Immigration, Innovation, and Growth*

Konrad B. Burchardi† **Thomas Chaney**‡ **Tarek A. Hassan**§
Lisa Tarquinio¶ **Stephen J. Terry**‖

February 2020

Abstract

We show a causal impact of immigration on innovation and dynamism in US counties. In order 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 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 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.

JEL Classification: J61, O31, O40.
Keywords: migrations, innovation, patents, endogenous growth, dynamism.

*We are grateful to Daron Acemoglu, David Autor, Murat Celik, Klaus Desmet, Ray Fisman, Kevin Lang, 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.

†Institute for International Economic Studies, 106 91 Stockholm, Sweden; E-mail: konrad.burchardi@iies.su.se.

‡Sciences Po, 28 rue des Saints Pères, 75007 Paris, France; E-mail: thomas.chaney@gmail.com.

§Boston University, NBER and CEPR, 270 Bay State Road, Boston MA 02215, USA; E-mail: thassan@bu.edu.

¶Boston University, 270 Bay State Road, Boston MA 02215, USA; E-mail: ltarq@bu.edu.

‖Boston University, 270 Bay State Road, Boston MA 02215, USA; E-mail: stephent@bu.edu.

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 United States over the last three decades. We find a positive causal impact of immigration on both innovation and economic dynamism at the county level.

Canonical models of endogenous growth and firm dynamics suggest that immigration should increase both innovation and economic dynamism. As population grows with the inflow of migrants, more people can work towards more and possibly harder innovations (Jones, 1995; Bloom et al., 2017). As innovation is embedded in creative destruction at the microeconomic level, those new innovations facilitated by the inflow of migrants should also lead to more creative destruction and a higher level of overall economic dynamism (Aghion and Howitt, 1992; Grossman and Helpman, 1991; Hopenhayn and Rogerson, 1993; Klette and Kortum, 2004). If local returns to scale are large enough, the arrival of migrants should therefore cause innovation, economic dynamism, and economic growth within the receiving local communities.[1]

Contrasting with 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 elusive. 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, it is plausible that migrants arriving in the US will tend to 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.

Our main contribution is to propose a formal identification strategy which 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

[1] Peters (2017) shows evidence of such a link using a historical experiment in post-war Germany.

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 United States, 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 measures of 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 United States 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 that 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*) a pull factor, causing immigration into a given US county from all origins. To further ensure that our predicted historical migration is not contaminated by endogenous unobserved factors, we carefully leave out large population groups when predicting ancestry. In particular, as our focus is on immigration to the US after 1970, primarily originating from non-European countries, we use the historical location choices of European migrants to predict where non-European migrants settled. 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 origin country. Iterating this procedure over 100 years, we are able to isolate quasi-random variation in the distribution of ancestry across US counties in 1970.

In our second step, we use this predicted pre-existing distribution of ancestry to predict where new migrants arriving in the US after 1970 will settle. Implicitly, we assume 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 will 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 1970. This predicted immigration is plausibly orthogonal to any origin-destination-specific factor which may make a destination US county more innovative and dynamic after 1970.

Finally, to further guard against any lingering concerns about identification, we estimate the impact of plausibly exogenous variations in immigration on *changes* in innovation, dynamism, and growth, not on levels. In many specifications, we are even able to control for 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: the arrival of 10,000 additional immigrants increases the flow of patents over a 5-year period by 1 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 one aspect 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 5%, all expressed as changes relative to their mean. The significant rise in local wages suggests that immigration does not only affect 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 the US state surrounding county d, innovation in d increases significantly. However, this spillover effect of immigration in 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 that 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. 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, while 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, relatively uneducated migrants (in the bottom third of the distribution of years of schooling among incoming migrants) have almost no effect on local innovation, while the increase in innovation induced by highly educated migrants (in the top third) is an order of magnitude larger than for the average migrant: it only takes 1,000 highly educated migrants to increase patenting by 1 patent per 100,000, the same marginal effect as that of 10,000 migrants with the average education level in our sample.

Related Literature. Our paper relates to several strands of the theoretical and applied literature. First, we contribute to the recent but growing literature on the link between immigration and innovation. Surveys by Hanson (2009, 2010) and Lewis (2013) emphasize the potential importance of a link between immigration and innovation. Early contributions providing empirical evidence for a link between migration and innovation are Kerr and Lincoln (2010) and Hunt and Gauthier-Loiselle (2010). Both papers present variants of a standard shift-share instrument. The former uses regional variation in the use of the H-1B visa program interacted with the national growth of the H-1B cap over time; the latter uses the baseline distribution of immigrants from a number of source countries and time-variation in the aggregate number of skilled immigrants arriving to the US over subsequent years. Peters (2017) investigates the link between historical refugees in Germany and industrialization. Using data on the age of mass migration and the expansion of railways at the time, Sequeira et al. (2020) show a long-run effect of immigration on local economic development. Akcigit et al. (2017) show that many inventors are immigrants. Bernstein et al. (2018) also show that many inventions originate in the work of immigrants. Kerr and Kerr (2016) study the case of immigrant entrepreneurship. Moving beyond impacts on narrowly measured patenting or innovation, Lewis (2011) and Lafortune et al. (2019) emphasize that the economic impact of migration may be mediated through its impact on production processes, skill mixes, and capital responses. Tabellini (2018) shows that while historical immigration to the US has fostered development, it has also sparked a political backlash, a contemporaneous environment further detailed in Kerr (2018).

Second, our empirical study is motivated by theories of endogenous growth. While the seminal contribution of Romer (1990) predicts that a larger population, in levels, ought to increase the rate of growth of the economy, subsequent refinements (Jones, 1995, 1999) in so-called semi endogenous growth theory predict a positive link between the growth rate of population and economic growth. This purposefully brief description does not do full justice to the rich debate on scale effects in growth, with other important theoretical contributions including Peretto (1998) and Young (1998). Empirical work by Laincz and Peretto (2006) further disentangles predictions of the varying classes of models. In these models, a larger population, or a higher population growth rate, allow an economy to grow because increasing returns to scale in the technology for innovating overcome decreasing returns to scale in production (Solow, 1956). Another key insight from this literature is that a larger population of (potential) innovators may be better able to achieve increasingly hard innovations. Bloom et al. (2017) show evidence that new ideas are in fact becoming harder to produce over time, so that an increasing amount of resources are needed to continue innovating. We contribute to this literature first by showing a positive impact of immigration –one channel through which population grows - on innovation and growth. We also show evidence that the scale effects necessary for endogenous growth in Jones (1995) operate and are statistically identifiable at the local level. Finally, we show that a key input in the technology to produce innovation is human capital, whereby more educated migrants contribute more to innovation than less educated migrants.

Endogenous growth theory also ties the innovation and growth process to creative destruction, as in the seminal contributions of Aghion and Howitt (1992), Grossman and Helpman (1991), and Klette and Kortum (2004). We confirm the predictions from those theories: an increase in population, induced by immigration, which feeds the innovation process, also increases measures of creative destruction. US counties receiving more migrants experience both higher rates of job creation and destruction and also see more positively skewed growth, with more "superstar" or right-tail growth of industries in their area.

Third, we contribute to the empirical literature which has investigated the decline in dynamism, in particular in the US. Using aggregate trend data and a structural approach, Karahan et al. (2016) show that the declining dynamism in the US may be due to lower population growth. Khanna and Lee (2018) show a positive association between high-skilled migration and detailed firm-level measures of dynamism and turnover. Using US census microdata, Decker et al. (2014) document the declining dynamism in the US. Gordon (2012) contemplates the idea that growth may have run out in the US. Alon et al. (2018) suggest that the decline in US dynamism may be due to older firms. Using an endogenous growth theory, Akcigit and Ates (2019) argue that the decline may be due to declining knowledge diffusion across firms. Using data on aggregate trends, and guided by the theory, Hopenhayn et al. (2018) propose that the decline in dynamism in the US is due to older firms and a lower population growth rate. Hathaway and Litan (2014) also document declining measures of dynamism in the US. We contribute to this literature by showing causal evidence that population growth, and in particular population growth driven by immigration, may counteract the decline in dynamism in the US economy.

Important recent advances identify conditions for shift-share instruments to be valid, notably the work by Borusyak, Hull, and Jaravel (2018) and Goldsmith-Pinkham, Sorkin, and Swift (2018). The latter highlight that exogeneity of the shares is a sufficient condition for consistency of shift-share instruments. Our approach precisely singles out plausibly exogenous variation in the shares – the spatial distribution of ancestry – and uses that variation only in a shift-share like approach. This distinguishes our approach from earlier papers which use raw baseline shares.

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

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. 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 5-year periods,[2] 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 United States between $t-1$

[2]Due to data limitations in the reporting of year of immigration in the 1970-1990 censuses, the length of periods vary slightly over this time frame.

and t. The exception to this rule is the 1880 census (the first in our sample), which also 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.[3] 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 that this measure captures self-reported (recalled) ancestry.[4]

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 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 (for example, "Spanish" or "Denmark"). When no direct mapping exists (for example, "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.[5] 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 the earlier periods, when there was more personalized US innovation, inventor location is the natural measure of geography (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. Nevertheless, in addition to this baseline choice we also explore alternative means of

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

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

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

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.[6] Appendix A.2 gives additional details.

Dynamism. A growing empirical literature emphasizes that measures of dynamism and creative destruction in the United States 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 computed from the underlying Longitudinal Business Data microdata on the employment levels of the universe of US business establishments. The BDS data include job creation and job destruction rates (gross flows representing the ratio of the number of jobs created or destroyed as a fraction of total employment) at the yearly level and spanning 1977–2015. The sum of these measures is known as the job reallocation rate, and the difference between them is the net employment growth rate. We apportion the native MSA geography to 1990 US counties by population. 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, i.e., a decline in the relative importance of "superstar" growth performance in driving US employment dynamism. 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.

[6] We manually check the patenting 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.

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. It 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 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-years 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 5 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.

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 5-year window level. The table reveals sensible patterns. Counties on average received around 1.4 thousand non-European immigrants in each 5-year period between 1975 and 2010, a meaningful contribution to overall population growth of around four thousand. 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, 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 or 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 specification also includes 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, it remains possible that migrants are disproportionately drawn to counties within the US which 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 United States from that origin country, and then summing over origin countries, as for example in Hunt and Gauthier-Loiselle (2010). However, it is likely that omitted factors which 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; and more Indian engineers may systematically migrate to Silicon Valley (and other information technology hubs) whenever there is a boom in the information technology industry. If this were the case, the canonical shift-share approach would falsely identify a causal effect of immigration on innovation, when in reality innovations in software are the reason why destinations with high pre-existing Indian ancestry shares receive more immigration. Thus, if ancestry shares are themselves endogenous – i.e. are potentially correlated with unobserved factors affecting innovation – this poses 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 in the US: (i) time-series variation in the relative attractiveness of different destinations within the United States to the average migrant arriving at the time (e.g. end of nineteenth century Midwest versus early twentieth century West) and (ii) the staggered arrival of migrants from different origins (e.g. end of nineteenth century China versus early twentieth century Japan). We argue that 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, like the affinity of Indian engineers for software development mentioned above. We then use only the exogenous component of the pre-existing distribution of ancestries to predict migrations into each US county post 1970. 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,

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

In both equations, the δ terms are fixed effects, and $X_{o,d}'\beta$ 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 'push-pull' and 'shift-share'. The 'push-pull' force is captured by the term $I_{o}^{t}\left(I_{d}^{t}/I^{t}\right)$: in time periods where 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 'pull' I_{d}^{t}/I^{t} factor), we expect many migrants from o to settle in d. This 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 'shift-share' force is captured by the term $I_{o}^{t}\left(A_{o,d}^{t-1}/A_{o}^{t-1}\right)$: migrants arriving from o (the 'shift' factor I_{o}^{t}) have a tendency to locate in destinations d with a pre-existing community from their home country (the 'share' factor

$A_{o,d}^{t-1}/A_o^{t-1}$). This corresponds to a social motive for migration: other things equal, migrants tend to prefer living near others of their own ethnicity.

We use the recursive set of equations (2)-(3) to construct an instrument for immigration.

The 'push-pull' force, cumulated over several periods, allows us to isolate plausibly exogenous variations in pre-existing ancestry inherited from historical shocks that operated prior to 1975. To fix ideas around these 'push-pull' forces more concretely, directly examining the underlying variation proves useful. Figure 1 plots the share of non-European immigration into the US from 14 of the 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 rising over time, in this case as the various US immigration exclusion acts were repealed.

As the different nations in Figure 1 sent immigrants at different points in time to the US, the location of relatively attractive destination counties – our source of variation for the pull factor also changed substantially over time. Figure 2 plots color-coded maps of migration into the US over Census waves from 1880 to 2010, with darker shades representing a higher intensity of migration to a given county. Early on in the sample during the late 19th century northeastern locations were particularly attractive destinations. But by the early 1900's the average immigrant's favored destinations shifted to the midwest and western regions, before shifting yet again to the coastal and southeastern regions later on.

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 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, say, the fact that certain southeastern and midwestern regions happened to be popular destinations ('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 variation in pre-existing ancestry composition, we can then confidently use the 'shift-share' force to predict contemporaneous immigration shocks in the period after 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 'push-pull' force in (2), and by cumulating successive migration waves using (3). We complement this model by adding 'leave-outs', to ensure that our instruments do not contain any, potentially confounding, origin-destination-specific factors . Formally, we estimate

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

The 'leave-outs' ensure that 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 (Census division, a grouping of several adjacent US states) where d is located over the 5-year period ending in τ; 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.[7] Our results are robust to using various alternative 'leave-out' strategies. $\delta^t_{o,r(d)}$ and $\delta^t_{c(o),d}$ are a series of origin country-destination region and continent of origin-destination county interacted fixed effects, while $X_{o,d}$ contains a series of time invariant controls for $\{o, d\}$ characteristics (including distance and latitude difference). We estimate (4) separately for each $t = 1980, 1985, 1990, 1995, 2000, 2005, 2010$ on all non-European countries.

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

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

[7]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 that this leave-out imposes a stricter requirement than simply removing migrants from o from I^τ_d/I^τ.

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 $X_{o,d}$. Again, our baseline specification uses region and continental leave-outs, $I^{\tau}_{o,leaveout} = I^{\tau}_{o,-r(d)}$ and $I^{\tau}_{leaveout,d}/I^{\tau}_{leaveout} = I^{\tau}_{Europe,d}/I^{\tau}_{Europe}$.

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 1970, we use the 'shift-share' 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

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

where 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 $\tilde{I}^{t}_{o,leave-out} = I^{t}_{o,-r(d)}\left({}^{I^{t}_{Europe,r(d)}}/{}_{I^{t}_{Europe,-r(d)}}\right)$. As we leave-out all migrants from o who settle in d's region from $I_{o,-r(d)}$, we include a scaling factor at the regional level, ${}^{I^{t}_{Europe,r(d)}}/{}_{I^{t}_{Europe,-r(d)}}$, to correct for differences in region sizes.

Step 3: Predicting aggregate migration. We are finally able to generate our main instrument for the total number of migrants settling in county d in period t, $\text{Immigration}^{t}_{d}$ in equation (1),

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

Identifying assumption. With our baseline regional and continental leave-outs, a sufficient condition for the validity of this instrument can be written 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 confounding factors that correlate with increases in a given county's innovation or dynamism post-1975 and historically made a given destination more attractive for migration from a given non-European origin country do not also correlate with past instances of the interaction of the settlement of European migrants with the total number of migrants arriving from that non-European origin who settle in other US regions. 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).[8]

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

We believe this assumption is plausible: consider again a shock to productivity of software development in Silicon Valley that attracts Indian software engineers. This confounding shock, and any other origin-destination specific factor that drives migration and might affect the destination's capacity for future innovation, in general affects neither $\hat{A}_{o,d}^{t-1}$ nor $\tilde{I}_{o,-r(d)}^{t}$; the former depends only on how the historical destination choices of Europeans coincided with the number of Indians arriving in the US who chose destinations other than the West Coast, while the latter depends only on the number of Indians arriving in t who again do not settle on the West Coast. In order to violate (8), the confounding shock would instead have to systematically affect both the destination choices of Indians and large numbers of Europeans (enough to sway shares), while also attracting large numbers of Indians to US counties outside of the West Coast and generating economic dynamism in these locations post-1975. We address this remaining (if unlikely-sounding) concern below by varying the way in which we construct the leave-out categories in our estimation.

3.2 The Construction and Performance of the Instrument

We now review the estimation results of each of the steps towards 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 predict ancestry levels by using historical push-pull factors in (5). Figure 3 reports the coefficients in this regression predicting 2010 ancestries, and reports the coefficients on the interaction term by time period (assuming for presentational purposes only that $b_{r(d)}^{\tau} = b^{\tau}\ \forall r(d)$). The results indicate that we identify variation in current ancestry levels based on push and pull factors from across the full range of time periods in our sample, with statistically precise contributions from periods as far back as the pre-1900's 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 the resulting predicted ancestry levels against realized ancestry in 2010. Realized and predicted ancestry are tightly aligned along the 45-degree line. (The corresponding regression of $A_{o,d}^{2010}$ on $\hat{A}_{o,d}^{2010}$ as defined in (5) yields an R^2 of 74.9%.) We conclude that Step 1 successfully predicts plausibly exogenous variation in 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 5-year period post-1970 to predict plausibly exogenous variation in immigration $I^t_{o,d}$ at the $\{o, d\}$ level in (6). We allow the coefficient γ_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 origin-county level in all seven time periods post-1970, 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, 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) 'shift-share' terms remain virtually unchanged. Remarkably, they even remain unchanged in column 5, where we control directly for contemporaneous economic forces shaping migration by including 'push-pull' interactions for each period post-1970;[9] 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 that our instruments for origin-destination-specific migration are orthogonal to a wide range of observables, and that Step 2 successfully predicts plausibly exogenous variation in immigration at the $\{o, d\}$ level.

In Step 3, we sum across origin countries to compute an instrument $\hat{I}^t_d$ for total non-European immigration to county d at time t (7). Figure 5 presents a series of maps displaying this "immigration shock" for each 5-year period from 1975 to 2010. To make those maps easier to read, we remove county and state-time fixed effects. As a sense-check, Appendix Table 1 shows that $\hat{I}^t_d$ indeed significantly predicts both non-European immigration and population growth conditional on these controls. For example, in column 5, a regression of a county's 5-year change in population on $\hat{I}^t_d$ yields a coefficient of 1.921 (s.e.=0.323), suggesting that a one standard deviation increase in our immigration shock (4.99) is associated with a 0.49 standard deviation increase in the county's population and, equivalently, a 0.65 standard deviation increase in the county's number of newly arrived immigrants.

[9]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.

4 The Impact of Immigration on Innovation and Growth

Classical endogenous growth theories link population growth to innovation, dynamism, and income growth at the local level, a result which we lay out in a straightforward theoretical derivation in Appendix B. In this section, we exploit our quasi-random variation in immigration above to test these predictions 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 arriving in the county during the 5-year period. The dependent variable is the change in patents *per capita* over the same 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 allocated by the social factor in (2)) on county-level innovation. It implies that the arrival of 10,000 additional immigrants in a given county on average increases the flow of patents filed over a 5-year period by 1 patent per 100,000 people. Comparing these magnitudes to the summary statistics above, an increase in immigration flows of one standard deviation - 12 thousand immigrants - causes around 1.2 more patents per 100,000 people, an increase of 27% relative to mean (4.45 patents per 100,000 people).[10] The F-statistic on the excluded instrument 911, and thus far above critical values.

Column 1 shows the OLS estimate of (1) for comparison. As expected, it is larger than our preferred estimate (by about 2 standard errors), consistent with the view that, other things equal, immigrants select into innovative counties in equilibrium – resulting in an upward bias in OLS estimates of the effect of immigration on innovation.

Columns 3 and 4 show our results are robust and the estimated impact of immigration on innovation vary little if we include interacted time and state fixed effects (0.100, s.e.=0.032 in column 3), or even county fixed effects (0.108, s.e.=0.033 in column 4).[11] Panel B of Table 3

[10] In order 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 predicted by endogenous growth models, to section 5.3.

[11] 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

shows similar results and similar magnitudes when we consider the impact of population growth on innovation, instrumenting population growth with immigration shocks as in panel B of Table 1.[12]

4.2 Robustness

We show below that our results are robust to a large array of alternative specifications.

Alternative Instruments. Table 4 shows how our instrumentation and identifying assumptions affect our estimates of (1). Column 1 freezes predicted ancestry at its 1975 level, instead of updating predicted ancestry each period. 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 towards d. 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' towards d when predicting migrations from o to d, we use instead all migrants to d originating from countries outside d's continent. Comforting for our identifying assumption, all of these variations yield estimates that are almost identical to the one in our standard specification (0.101).

In particular, recall that a confounding factor violating our assumption (8) would have to systematically and repeatedly attract immigrants of a given ethnic group (Indians) to a given US destination (Silicon Valley), while at the same time also attracting and large numbers of Europeans (enough to sway shares) to the same county, and also attracting large numbers of migrants from the same origin to US locations in other census regions (regions other than the West Coast), while also generating economic dynamism in these locations post-1975. If such (complicated) confounding factors 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

prior shocks, a common problem with traditional implementations of the 'shift-share' approach emphasized in recent work by Jaeger et al. (2018).

[12]Consistent with this positive effect of immigration on innovation we also find that a positive immigration shock has an agglomerative effect. 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 3 for details.

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 'shift' ($\tilde{I}_{o,leave-out}^t$), 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 re-traces each of these steps to make clear how each modification affects our estimates, and how they help to address econometric shortcomings of canonical applications of the 'shift-share' approach highlighted in the recent literature (Adão et al., 2019).

Column 1 replicates our standard specification for comparison. Column 2 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, as predicted ancestry from (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 that it allows us to test whether our instrumentation approach successfully addresses an over-rejection problem in standard 'shift-share' applications, which take pre-existing ancestry shares as given. 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).[13] 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

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

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%.[14]

For comparison, column 3 repeats the same estimation as that of column 2 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 remain similar). Finally, in column 4, 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.

We conclude that our instrumentation strategy is sufficiently powerful to isolate quasi-random variation in ancestry levels, or shares, and that it effectively removes spurious correlation with the error term, bolstering our confidence in a causal interpretation of the results. Finally, it is worth noting that, despite their various challenges, 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 growth: population density in 1970, the number of patents generated in 1975 per 1,000, 1970 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 are themselves outcomes of migration and should thus more appropriately be thought of as various channels through which historical migration and ancestry may affect innovation. Nevertheless, it is comforting for our identifying assumption that controlling for these initial conditions has only modest effects on our result. 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

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

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 sequentially exclude migrations from the five largest sending countries post 1975 (Mexico, China, India, Philippines, and Vietnam) from the sum in (7), thus treating migration from these countries as endogenous. While dropping Mexican immigration from our instrument lowers the F-statistic in the first stage by about half, 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, while 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 5-year period, as in Table 3 panel A column 4. Columns 3 and 4 consider the impact of immigration on patenting over a 10- or 15-year period. We find that 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 2 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 immigration's impact 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 that 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 that immigration may help to dampen such declines. Turning to magnitudes, a one standard deviation increase in immigration in a county - around 12 thousand more people - leads to 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).[15]

Exploring an alternative measure of dynamism, higher immigration causes an increase in the skewness of employment growth (column 3). Intuitively, when more immigrants arrive, the importance of 'superstar' employment growth experiences across sectors in an area increases. 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 thousand more people to a county on average increases wages per capita by around 5% relative to the mean observed growth. Because the QCEW wage data does 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

[15] It is worth noting that although most endogenous growth theories link higher dynamism to innovation, higher income growth, and higher welfare, the impact of dynamism on the subjective wellbeing of individuals exposed to such creative destruction is more ambiguous (Aghion et al., 2016).

county 5 years prior to the census. We find a positive, and statistically significant, effect of immigration on the average wage of both groups.

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.

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, since most theoretical models predict that 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, finding 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 two 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 follow a symmetric procedure, adding the immigration shocks up 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. 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 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.[16] 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 thousand 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 immigration flow constant, increases patenting per capita by around 31% relative to mean. In other words, both local immigration and immigration to the surrounding state positively impact local innovation. Although migrants to other counties matter less individually for a county's innovation, the larger scale of those flows means that 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 that immigration to nearby counties, where we discount distant counties inversely with distance.[17] We show that 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 that immigration to close-by 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 mean; a one standard deviation increase in

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

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

immigration between 100km and 250km increases innovation by 42%; but immigration beyond 250km no longer has a statistically detectable effect on innovation.

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 20% increase relative to mean. However, the effect is statistically indistinguishable from zero beyond 100km.

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 greater, 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 - total years of education or total years of college, in two alternative versions - adding a second endogenous variable to our baseline specification.

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. Our identification strategy allows us to construct a separate instrument for each origin country o - destination county d pair and each time period t, $\hat{I}^t_{o,d} = \hat{A}^{t-1}_{o,d} \times \tilde{I}^t_{o,-r(d)}$. 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 variable takes the form

$$Average\ Years\ Education^t_d \times Immigration^t_d = \delta_s + \delta_t + \sum_{o=1}^{20} \kappa_o \hat{I}^t_{o,d} + \nu^t_d.$$

Since migrants from different countries have different schooling levels, and they emigrate to different counties, we are able to isolate variations in the level of education of migrants across counties.

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.[18] Column 2 adds the interaction of immigration with (demeaned) average years of education for immigrants to the same county. The estimates indicate that more highly educated immigrants cause a bigger increase in innovation. To inspect the magnitude of the heterogeneity at work here, consider two counties both receiving ten thousand more migrants. A county receiving migrants of average education (around 11 years per person) 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 (about 3.5) extra years of education per person would see $10 \times (0.200 + 3.5 \times 0.221) = 9$ 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 that columns 2 through 4 rely on a linear interaction of immigration and education, imposing functional form restrictions on the link between education and innovation responses. Column 4 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 immigrants see an order of magnitude higher impact on patenting relative to counties receiving medium-education migrants, while 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 increases average wages

[18] 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.

in column 1. Column 2 reveals a higher impact in counties which receive immigrants with a higher average education level. To evaluate magnitudes, once again consider two counties both receiving 10 thousand more migrants. A county receiving migrants of average education (around 11 years per person) would see wage increases by around $10 \times 0.290 \times \$100 = \$290$ more per person over 5 years. A county receiving the same number of immigrants with one standard deviation (about 3.7) extra years of education per person would see more wage growth by around $10 \times (0.290 + 3.7 \times 0.231) \times \$100 = \$1{,}145$ per person over 5 years. Column 4 reveals similar - 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 growth to total population change, rather than immigration per se. Furthermore, different models imply different functional forms. For example, innovation growth rates respond to absolute population in "strong scale effects" models (Romer, 1990), suggesting a positive semi-elasticity of innovation to people. By contrast, "weak scale effects" models (Jones, 1995, 1999) suggest a positive growth response to population growth rates or elasticity specifications. Although we build a suggestive motivating model in Appendix B, our data does not have the sometimes centuries-long time scale required to cleanly differentiate alternative models, and our results relate to local or relative innovation rather than aggregate dynamics. However, with those limitations in mind, Table 12 explores a range of variations on our baseline specification loosely inspired by the endogenous growth literature.

First, column 1 duplicates our baseline results. We then note that if population growth rates determine innovation growth, as in models with weak scale effects, counties with larger absolute immigration flows - and lower corresponding population growth rates induced by immigration - should see smaller 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 that 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 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, with different classes of growth models mapping to each concept as discussed above. Column 5 reveals a positive semi-elasticity of innovation to immigration. Although our instrument was designed to predict immigration rather than overall population change, and has less power for predicting overall population change as measured by the first-stage F-statistic. Column 6 reveals a positive semi-elasticity to total population changes induced by predicted immigration. 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 one-percent 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 1% higher population change.

6 Conclusion

We find that plausibly exogenous variation in migration at the county level induced by quasi-random variation in historical "push" and "pull" factors for immigration strongly predicts overall realized immigration. We then estimate that over the past four decades immigration at the local level has spurred increased innovation or patenting per capita. A strong tradition in endogenous growth theory predicts exactly these links, with more immigrants bringing new ideas, purchasing power, and scale which contribute to innovation. The impact of immigration on innovation spills over positively to nearby counties and unsurprisingly increases with the schooling of immigrants.

However, we also find that immigration increases dynamism or creative destruction at the local level. Classic theoretical models predict that such creative destruction, which has been declining on average in the US overall in recent decades, should move with overall innovation, consistent with our findings on patents above. We conclude that increased immigration in recent decades may have prevented even worse declines in dynamism at the local level, and might contribute to further mitigating the downward trend in dynamism if allowed in the future.

Finally, we show that immigration to an area also has increased income growth or wages per capita at the local level.

Given the relevance of innovation, dynamism, and growth for the long-run prospects of the US economy as well as for macro models of growth, we view our results as relevant to the ongoing debate about the appropriate choice of immigration policy.

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. 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.

Altonji, J. G. and D. Card (1991): "The effects of immigration on the labor market outcomes of less-skilled natives," in *Immigration, trade, and the labor market*, University of Chicago Press, 201–234.

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.

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

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.

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

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.

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.

Hopenhayn, H. and R. Rogerson (1993): "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of political Economy*, 101, 915–938.

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, S. P. and W. R. Kerr (2016): "Immigrant entrepreneurship," National Bureau of Economic Research working paper.

Kerr, W. R. (2018): *The Gift of Global Talent: How Migration Shapes Business, Economy, & Society*, Stanford University Press.

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.

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

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

Peters, M. (2017): “Refugees and Local Agglomeration – Evidence from Germany’s 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.

Solow, R. M. (1956): “A contribution to the theory of economic growth,” *The Quarterly Journal of Economics*, 70, 65–94.

Tabellini, M. (2018): “Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration,” 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				
Non-European immigration (1000s)	21,987	1.42	12.21	0.22
Difference in population (1000s)	21,986	4.03	19.56	2.56
Instrument for Non-European immigration (1000s)	21,987	0.00	4.99	0.24
Patents				
Patenting per 100,000 people	21,987	32.60	94.73	20.66
Difference in patenting per 100,000 people	18,846	4.45	47.83	6.45
Difference in patenting per 100,000 people (Inventors)	18,846	18.85	90.45	41.13
Difference in patenting per 100,000 people (Citation Weighted)	18,846	5.23	63.91	6.04
Dynamism and Wages				
Difference in job creation rate	6,600	-32.47	209.90	50.00
Difference in job destruction rate	6,600	-17.37	199.58	38.46
Difference in skewness of employment growth	12,564	-6.82	48.91	51.87
Difference in average annual wage ($100)	21,976	19.00	56.23	25.80
Difference in average annual wage of natives ($100)	12,546	10.75	25.80	32.20
Difference in average annual wage of native non-movers ($100)	6,274	16.85	27.19	33.08
Immigration and Education				
Non-Euro. immigration, age 25+	21,987	0.80	6.91	0.11
Non-Euro. immigration, age 25+: Average years college	21,987	1.50	1.41	1.82
Non-Euro. immigration, age 25+: : Average years education	21,987	10.88	3.65	4.59
Immigration for Spillovers				
State (minus own county) non-Euro. immigration (1000s)	21,987	808.00	1,438.90	557.47
Instrument for state (minus own county) non-Euro. immigration (1000s)	21,987	31.92	366.25	60.89
Inverse-distance weighted non-Euro. immigration (1000s)	21,987	1.15	0.78	0.65
Instrument for inverse-distance weighted non-Euro. immigration (1000s)	21,987	-0.01	0.15	0.03
Non-Euro. immigration to counties within 100km of d (1000s)	21,987	18.58	64.65	9.21
Instrument for non-Euro. immigration to counties within 100km of d (1000s)	21,987	-0.24	11.69	2.93
Non-Euro. immigration to counties within 250km of d (1000s)	21,987	74.96	133.50	67.60
Instrument for non-Euro. immigration to counties within 250km of d (1000s)	21,987	-2.46	26.90	10.81
Non-Euro. immigration to counties within 500km of d (1000s)	21,987	123.10	149.52	143.69
Instrument for non-Euro. immigration to counties within 500km of d (1000s)	21,987	0.00	36.88	10.01

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 instruments. Variables for immigration, population growth and education are all for 5-year periods as are the differenced outcomes except in the case of differences in average annual wage for natives and non-native movers, which are over 10-year periods.

c

Table 2: Regressions of Immigration on Shift-Share 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
Controls:					
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-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: Panel Regressions of 5-year Difference in Patenting per 100,000 People on Immigration (and Population Growth) at the County Level

	5-year Difference in Patenting per 100,000 People Post-1980			
	(1)	(2)	(3)	(4)
Panel A: Non-European Immigraton				
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
First-Stage F Stat.		911	807	85
Panel B: Population Growth				
$\Delta\ \text{Population}_d^t$	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
First-Stage F Stat.		112	105	53
Controls:				
Specification	OLS	IV	IV	IV
Geography FE	state	state	state	county
Time FE	yes	yes	yes	yes
State-Time FE	no	no	yes	no

Notes: 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 5-year period ending in t and the endogenous variable is non-European immigration (1,000s) (Panel A) or population growth (1,000s) (Panel B) in d and period t. Column 1 provides the results of the OLS estimation of Equation (1) while columns 2 through 4 provide an IV estimate of the second stage. 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

	5-year Difference in Patenting per 100,000		
Specification:	*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
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. 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 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:	*Baseline Instrument*	*Predicted Ancestry Shares*	*Realized Ancestry Shares*	*Realized Ancestry No Leave-Out*
	(1)	(2)	(3)	(4)
Immigration$_d^t$	0.101***	0.195**	0.106***	0.132**
	(0.031)	(0.090)	(0.035)	(0.055)
N	18,846	18,846	18,846	18,846
First Stage F-Stat	911	656	265	361
Adã o et al (2019) First-Stage False Rejection Rate:		4.5	28.2	28.2
Instrument Functional Form:				
Ancestry Measure	Levels	Shares	Shares	Shares
Instrumented Ancestry	yes	yes	no	no
Shift Leave-Out	yes	yes	yes	no
Controls:				
Geography FE	state	state	state	state
Time FE	yes	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 our baseline instrument which utilizes prediced ancestry in levels while column 2 utilizes predicted ancestry shares. Column 3 utilizes the same instrument as column 2, with predicted ancestry shares, but with no leave-out in the shift factor. Finally, column 4 takes the instrument in column 3 but replaces prediced ancestry shares with actual ancestry but with actual ancestry shares as in the traditional Card-style shift-share instrument. We report the first-stage F-statistic on the excluded instrument for each specification. For columns 2 through 4, 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). 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
Controls:						
Geography FE	state	state	state	state	state	county
Time FE	yes	yes	yes	yes	yes	yes

Notes: 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) and the endogenous variable is non-European immigration (1,000s) to d in t. Column 1 repeats our main specification while columns 2 through 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				
Excluding:	*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
Controls:					
Geography FE	state	state	state	state	state
Time FE	yes	yes	yes	yes	yes

Notes: This table reports the results of our second stage 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 that for 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
Controls:				
Geogrpahy FE	county	county	county	county
Time FE	yes	yes	yes	yes

Notes: This table reports the results of our second stage 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 while column 2 repeats the standard specification. Columns 3 and 4 then utilize the two-period and 3-period chanage 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: Panel Regressions of All Outcomes on Immigration

	5-Year Difference in:				10-Year Difference in:	
	Job Creation Rate	*Job Destruction Rate*	*Job Growth Rate Skewness*	*Average Annual Wage*	*Avg. Annual Wage: Native*	*Avg. Annual Wage: Native Non-Mover*
	(1)	(2)	(3)	(4)	(5)	(6)
Immigration$_d^t$	0.176***	0.152***	0.019***	0.083***	0.049***	0.056***
	(0.033)	(0.035)	(0.004)	(0.019)	(0.016)	(0.020)
N	6,600	6,600	12,564	21,976	9,411	6,274
First Stage F-Stat	951	951	151	1,202	750	1,178
Controls:						
Geography FE	state	state	state	state	state	state
Time FE	yes	yes	yes	yes	yes	yes

Notes: This table reports the results of our second stage 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 while the dependent variable for the specification shown in column 4 is the change in the average annual real wage ($100) over the 5-year period ending in t. Columns 5 and 6 reports results of a regression of the change in the average annual real wage ($100) for natives and non-native 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)
	5-Year Difference in Patenting per 100,000 People Post-1980			
Immigration^t_d	0.130***	0.107***	0.072**	0.080**
	(0.039)	(0.035)	(0.032)	(0.037)
$\text{Immigration}^t_{State}$		0.001***		
		(0.000)		
Neighbors' Immigration^t_d (Inverse Distance Weight)			6.600***	
			(1.593)	
$\text{Immigration}^t_{100km}$				0.056***
				(0.018)
$\text{Immigration}^t_{250km}$				0.014***
				(0.005)
$\text{Immigration}^t_{500km}$				0.006
				(0.005)
N	18,846	18,846	18,846	18,846
First Stage F-Stat d	876	1,792	2,175	6,065
First Stage F-Stat Spillover		470	162	383
First Stage F-Stat Spillover				150
First Stage F-Stat Spillover				66
	5-Year Difference in Average Annual Wage ($1,000) Post-1975			
Immigration^t_d	0.010***	0.009***	0.005***	0.005***
	(0.002)	(0.003)	(0.001)	(0.002)
$\text{Immigration}^t_{State}$		0.000		
		(0.000)		
Neighbors' Immigration^t_d (Inverse Distance Weight)			0.560***	
			(0.191)	
$\text{Immigration}^t_{100km}$				0.006***
				(0.002)
$\text{Immigration}^t_{250km}$				-0.001
				(0.001)
$\text{Immigration}^t_{500km}$				-0.000
				(0.001)
N	21,976	21,976	21,976	21,976
First Stage F-Stat d	1,166	2,289	3,482	7,967
First Stage F-Stat Spillover		434	165	395
First Stage F-Stat Spillover				157
First Stage F-Stat Spillover				67
Controls:				
Geography FE	division	division	division	division
Time FE	yes	yes	yes	yes

Notes: This table reports the results of our second stage specification for the change in patenting per 100,000 people (population is based on baseline 1970 levels) (Panel A) and the change in the average annual wage ($1,000s) (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 controls. 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 in the same way. Column 4 includes variables, and appropriate instruments, for non-European immigration to counties within 100km, 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)
	5-year Difference in Patenting per 100,000 People Post-1980				
Immigration$_d^t$	0.166***	0.200***	0.485***	0.415***	
	(0.053)	(0.070)	(0.165)	(0.076)	
Average Years Education$_d^t$ × Immigration$_d^t$		0.221***	0.251***		
		(0.068)	(0.079)		
Average Years College$_d^t$ × Immigration$_d^t$				0.887***	
				(0.166)	
$\mathbf{1}${Low Avg. Years Education} × Immigration$_d^t$					1.863
					(4.539)
$\mathbf{1}${Medium Avg. Years Education} × Immigration$_d^t$					0.084*
					(0.044)
$\mathbf{1}${High Avg. Years Education} × Immigration$_d^t$					1.401*
					(0.792)
N	18,846	18,846	18,846	18,846	18,846
	5-year Difference in Average Annual Wage ($100) Post-1975				
Immigration$_d^t$	0.239**	0.290***	0.770*	0.400***	
	(0.091)	(0.058)	(0.419)	(0.078)	
Average Years Education$_d^t$ × Immigration$_d^t$		0.231***	0.221**		
		(0.051)	(0.096)		
Average Years College$_d^t$ × Immigration$_d^t$				0.569***	
				(0.084)	
$\mathbf{1}${Low Avg. Years Education} × Immigration$_d^t$					-0.296
					(0.249)
$\mathbf{1}${Medium Avg. Years Education} × Immigration$_d^t$					0.189***
					(0.069)
$\mathbf{1}${High Avg. Years Education} × Immigration$_d^t$					1.514***
					(0.473)
N	21,976	21,976	21,976	21,976	21,976
Controls:					
Geogrpahy FE	state	state	county	state	state
Time FE	yes	yes	yes	yes	yes

Notes: The table reports the results of our second stage specification for the change in patenting per 100,000 people (population is based on baseline 1970 levels) in the first panel and the 5-year difference in county-level average annual wages ($100s) in the second panel. 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 while 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 .418 (.095) and .299 (.083) 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 × 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 Patents 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)			
sq(Immigration$_d^t$)		-0.001***		-0.004***				
		(0.000)		(0.000)				
Δ Population$_d^t$						0.033***		
						(0.012)		
IHS(Immigration$_d^t$)							1.723***	
							(0.111)	
IHS(Δ 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	911	95	1,202	102	1,202	102	94	16
First Stage F-Stat		11,231		11,879				
Controls:								
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 second stage 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 IHS of patents (columns 5 through 8). Column 1 repeats our main specification while column 2 adds as a dependent variable the square of non-European immigration (1,000s) to d in t as the endogenous variable. 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, while 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.

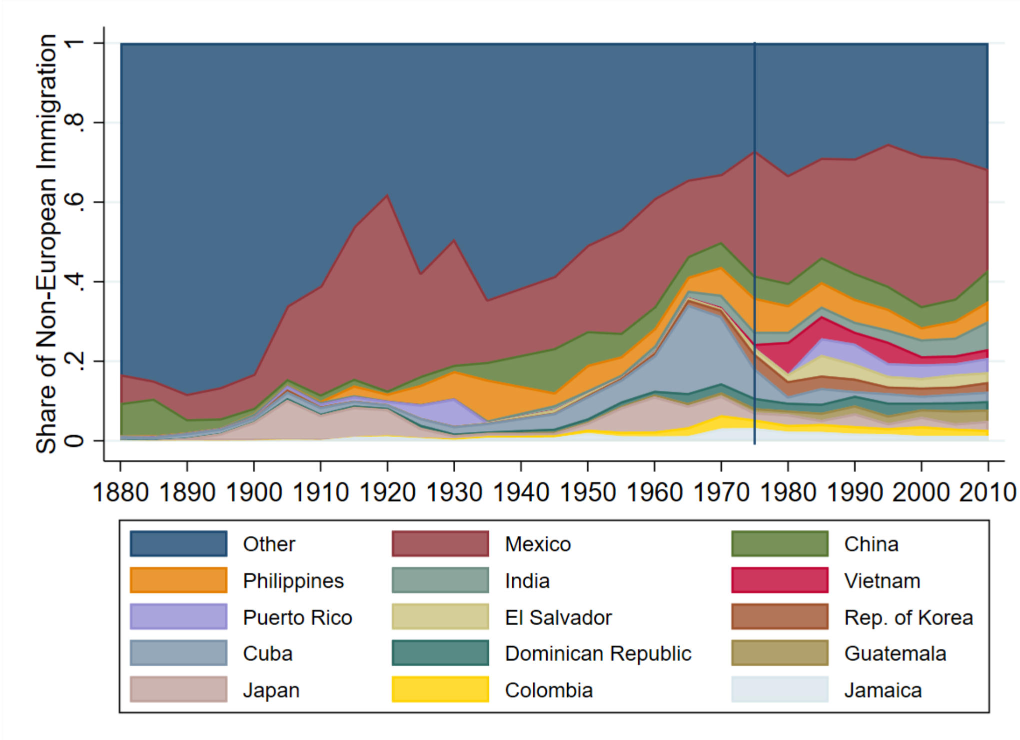

Figure 1: Origins of Non-European Immigrants to the U.S.

Notes: This figure plots the share of non-European immigration into the US from the 14 non-European origin nation (except for Canada which is included in "Other") with the largest cumulative immigration to the US. The figure highlights variation in the push factor, showing that the number of migrants from a given source country o to the US varies by period t. The year 1975 is marked with a vertical line.

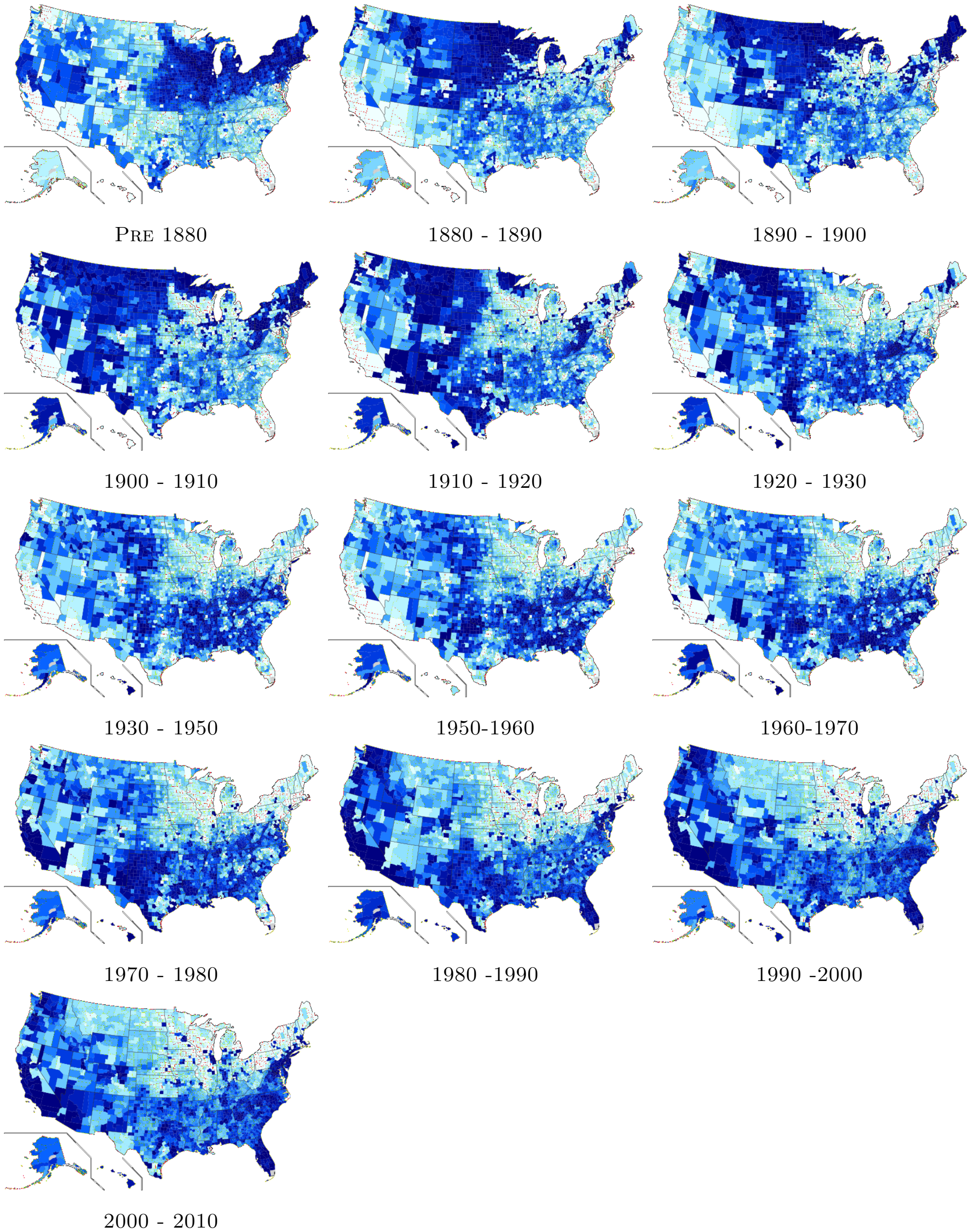

Figure 2: Destinations of Immigrants to the United States, pre-1880 to 2010

Notes: This figure maps immigration flows into US counties by 10-year periods (except between 1930 and 1950). We regress the number of 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.

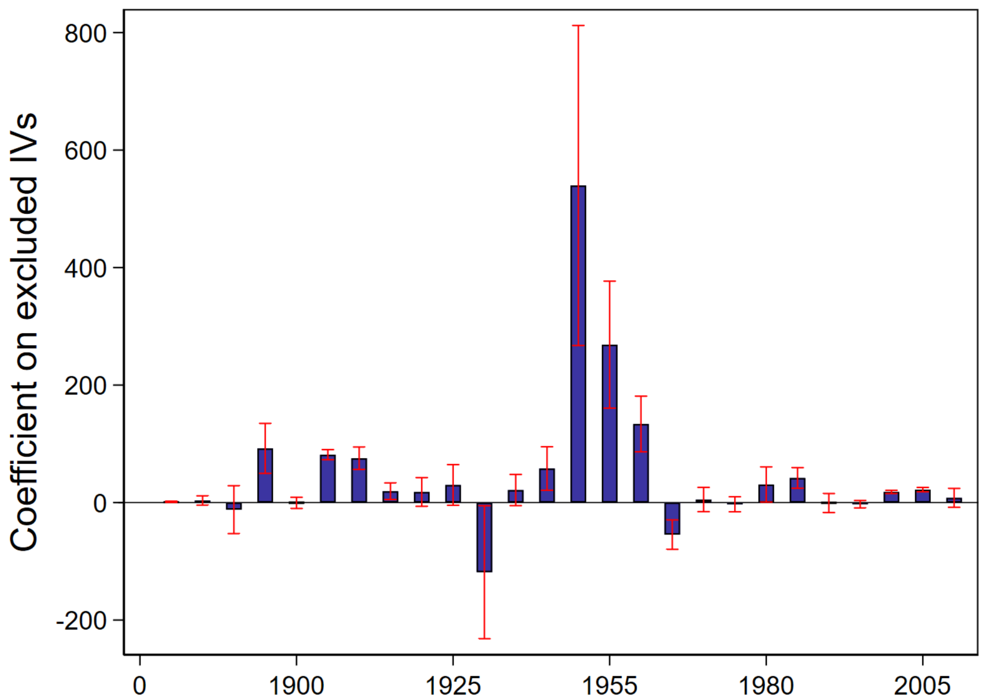

FIGURE 3: STAGE 0: PREDICTING ANCESTRY

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 $b^{\tau}_{r(d)} = b^{\tau}\ \forall r(d)$). The figure shows that we identify variation in current ancestry levels based on push-pull interactions from the full range of time periods in our sample. Standard errors are clustered at the origin country level. (R^2 .5041)

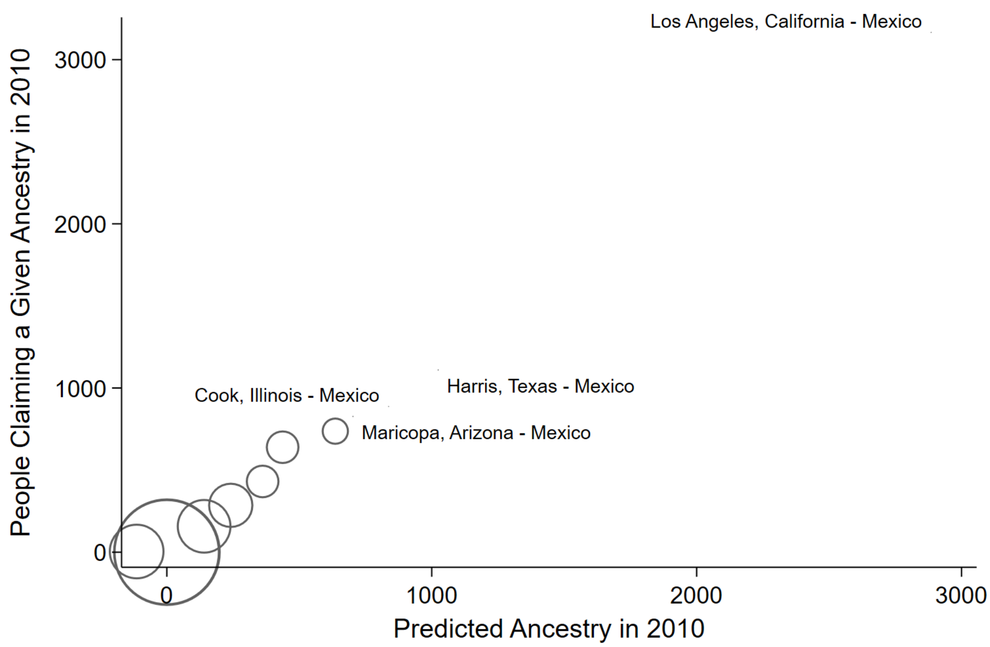

Figure 4: Stage 0: Predicting Ancestry (2010)

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.

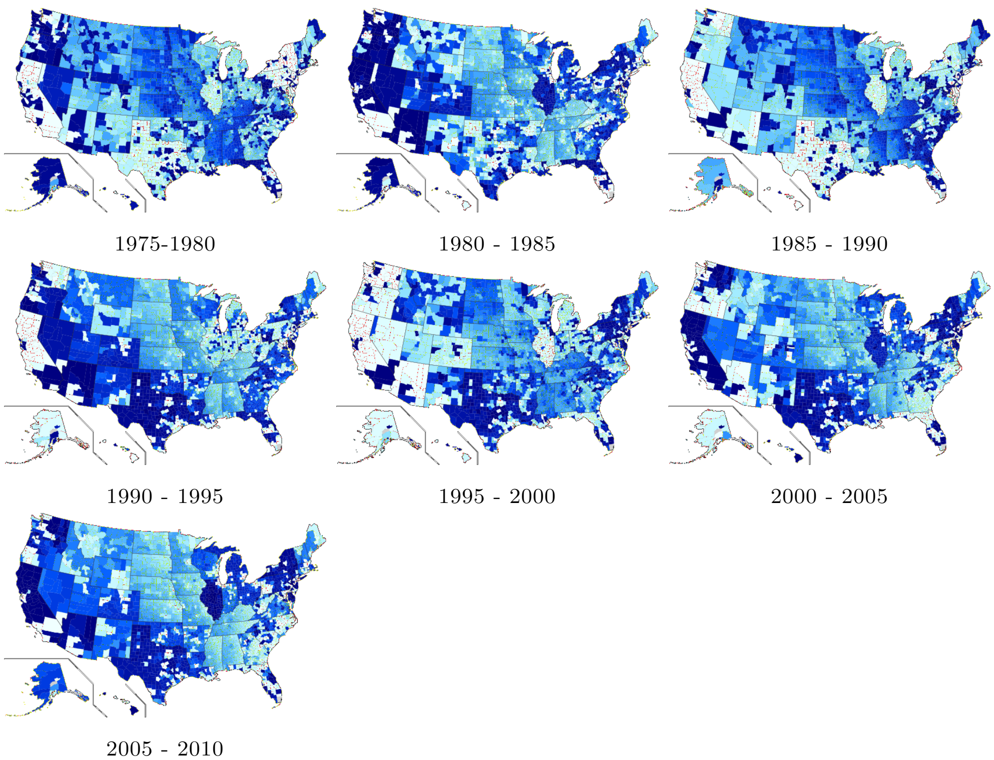

Figure 5: Immigrantion Shock Conditional on County and State-Time FE

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 labor shock used as in instrument in the regression shown in column 3 of Table 3. The map's color coding depicts the 20 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

In order to construct county-level data on migration, ancestry, and ethnicity, we follow the approach of Burchardi et al. (2019). The following section summarizes this approach, highlighting any difference in data construction made in this paper.

Construction of post-1880 immigrantion 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 a 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) 5-year periods based on the year in which each migrant arrived to the US.[19] 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.

[19] For the 1970, 1980 and 1990 censuses, the exact year of arrival for migrants is not provided and so we use as our '5-year periods' the following bins: 1925-34, 1935-44, 1945-49, 1950-54, 1955-59, 1960-64, 1965-70, 1970-74, 1975-80, 1980-84 and 1985-90.

Construction of pre-1880 immigration stock

From the 1880 census, we develop the count of 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 parents' 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 that either were born or migrated to the US after the midyear. Again, we follow Burchardi et al. (2019) in transforming ancestries to 1990 countries and US geographies to 1990 US counties.

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 migrants aged 25 years or more at the time of each respective census. For each 1990 US county d, we sum the number of years an individual is reported to have over all migrants, assigning the midpoint when a range of years of education is provided instead of an exact number of years.

For the 1970 census, we consider 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 4 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 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 5-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.3 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 5-year difference in the dynamism or wage variable.

Wages. The county-level average annual wage for every five years from 1975 to 2010 are taken from the Quarterly Census of Employment and Wages. The data for each period is 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.

Growth Rate Skewness. The growth rate skewness variable for 2010 US counties for each 5-years from 1995 to 2010 is estimated using data from the Longitudinal Business Database. We compute hte Kelly Skewness of employment growth rates for 4-digit sectors and then transition this measure from 2010 to 1990 US counties.

Job Creation and Destruction Rates. Job creation and destruction data is 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.4 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 United States 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 5 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. 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, & 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. These two outcomes concisely justify 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 instantaneous increase in the quality level q_{jt} in that good's production, i.e., 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 there must be 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.

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

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

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

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

i.e., 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: Panel Regressions of Immigration and Population Change on Predicted Immigration Flows at the County Level for 1980 to 2010

	(1)	(2)	(3)	(4)	(5)
	5-Year Non-European Immigration				
$\widehat{Immigration}_d^t$	2.107***	2.107***	2.100***	2.111***	1.580***
	(0.046)	(0.062)	(0.061)	(0.068)	(0.196)
N	21,987	21,987	21,987	6,600	21,987
F-Stat	2,139	1,168	1,202	951	65
R^2	0.741	0.768	0.777	0.771	0.947
	5-Year Population Growth				
$\widehat{Immigration}_d^t$	1.890***	1.890***	1.818***	1.767***	1.921***
	(0.168)	(0.190)	(0.180)	(0.157)	(0.323)
N	21,986	21,986	21,986	6,600	21,986
F-Stat	127	99	102	126	35
R^2	0.233	0.272	0.314	0.370	0.795
Controls:					
Geogrpahy FE	none	division	state	state	county
Time FE	no	yes	yes	yes	yes
MSA Counties	no	no	no	yes	no

Notes: This table reports the results for step 3 of instrument construction, or the coefficient estimates for the first stage specification for non-European immigration (1,000s) (first panel) and population change (1,000s) (second panel). Column 1 provides the results from a regression of non-European immigration or population change on the instrument described in equation (7). Column 2 then adds to that regression time and census division effects while Column 3, our main specification, includes state and time fixed effects. Column 4 shows the first stage estimated on a restricted sample of counties, which is used in analyses of natively MSA-level BDS data, and Column 6 reports results with county and time fixed effects. 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.

Appendix Table 2: 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
Controls:				
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 3: 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
Controls:		
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 4: Placebo Tests of Main Specification

	Coefficient Mean (1)	Coefficient Std. Dev. (2)	Median Standard Error (3)	RHS Rejection Rate (%) (4)
Panel A: First Stage				
Placebo 1	-0.0003	0.013	0.005	0.20
Placebo 2	0.0001	0.013	0.005	0.00
Placebo 3	-0.0117	0.031	0.019	2.70
Panel B: Reduced Form				
Placebo 1	0.0016	0.073	0.036	0.80
Placebo 2	0.0026	0.069	0.034	0.70
Placebo 3	-0.0023	0.069	0.035	4.40

Notes: This table reports the results of a placebo test of the first stage of our standard specification (Panel A) and a reduced form analysis of Equation (1) (Panel B), for changes in patenting per 100,000 people with non-European immigration to d in t as the endogenous variable and state and time fixed effects as controls. Columns 1 and 2 report the average and standard deviation for the coefficient of interest for 1,000 placebo runs. Column 3 then reports the median standard errors of the 1,000 runs. Finally, we report the percentage of runs for which we reject that the coefficient of interest is different from 0 at the 5 percent level on the right-hand side (Column 4). Each row represents a different random assignment strategy for the placebo analysis where we randomly reassign an observation the instrument of another observation: in the sample (Placebo 1), in the same period t (Placebo 2), or in the same period t and census division $r(d)$. Standard errors are clustered by state for all specifications.

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

	Coefficient Mean (1)	Coefficient Std. Dev. (2)	Median Standard Error (3)	Rejection Rate (%) (4)
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, difference in patenting flows per 100,000 people. Column 1 reports the mean value of the coefficient over all placebo regressions while column 2 reports the standard deviation. Column 3 then reports the median standard error for the coefficient of interest over all placebo regressions and finally 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' 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 consistent with the false rejection rates achieved by the AKM method for correcting standard errors outlined in Adão et al. (2019).