019***	0.112***	0.049***	0.056***
	(0.033)	(0.035)	(0.004)	(0.036)	(0.016)	(0.020)
N	6,600	6,600	12,564	21,978	9,411	6,274
First Stage F-Stat	951	951	151	1,202	750	1,178
Geography FE	State	State	State	State	State	State
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), for each of our dependent variables with non-European immigration (1,000s) to d in t as the endogenous variable. Columns 1 and 2 report the results of our second stage with the job creation rate and job destruction rate as the dependent variable, respectively. Column 3 then provides results for job growth rate skewness as the dependent variable, whereas the dependent variable for the specification shown in column 4 is the change in the average annual real wage ($100s, at 2010 prices) over the five-year period ending in t. Columns 5 and 6 reports results of a regression of the change in the average annual real wage ($100s) for natives and native non-movers over the 10-year period ending in t on instrumented non-European immigration for the 10-year period ending in t. We report the first-stage F-statistic on the excluded instrument for each specification. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 10: Spillovers Analysis

	(1)	(2)	(3)	(4)
Panel A	*5-Year Difference in Patenting per 100,000 People Post 1980*			
$\text{Immigration}_{d}^{t}$	0.130***	0.107***	0.072**	0.080**
	(0.039)	(0.035)	(0.032)	(0.037)
$\text{Immigration}_{State}^{t}$		0.001***		
		(0.000)		
Neighbors' $\text{Immigration}_{d}^{t}$ (Inverse Distance Weight)			6.600***	
			(1.593)	
$\text{Immigration}_{100km}^{t}$				0.056***
				(0.018)
$\text{Immigration}_{250km}^{t}$				0.014***
				(0.005)
$\text{Immigration}_{500km}^{t}$				0.006
				(0.005)
N	18,846	18,846	18,846	18,846
First Stage F-Stat (first coefficient)	876	1,792	2,175	6,065
First Stage F-Stat (second coefficient)		470	162	383
First Stage F-Stat (third coefficient)				150
First Stage F-Stat (fourth coefficient)				66
Panel B	*5-Year Difference in Average Annual Wage ($100) Post 1975*			
$\text{Immigration}_{d}^{t}$	0.124***	0.111***	0.077***	0.071***
	(0.042)	(0.036)	(0.021)	(0.026)
$\text{Immigration}_{State}^{t}$		0.001*		
		(0.000)		
Neighbors' $\text{Immigration}_{d}^{t}$ (Inverse Distance Weight)			5.502***	
			(1.278)	
$\text{Immigration}_{100km}^{t}$				0.067***
				(0.009)
$\text{Immigration}_{250km}^{t}$				0.007
				(0.006)
$\text{Immigration}_{500km}^{t}$				0.004
				(0.007)
N	21,978	21,978	21,978	21,978
First Stage F-Stat (first coefficient)	1,166	2,312	3,487	7,969
First Stage F-Stat (second coefficient)		437	165	394
First Stage F-Stat (third coefficient)				156
First Stage F-Stat (fourth coefficient)				66
Geography FE	Division	Division	Division	Division
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification (1) for the change in patenting per 100,000 people (population is based on baseline 1970 levels) (Panel A) and the change in the real average annual wage ($100s, at 2010 prices) (Panel B) with non-European immigration (1,000s) to d in t as the endogenous variable. The first column repeats our baseline specification but with census division fixed effects. Column 2 adds as a second endogenous variable: total non-European immigration to the state in which d is located, excluding own-immigration to d, in period t and a comparable instrument. Column 3 adds as a second endogenous variable the inverse-distance-weighted sum of non-European immigration to all counties in the US, excluding own-immigration, and an instrument constructed analogously. Column 4 includes variables, and appropriate instruments, for non-European immigration to counties within 100km (excluding d), 100km to 250km, and 250km to 500km of county d. For each specification we report the first-stage F-statistic(s), utilizing the F-statistic described in Angrist and Pischke (2009, p. 217-218) in the case of multiple endogenous variables. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 11: Education Analysis

	(1)	(2)	(3)	(4)	(5)
Panel A	*5-Year Difference in Patenting per 100,000 People*				
Immigration_d^t	0.166***	0.200***	0.485***	0.415***	
	(0.053)	(0.070)	(0.165)	(0.076)	
$\text{Average Years Education}_d^t \times \text{Immigration}_d^t$		0.221***	0.251***		
		(0.068)	(0.079)		
$\text{Average Years College}_d^t \times \text{Immigration}_d^t$				0.887***	
				(0.166)	
$\mathbf{1}\{\text{Low Avg. Years Education}\} \times \text{Immigration}_d^t$					1.863
					(4.539)
$\mathbf{1}\{\text{Medium Avg. Years Education}\} \times \text{Immigration}_d^t$					0.084*
					(0.044)
$\mathbf{1}\{\text{High Avg. Years Education}\} \times \text{Immigration}_d^t$					1.401*
					(0.792)
N	18,846	18,846	18,846	18,846	18,846
Panel B	*5-year Difference in Average Annual Wage ($100s)*				
Immigration_d^t	0.289**	0.367***	0.423*	0.617***	
	(0.115)	(0.065)	(0.229)	(0.155)	
$\text{Average Years Education}_d^t \times \text{Immigration}_d^t$		0.352***	0.315**		
		(0.078)	(0.119)		
$\text{Average Years College}_d^t \times \text{Immigration}_d^t$				1.159***	
				(0.313)	
$\mathbf{1}\{\text{Low Avg. Years Education}\} \times \text{Immigration}_d^t$					-0.060
					(0.091)
$\mathbf{1}\{\text{Medium Avg. Years Education}\} \times \text{Immigration}_d^t$					0.160**
					(0.062)
$\mathbf{1}\{\text{High Avg. Years Education}\} \times \text{Immigration}_d^t$					2.534***
					(0.752)
N	21,978	21,978	21,978	21,978	21,978
Geogrpahy FE	State	State	County	State	State
Time FE	Yes	Yes	Yes	Yes	Yes

Notes: The table reports the results of our IV specification (1) for the change in patenting per 100,000 people (population is based on baseline 1970 levels) in Panel A and the 5-year difference in county-level average real annual wages ($100s) in Panel B. Column 1 repeats our main specification but adjusting the migrant pool to those aged 25+ (1,000s). Columns 2 and 3 then add a second endogenous variable for the interaction of immigration with the (demeaned) average years of education of the migrants arriving in the destination county, whereas column 4 adds (demeaned) average years of college education of those migrants. Repeating the regression in column 2 of the second panel for the 10-year difference in average annual wages ($100s) of native non-movers (US-born working individuals who have not moved outside of the county within the past 5 years) on 10-year migration and corresponding education results in coefficients of 0.246 (0.057) and 0.142 (0.040) on immigration and average years of education times immigration, respectively. Column 5 uses as endogenous variables adult immigration interacted with indicators for the terciles of average years of education of migrants across counties in period t. In all specifications, for instrumentation, we exploit the fact that in our initial instrument construction we created quasi-exogenous immigration shocks for each origin country-o $\times$ destination county-d pair in each time period t; each specification utilizes the predicted immigration shocks for each of the the top 20 origin nations as a joint set of instruments. For column 1, the Montiel Olea and Pflueger (2013) effective F-statistic is 39 (critical value 32 for τ of 5%) for the first panel and 40 (critical value 31 for τ of 5%) for the second panel. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 12: Growth Models and Population Change

	Difference in Patenting per 100,000 People Post 1980		*Patenting per 100,000 People Post 1975*		*IHS of Patenting Post 1975*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigration_d^t	0.101***	0.509***	0.501**	2.505***	0.028***			
	(0.031)	(0.090)	(0.190)	(0.268)	(0.011)			
$\text{sq}(\text{Immigration}_d^t)$		-0.001***		-0.004***				
		(0.000)		(0.000)				
$\Delta\ \text{Population}_d^t$						0.033***		
						(0.012)		
$\text{IHS}(\text{Immigration}_d^t)$							1.723***	
							(0.111)	
$\text{IHS}(\Delta\ \text{Population}_d^t)$								2.471***
								(0.510)
N	18,846	18,846	21,987	21,987	21,987	21,986	21,987	21,986
First Stage F-Stat (first coefficient)	911	95	1,202	102	1,202	102	94	16
First Stage F-Stat (second coefficient)		11,231		11,879				
Geography FE	State	State	State	State	State	State	State	State
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), for changes in patenting per 100,000 people (columns 1 and 2), patenting per 100,000 people (columns 3 and 4), and the inverse hyperbolic sine (IHS) of patenting (columns 5-8), where patenting is defined as the number of patents filed in the five-year period. Column 1 repeats our main specification while column 2 adds as a second endogenous variable the square of non-European immigration (1,000s) to d in t. Columns 3 and 4 repeat the specifications in columns 1 and 2 but with patenting as the dependent variable. Finally, columns 5 and 6 include as endogenous variables non-European immigration and population change in d at t, respectively, whereas columns 7 and 8 include as endogenous variables the IHS of non-European immigration and population change in d at t, respectively. For each specification, we report the first-stage F-statistic(s), utilizing the F-statistic described in Angrist and Pischke (2009, p. 217-218) in the case of multiple endogenous variables. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

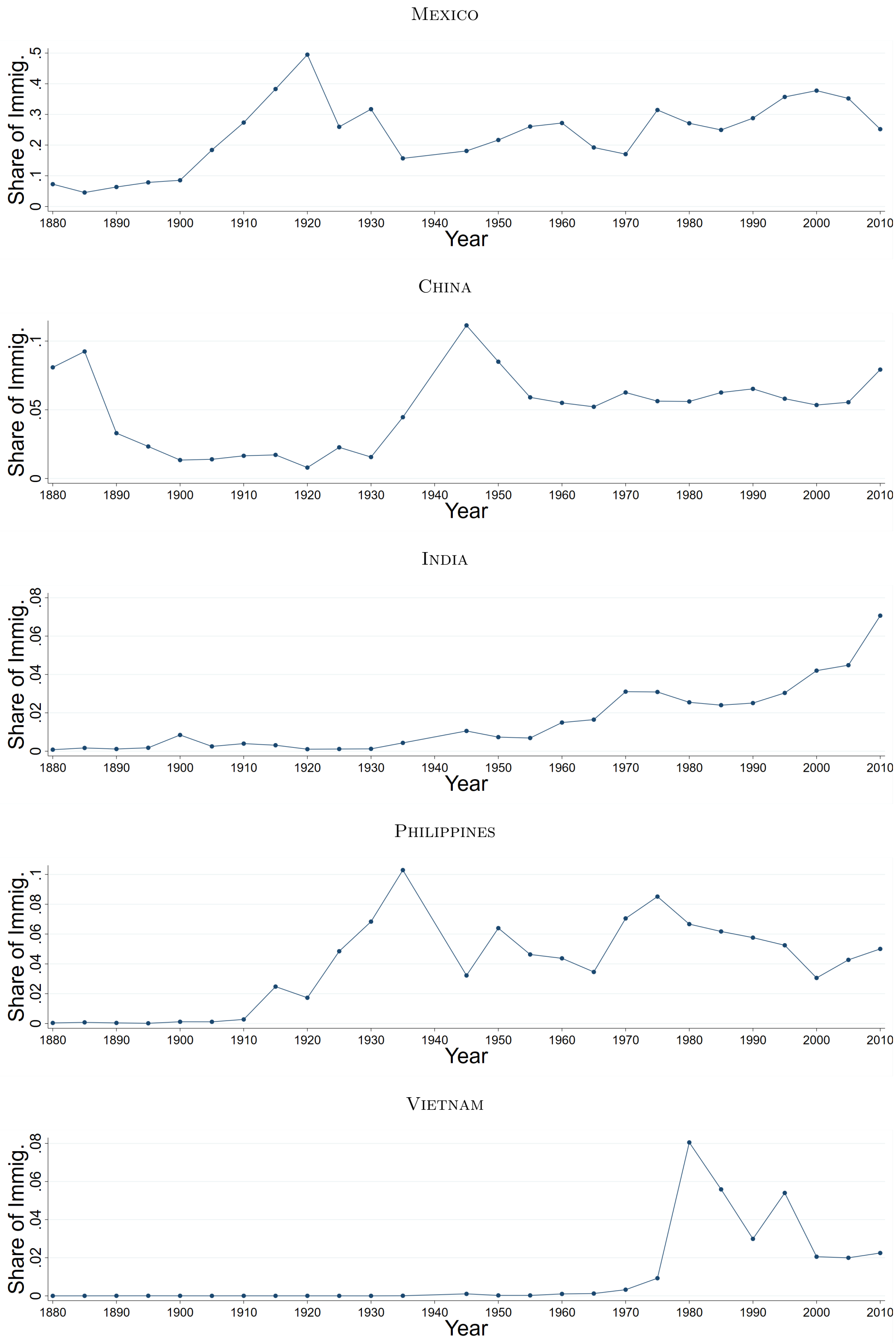


Figure 1: Share Non-European Immigrants to the US by Origin Country

Notes: This figure plots the share of non-European immigration into the US from the 5 non-European origin nations with the largest cumulative immigration to the US: Mexico, China, India, Philippines, and Vietnam. The figure highlights variation in the push factor, showing the number of migrants from a given source country o to the US varies by period t.

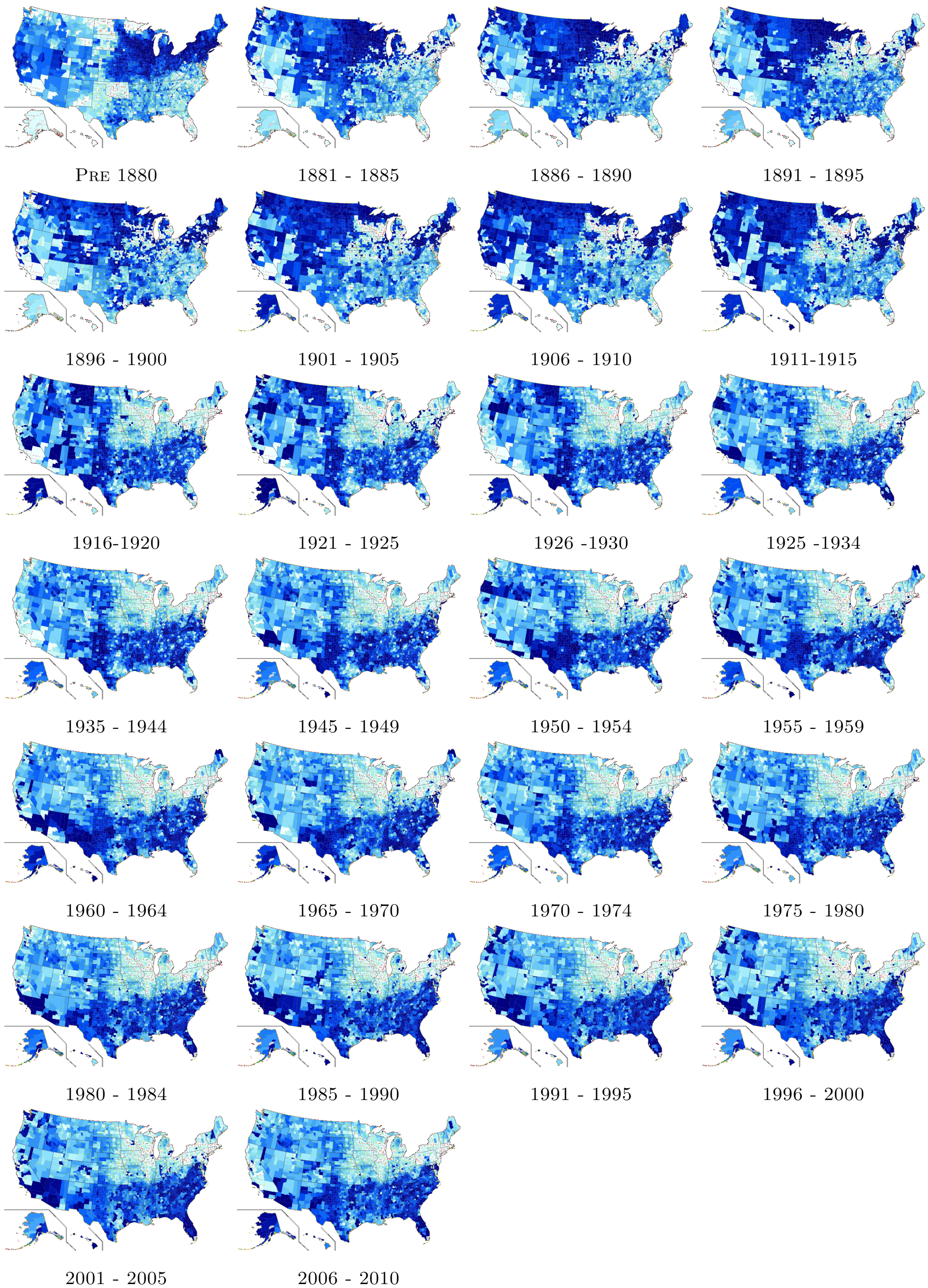


Figure 2: Destinations of European Immigrants to the US

Notes: This figure maps immigration flows into US counties by 5-year periods (except between 1930 and 1950). We regress the number of European immigrants into US county d at time t, I_d^t, on destination county d and year t fixed effects, and calculate the residuals. The map's color coding depicts the 20 quantiles of the residuals across counties and within census periods. Darker colors indicate a higher quantile.

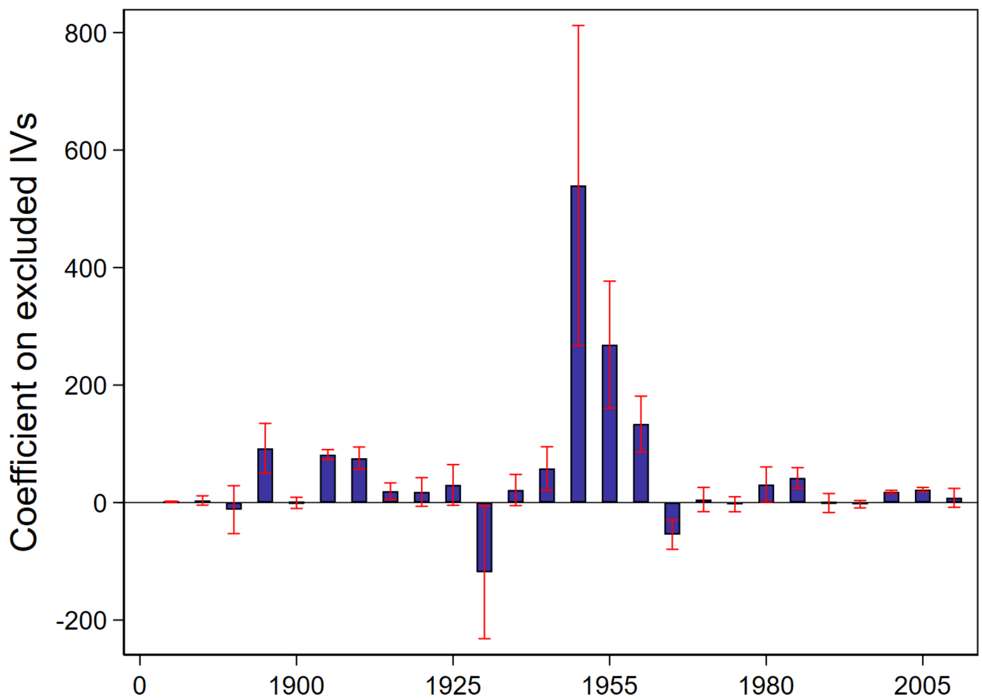
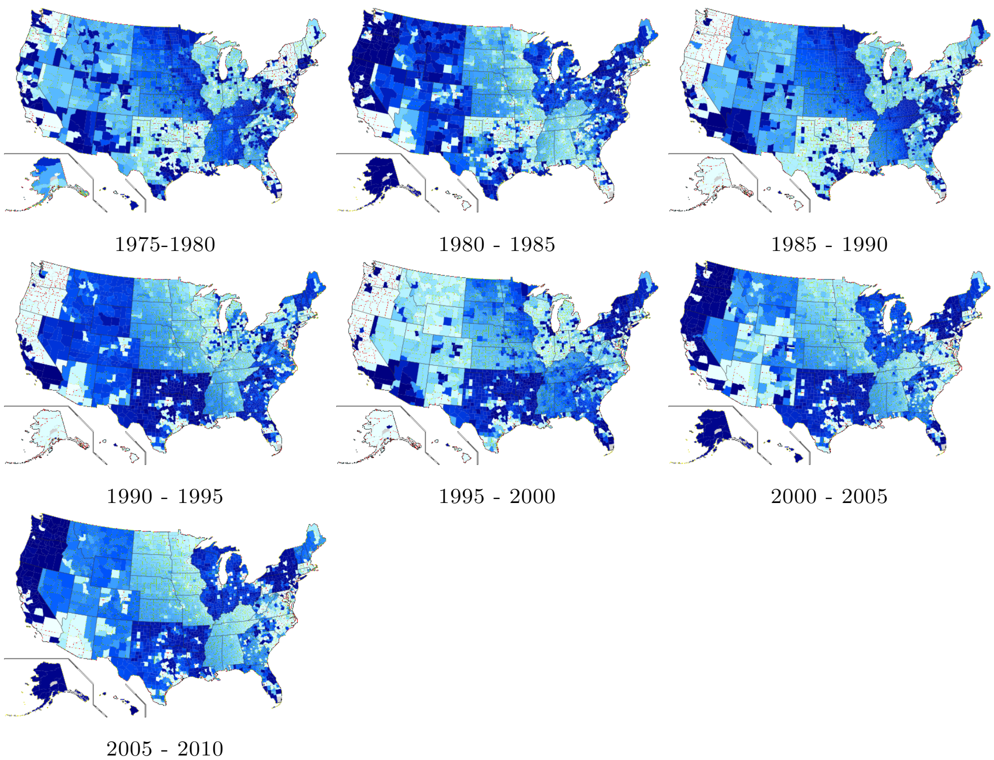


Figure 3: Step 1 – Predicting Ancestry

Notes: This figure displays the coefficients (bars) and 95% confidence intervals (red lines) in the ancestry prediction regression, equation (5), for estimating 2010 reported ancestry (assuming for presentational purposes only that $a^{\tau}_{r(d)} = a^{\tau}\ \forall r(d)$). The figure shows we identify variation in current ancestry levels based on push-economic pull interactions from the full range of time periods in our sample. Standard errors are clustered at the origin country level.

FIGURE 4: STEP 1 – PREDICTING ANCESTRY (2010)

Notes: This figure plots actual ancestry in 2010 against predicted ancestry, as given in equation (5), with the size of each circle indicating the log number of observations in a given bin of predicted ancestry. The labeled counties are those with the highest number of individuals declaring a given ancestry in 2010. The corresponding regression of $A^{2010}_{o,d}$ on $\hat{A}^{2010}_{o,d}$, as defined in equation (5), yields an R^2 of 74.9%.



1975-1980 | 1980 - 1985 | 1985 - 1990

1990 - 1995 | 1995 - 2000 | 2000 - 2005

2005 - 2010

Figure 5: Immigration Shock Conditional on County and Time FE

Notes: This figure maps the instrumented non-European immigration flows into US counties by 5-year periods. We regress the instrument for immigration into US county d at time t on county and state-year fixed effects, and calculate the residuals. This figure provides a visualization for the immigration shocks used as in instrument in the regression shown in column 3 of Table 3. The map's color coding depicts the 200 quantiles of the residuals across counties and within census periods. Darker colors indicate a higher quantile.

Online Appendix

"Immigration, Innovation, and Growth"

Konrad B. Burchardi

Thomas Chaney

Tarek A. Hassan

Lisa Tarquinio

Stephen J. Terry

A Data Appendix

A.1 Details on the construction of migration and ethnicity data

To construct county-level data on migration, ancestry, and ethnicity, we follow the approach of Burchardi et al. (2019). We utilize data from each available IPUMS wave from 1880 to 2010. Specifically, we use the 10% sample of the 1880 Census, the 5% sample of the 1900 Census, the 1% sample of the 1910 Census, the 1% sample of the 1920 Census, the 5% sample of the 1930 Census, 1% Form 1 Metro sample of the 1970 Census, 5% State sample of the 1980 Census, 5% State sample of the 1990 Census, 5% sample of the 2000 Census, and the American Community Service 5-Year sample of the 2010 Census. The following section summarizes this approach, highlighting any difference in data construction made in this paper.

Construction of post-1880 immigration flows

We start the construction of our immigration variable by identifying the number of individuals located in a given US geography d at the time of each census who immigrated to the US since the prior census and were born in a historic origin country o (based on the detailed birthplace variable). For each census wave, we then separate this immigration count into (roughly) five-year periods based on the year in which each migrant arrived to the US. For the 1970, 1980, and 1990 censuses, the exact year of arrival for immigrants is not provided, and instead the year of arrival is provided in bins (e.g., a person who arrived in 1964 has a year of arrival of 1960-1964). For these years, we use as our five-year periods the bins that are reported in each census: 1925-34, 1935-44, 1945-49, 1950-54, 1955-59, 1960-64, 1965-70, 1970-74, 1975-80, 1980-84, and 1985-90. We then follow the approach outlined in Burchardi et al. (2019) to transform foreign origin countries, given as birthplaces, to 1990 foreign countries and non-1990 counties and county groups into 1990 counties. Because some foreign birthplaces do not refer to any modern (1990) country, we use population-based weights for transitioning birthplaces to countries (for more details on the weighting scheme, see Burchardi et al. (2019)).

Construction of pre-1880 immigration stock

From the 1880 census, we count all individuals who were born in a foreign origin country o and reside in a historic US geography d, regardless of the date of arrival to the US. We then add to this count all individuals residing in d who were born in the US but whose parents were born in origin country o (if an individual's parents were born in different countries, the individual is assigned a count of one half for each parent's origin country o). We then transform the given birthplace to 1990 foreign countries and the pre-1880 US geography to 1990 US counties following the transition method outlined in Burchardi et al. (2019).

Construction of ancestry stock

For the years 1980, 1990, 2000, and 2010, we take from the respective census all individuals in a US county or county group that list as their primary ancestry a foreign nationality or area. We then estimate the ancestry stock in each midyear (1975, 1985, 1995, and 2005) by taking the individuals identified in each census year as belonging to a given ancestry and removing all individuals who either were born or migrated to the US after the midyear. Ideally, we would also remove all individuals who moved to the county after the midyear, but data is not available for all census years; thus, for consistency, we do not remove these individuals. Again, we follow Burchardi et al. (2019) in transforming ancestries to 1990 countries and US geographies to 1990 US counties. As with the data on foreign birthplaces, some ancestries do not correspond directly to a modern (1990) country; again, we follow the weighting scheme outlined in Burchardi et al. (2019) for transitioning stated ancestries to 1990 foreign countries.

Construction of education data for migrants

For the five-year migration periods from 1975 to 2010, whose construction is previously described, we also identify the total number of years of education for each set of immigrants. Specifically, we take the set of individuals that make up each five-year immigration flow and limit to those individuals who are aged 25 years or older at the time of each respective census. For each 1990 US county d, we then sum the number of years each individual is reported to have over all immigrants in this set, assigning the midpoint when a range of years of education is provided instead of an exact number of years. We then generate the average years of education for immigrants to county d in each period t and demean these values. Finally, we take the demeaned average years of education and multiply by the count of immigrants aged 25 or older to generate the (demeaned) total years of education. We construct this variable for total years of education as well as for years of college education.

We also utilize information on education from the census to construct county-level demographic controls for the share of the county's population that has a specified level of education in a baseline year, 1970. Using data from the 1970 census, we calculate the share of all individuals, regardless of birthplace, residing in a historic US county d who report having at least a Grade 12 education (share of high-school educated) and those who report having at least four years of college education (share of college educated). These values are then transformed from 1970 US counties to 1990 US counties, again using the transition matrices described by Burchardi et al. (2019).

A.2 Construction of population data

For the period 1970 to 2010, we collect county-level population data in each census year and intercensal year. Population counts for the 2010 and 2000 census were taken directly from the US Census Bureau. The 1990 census data are taken from the US Census Bureau's Population of Counties by Decennial Census: 1900 to 1990. The 1970 and 1980 data are taken from Inter-University Consortium for Political and Social Research (ICSPR) County Population by Singe Years of Age, Sex, Race, 1970, 1980, 1990. All intercensal population counts are taken from the NBER's Census U.S. Intercensal County Population Data, 1970-2014. For each period, data are transformed from the given US counties to 1990 US counties using the transition matrices described by Burchardi et al. (2019).

A.3 Construction of patenting data

We utilize data on corporate utility patents with a US assignee from the the US Patent and Trademark Office microdata for the period 1975 to 2010. We translate the location of patents from assignee (or inventor) location to 2010 US counties and then transition to 1990 counties using area weights as in Burchardi et al. (2019) to estimate the number of patents granted to assignees in each county and year. For our main measure of patenting, we utilize unweighted patent counts with locations based on assignee, but we also consider location based on inventors and weighted patent counts as in Hall et al. (2001). We then construct a variable for the total number of patents filed in each five-year period ending in t, for each measure of patenting, and divide by the 1970 population (100,000 people) to get "per-capita patenting" in t. We then winsorize the variables at the 1% and 99% levels. The main patenting outcome variable is then the difference in this per-capita-patenting variable between $t-1$ and t.

A.4 Construction of business dynamism data

In this section, we explain the construction of variables used to measure business dynamism. In each case, we take the five-year difference in the dynamism or wage variable.

Wages. The county-level average annual wage for every five years from 1975 to 2010 is taken from the Quarterly Census of Employment and Wages. The data for each period are then transformed from the US counties for that period to 1990 US counties using the transition matrices developed in Burchardi et al. (2019) and then converted to 2010 US dollars using the Personal Consumption Expenditures Price Index from the Bureau of Economic Analysis. We generate this county-level average annual wage for all industries as well as manufacturing (SIC 20-39 and NAICS 31-33) and services (SIC 60-67 and NAICS 52-53).

Growth Rate Skewness. The growth rate skewness variable for 2010 US counties for each five years from 1995 to 2010 is estimated using data from the Longitudinal Business Database. We compute the Kelly Skewness of employment growth rates across 4-digit sectors, and then transition this measure from 2010 to 1990 US counties.

Job Creation and Destruction Rates. Job creation and destruction data are taken from the Business Dynamics Statistics for metropolitan statistical areas (MSAs) and transitioned to 1990 US counties based on weights derived from 1990 population data.

A.5 Construction of native wages data

We construct variables for native wages in each census year from 1970 to 2010 using data from the 1970 1% Form 1 Metro sample, 1980 5% State sample, 1990 5% State sample, 2000 5% Census sample, and 2010 American Community Service (ACS). In each year, we limit the sample to the pre-tax wage and salary income (incwage) for individuals born in the US who are employed (empstat is equal to 1), referred to here as natives. For the census years 1980 to 2000, we also generate a wage measure for the subset of natives who report that they lived in the same county five years prior to the census year, referred to as native non-movers. We use the Consumer Price Index provided in IPUMS USA (CPI99) to adjust wages to a common dollar year, 1999. We then follow the same method as that used in Burchardi et al. (2019) to transform wages for county groups into 1990 US counties. Finally, we determine average wages in each county using the person weight (PERWT) for the selected sample and generate a variable for wage growth in each county that is the 10-year difference in average annual wages for natives (or native non-movers).

B Growth, Population Growth, Innovation, and Dynamism

In this appendix, we sketch out a deliberately simple theoretical mechanism linking innovation, income growth, dynamism, and population growth. We present the minimum ingredients needed from a combination of the "semi-endogenous growth" model outlined in Jones (1995) and the micro-level distribution of creative destruction from Schumpeterian growth models (Aghion and Howitt, 1992; Grossman and Helpman, 1991; Klette and Kortum, 2004). We show that in such a model, the long-run balanced growth path per-capita growth rate of the economy must be proportional to the growth rate of labor input in the economy and that the economy-wide growth rate links positively to the rates of creative destruction and innovation at the micro level. To the extent that local economies are segmented, these two outcomes concisely motivate our empirical analysis linking population dynamics to measures of scaled innovation, dynamism rates, and income growth, abstracting from cross-economy spillovers and heterogeneity in labor input, both of which we nevertheless explore empirically.

B.1 Environment

Final Goods Production We examine a closed local economy in continuous time t. Final output Y_t is produced according to the technology

$$\log Y_t = \int \log y_{jt} dj$$

utilizing a unit mass of intermediate varieties j.

Intermediate Goods Production Intermediate goods are each produced with a symmetric technology combining production labor l^P_{jt} and variety-specific quality q_{jt}, with $y_{jt} = q_{jt} l^P_{jt}$. Incumbent intermediate goods firms f produce portfolios of intermediate varieties j for which they operate the current leading-edge quality level q_{jt}. Let $\log Q_t = \int \log q_{jt} dj$ be the average quality level in the economy.

Innovation For an individual variety, innovation is embodied in an instantaneous increase in the quality level q_{jt} in that good's production, that is, a switch from q_{jt} to $q_{jt+\Delta} = \lambda q_{jt}$, where $\lambda > 1$ is a quality ladder or innovation step size. Incumbent firms f may innovate by hiring labor for innovation in the amount s^I_{ft} to guarantee an innovation arrival rate p^I_{ft} satisfying

$$p^I_{ft} \propto {s^I_{ft}}^{\gamma} Q_t^{-\alpha},$$

where $\alpha, \gamma > 0$. A mass of potential entrants each hires labor for innovation s_t^E to guarantee an innovation arrival rate p_t^E satisfying

$$p_t^E \propto s_t^{E\gamma} Q_t^{-\alpha}.$$

In both of the innovation technologies, innovation arrival probabilities depend positively on innovation input – labor – but negatively on the current average quality level in the economy Q_t. Solving harder problems to improve upon a higher existing average quality level requires more input. When an innovation occurs, for either an entrant or incumbent, they become the leading-edge incumbent producer of a random variety.

Labor Input The exogenous instantaneous growth rate of labor input or the population of the economy L_t is n, and total labor input in any period must equal the sum of the total amounts of labor used for production, incumbent innovation, and entrant innovation:

$$L_t = L_t^P + S_t^I + S_t^E.$$

B.2 Balanced Growth

A range of straightforward and standard additional machinery needed for description of a decentralized equilibrium along a stationary balanced growth path – along the lines of the equilibria described in Klette and Kortum (2004) or Grossman and Helpman (1991) – could be added to the framework already outlined above. But we do not need additional elements for our desired implications. Instead, we simply note that in standard decentralizations, output per capita is proportional to the average quality level Q_t. We also note that along any stationary balanced growth path in this economy, by definition, constant output growth rates, constant quality growth rates, constant ratios of production labor and innovation labor to total labor input, and a stationary distribution of outcomes at the firm and variety levels must exist.

But then note that constant quality growth rates and constant innovation rates for incumbents and entrants - given the innovation technologies - imply

$$Q_t^\alpha \propto S_t^{I\gamma} \propto S_t^{E\gamma} \propto L_t^\gamma \rightarrow \alpha g_Q = \gamma n \rightarrow g_Q = \frac{\gamma}{\alpha} n.$$

In other words, average quality growth, which is equal to per-capita growth in this economy, must be positively proportional to the population growth rate n. This is our first desired result, echoing Jones (1995). Then, given the definition of average quality Q_t, the implication of a constant growth rate $g_Q = \frac{\partial \log Q_t}{\partial t}$ is that

$$g_Q = p \log \lambda,$$

where $p = p^I + p^E$ is the sum of the constant incumbent and entrant innovation rates, and λ is the quality ladder step size described above. But note that

$$p = \mathbb{P}(\text{Innovation}) = \mathbb{P}(\text{Displacement})$$

in this Schumpeterian economy. So we obtain that

$$\mathbb{P}(\text{Innovation}) = \mathbb{P}(\text{Displacement}) = \frac{g_Q}{\log \lambda} = \frac{\gamma}{\alpha \log \lambda} n;$$

that is, the rate of creative destruction and the innovation rate are positively proportional to population growth. This is our second result, following directly from the logic of creative destruction-based growth models.

B.3 Implications

Along a balanced growth path, in models with the ingredients outlined above, we must have the following implications.

- Per-capita output and income growth rates positively link to population growth rates.
- Innovation rates positively link to population growth rates.
- Creative destruction or displacement rates positively link to population growth rates.

Appendix Table 1: Assignment of States to Census Divisions

Census Region	State Names
New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Middle Atlantic	New Jersey, New York, Pennsylvania
East North Central	Illinois, Indiana, Michigan, Ohio, Wisconsin
West North Central	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
South Atlantic	Delaware, District Of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia
East South Central	Alabama, Kentucky, Mississippi, Tennessee
West South Central	Arkansas, Louisiana, Oklahoma, Texas
Mountain	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
Pacific	Alaska, California, Hawaii, Oregon, Washington

Appendix Table 2: Permutation Tests for Main Specification

	(1)	(2)	(3)	(4)
	Coefficient		Standard Error	RHS Rejection
	(Mean)	(St. Dev.)	(Mean)	Rate (%)
Panel A: First Stage				
Placebo 1	-0.0003	0.013	0.007	0.20
Placebo 2	0.0001	0.013	0.008	0.00
Placebo 3	-0.0117	0.031	0.020	2.70
Panel B: Reduced Form				
Placebo 1	0.0016	0.073	0.049	0.80
Placebo 2	0.0026	0.069	0.048	0.70
Placebo 3	-0.0023	0.069	0.044	4.40

Notes: This table reports the results of three different placebo tests on our standard specification, corresponding to column 2 of Table 3. For each of the placebo tests, we randomly reassign the instrument across observations: in the first version, we randomly reassign within the entire sample (Placebo 1); in the second version, we randomly reassign within the same period t (Placebo 2); and in the third version, we reassign within the same period t and census division $r(d)$ (Placebo 3). For each version, we perform 1000 placebo runs. We present summary statistics on the first stage (Panel A) and reduced form (Panel B) coefficients of interest across placebo runs. Columns 1 and 2 report the average and standard deviation for the coefficient of interest, column 3 reports the mean standard errors, and columns 4 reports the percentage of runs for which we reject that the coefficient of interest is different from 0 at the 5% level on the right-hand side. The standard errors are clustered by state in our standard specification and hence all placebo runs.

Appendix Table 3: County-Level Panel Regressions of Difference in Patenting on Population Growth

	(1)	(2)	(3)	(4)
Panel A: Second Stage	*5-Year Difference in Patenting per 100,000 People Post 1980*			
	(OLS)	(IV)	(IV)	(IV)
Δ Population$_d^t$	0.223***	0.113***	0.113***	0.087***
	(0.066)	(0.030)	(0.031)	(0.027)
N	18,846	18,846	18,840	18,846
Panel B: First Stage	Δ *Population*$_d^t$			
Immigration Shock ($\hat{I}_d^t$)		1.885***	1.877***	2.002***
		(0.178)	(0.183)	(0.276)
F-Stat		112	105	53
R^2		0.324	0.338	0.808
Geography FE	State	State	State	County
Time FE	Yes	Yes	Yes	Yes
State-Time FE	No	No	Yes	No

Notes: The first panel of this table reports the results of our second-stage specification, described in equation (1), where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) in county d in the five-year period ending in t and the endogenous variable is population growth (1,000s) in d and period t. The second panel reports the results for step 3 of instrument construction, or the coefficient estimates for the first-stage specification for population change (1,000s) for the instrument described in equation (7). Column 1 provides the results of the OLS estimation of equation (1), whereas columns 2-4 provide an IV estimate of the second stage (first panel) and first stage (second panel). The table includes the first-stage F-statistic on the excluded instrument for each of the IV specifications. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix Table 4: Panel Regressions of Inflows of Native Migrants on Non-European Immigration

	Inflows of Internal Migrants:	
	All Natives	*Non-Hispanic White Natives*
	(1)	(2)
Immigration_d^t	3.675***	2.100***
	(0.616)	(0.406)
N	9,415	9,415
First Stage F-Stat	3,484	3,484
Geography FE	State	State
Time FE	Yes	Yes

Notes: This table reports the results of our second-stage specification, described in equation (1), for the migration of natives (1,000s) into county d in period t (for 1980, 1990, and 2000) with non-European immigration (1,000s) to d in t as the endogenous variable. Note, migrants who moved into county d from a foreign country are excluded. Standard errors are clustered by state for all specifications and *,**, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix Table 5: Results from Placebo Analysis Based on Adão et al (2019)

	(1)	(2)	(3)	(4)
	Coefficient (Mean)	Coefficient (St. Dev.)	Standard Error (Median)	Rejection Rate (%)
Panel A: Realized Ancestry Shares				
First Stage	-0.003229	0.0776	0.0403	28.2
Reduced Form	-0.000471	0.0168	0.0112	18.8
Panel B: Predicted Ancestry Shares				
First Stage	-0.002000	0.0388	0.0240	4.5
Reduced Form	-0.002597	0.1088	0.0904	8.2

Notes: Following Adão et al. (2019), we randomly generate immigration shocks (for each $\{o, r, t\}$ country-region-time triplet), and construct placebo instruments by interacting these random shocks with actual baseline ancestry shares (as in a traditional shift-share instrument) and our predicted baseline ancestry shares (as in the ancestry-share version of our baseline instrument). We then run 1,000 placebo regressions of actual immigration on the randomly generated Card-style instrument (Panel A) and our randomly generated instrument (Panel B); we also run the comparable reduced-form regressions where the dependent variable is our primary measure of patenting, the five-year difference in patenting flows per 100,000 people. Column 1 reports the mean value of the coefficient over all placebo regressions, whereas column 2 reports the standard deviation. Column 3 then reports the median standard error for the coefficient of interest over all placebo regressions, and column 4 reports the fraction of placebo regressions for which we reject the null hypothesis of no effect at the 5% statistical significance threshold. As shown, the traditional shift-share instrument suffers from the over-rejection identified in Adão et al. (2019) with false rejection rates of 28.2% in the first stage and 18.8% in the reduced-form specification. The ancestry-share version of our baseline instrument has false rejection rates of 4.5% (first stage) and 8.2% (reduced form). The latter is similar to the false rejection rates reported in Adão et al. (2019) when using their proposed standard error correction (labelled "AKM").

Appendix Table 6: Panel Regression of 5-Year Difference in Patenting per 100,000 People on Immigration using Alternative Patent Counts

	Difference in Patenting per 100,000 People Post 1980			
	Assignee (Unweighted)	*Assignee (Cite Weight)*	*Inventors (Unweighted)*	*Inventors (Cite Weight)*
	(1)	(2)	(3)	(4)
Immigration$_d^t$	0.101***	0.162***	0.269***	0.487***
	(0.031)	(0.042)	(0.092)	(0.140)
N	18,846	18,846	18,846	18,846
First Stage F-Stat	911	911	911	911
Geography FE	State	State	State	State
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of our second-stage specification, described in equation (1), for the change in patenting per 100,000 people (population is based on baseline 1970 levels) with non-European immigration (1,000s) to d in t as the endogenous variable. Column 1 repeats our main specification where patent location is based on assignees and raw patent counts are used. Column 2 also uses the assignee for patent location but uses citation-weighted patent counts. Columns 3 and 4 then provide results when inventors are used for identifying patent location where patent counts are unweighted and citation-weighted, respectively. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix Table 7: Spillover Analysis – First Stage

	Column 1	Column 2		Column 3		Column 4			
	Immigration$_d^t$	Immigration$_d^t$	Immigration$_{State}^t$	Immigration$_d^t$	Neighbors' Immigration$_d^t$	Immigration$_d^t$	Immigration$_{100km}^t$	Immigration$_{250km}^t$	Immigration$_{500km}^t$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immigration Shock ($\hat{I}_d^t$)	2.130***	2.120***	-1.212	2.093***	0.001	2.094***	-0.379	-0.080	0.345
	(0.072)	(0.070)	(1.995)	(0.055)	(0.002)	(0.058)	(0.257)	(0.264)	(0.450)
State Immigration Shock ($\hat{I}_{State}^t$)		0.001***	2.915***						
		(0.000)	(0.217)						
Neighbors' Immigration Shock ($\hat{I}_N^t$)				4.938*	2.388***				
				(2.730)	(0.369)				
Immigration Shock 100km ($\hat{I}_{100km}^t$)						0.058	3.404***	-0.071	-1.264
						(0.040)	(0.993)	(0.322)	(0.764)
Immigration Shock 250km ($\hat{I}_{250km}^t$)						0.006	-0.047	2.623***	-0.617*
						(0.011)	(0.095)	(0.387)	(0.315)
Immigration Shock 500km ($\hat{I}_{500km}^t$)						-0.006	-0.201*	-0.339	2.030***
						(0.007)	(0.120)	(0.232)	(0.263)
N	18,846	18,846	18,846	18,846	18,846	18,846	18,846	18,846	18,846
First Stage F-Stat	876	1,792	470	2,175	162	6,065	383	150	66
Geogrpahy FE	Division	Division	Division	Division	Division	Division	Division	Division	Division
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of the first-stage regressions for the IV regressions shown in Table 10. Column 1 of this table provides the first stage regression for column 1 of Table 10. The first stages for column 2 of Table 10 are shown in columns 2 and 3 of this table while those for column 3 in Table 10 are shown in columns 4 and 5 of this table. Finally, columns 6-9 display the first-stage regressions for column 4 of Table 10. For each specification, we report the first-stage F-statistic for the IV estimation in Table 10, utilizing the F-statistic described in Angrist and Pischke (2009, p. 217-218) in the case of multiple endogenous variables. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.