Online Appendix
"Immigration, Innovation, and Growth"

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

A Data Appendix

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

To construct county-level data on migration, ancestry, and ethnicity, we follow the approach of Burchardi et al. (2019). We utilize data from each available wave of data from 1880 to 2010 from the Integrated Public Use Microdata Series (IPUMS) (Ruggles et al., 2018-2020). Specifically, we use the 10% sample of the 1880 Census, the 5% sample of the 1900 Census, the 1% sample of the 1910 Census, the 1% sample of the 1920 Census, the 5% sample of the 1930 Census, 1% Form 1 Metro sample of the 1970 Census, 5% State sample of the 1980 Census, 5% State sample of the 1990 Census, 5% sample of the 2000 Census, and the American Community Service 5-Year sample of the 2010 Census. The following section summarizes this approach, highlighting any difference in data construction made in this paper.

Construction of post-1880 immigration flows

We start the construction of our immigration variable by identifying the number of individuals located in a given US geography d at the time of each census who immigrated to the US since the prior census and were born in a historic origin country o (based on the detailed birthplace variable). For each census wave, we then separate this immigration count into (roughly) five-year periods based on the year in which each migrant arrived to the US. For the 1970, 1980, and 1990 censuses, the exact year of arrival for immigrants is not provided, and instead the year of arrival is provided in bins (e.g., a person who arrived in 1964 has a year of arrival of 1960-1964). For these years, we use as our five-year periods the bins that are reported in each census: 1925-34, 1935-44, 1945-49, 1950-54, 1955-59, 1960-64, 1965-70, 1970-74, 1975-80, 1980-84, and 1985-90. We then follow the approach outlined in Burchardi et al. (2019) to transform foreign origin countries, given as birthplaces, to 1990 foreign countries and non-1990 counties and county groups into 1990 counties. Because some foreign birthplaces do not refer to any modern (1990) country, we use population-based weights for transitioning birthplaces to countries (for more details on the weighting scheme, see Burchardi et al. (2019)). We define adult immigrants as those aged 25 years and older at the time of the census.

Construction of pre-1880 immigration stock

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

Construction of ancestry stock

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

Construction of education data for migrants

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

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

A.2 Construction of population data

For the period 1970 to 2010, we collect county-level population data in each census year and intercensal year. The population counts for 2010 were taken directly from the US Census Bureau (the American Community Survey 5-year estimates). All other population counts are taken from the NBER (2018)'s Census U.S. Intercensal County Population Data, 1970-2014. For each period, data are transformed from the given US counties to 1990 US counties using the transition matrices described by Burchardi et al. (2019).

A.3 Construction of patenting data

We utilize data on corporate utility patents with a US assignee from the the US Patent and Trademark Office (USPTO) microdata for the period 1975 to 2010 (USPTO, Accessed: Mar. 28-29, 2022). We translate the location of patents from assignee (or inventor) location to 1990 US counties by mapping the latitude and longitude coordinates onto a shapefile of 1990 counties (obtained from IPUMS NHGIS (Manson et al., 2023)) to estimate the number of patents granted to assignees in each county and year. For our main measure of patenting, we utilize unweighted patent counts with locations based on assignee, but we also consider location based on inventors and weighted patent counts as in Hall et al. (2001). We then construct a variable for the total number of patents filed in each five-year period ending in t, for each measure of patenting, and divide by the 1970 population (100,000 people) to get "per-capita patenting" in t. We then winsorize the variables at the 1% and 99% levels. The main patenting outcome variable is then the difference in this per-capita-patenting variable between $t-1$ and t.

For the inventor-based measure of patenting, we also identify the subset of patents for which all inventors are designated as natives (as opposed to immigrants). Because we do not have information on inventor citizenship, we define native inventors as those whose first patent in the USPTO dataset is filed in the US. This definition of *native inventors* may include patents filed by foreign-born inventors if they first file a patent after moving to the US. Therefore, we also construct the count of of patent flows for *native inventors with a prior US patent* that further restricts the sample to patents filed in period t by only native inventors who have filed another patent in the US prior to t. While this latter definition does not account for individuals who are long-term foreign-born residents in the US that file their first patent at least 5 years after moving, it does remove patents that may have been filed by recent immigrants (as well as removing all patents by first-time native-born inventors or inventors whose previous patents are not contained in the USPTO dataset). Finally, we apply a further restriction limiting the set of patents to those for which all inventors have filed a patent prior to the current period in the same US county (or at least one of the same counties if the inventor filed in multiple locations in period t).

A.4 Construction of native wages data

We construct variables for native wages in each census year from 1980 to 2000 using data from the 1980 5% State sample, 1990 5% State sample, and 2000 5% Census sample (Ruggles et al., 2018-2020). In each year, we limit the sample to the pre-tax wage and salary income (incwage) for individuals aged 25 and older who were born in the US and are employed (empstat is equal to 1), referred to here as natives. We then further limit the sample to natives who report that they lived in the same county five years prior to the census year to identify wages of native non-movers. Additionally, we subset the data based on the education level of the individuals to estimate the wages of native non-movers with education levels of: less than high school, high school, some college (1-3 years), 4 years of college, and 5 or more years of college. We use the Consumer Price Index provided in IPUMS USA (CPI99) to adjust wages to a common dollar year, 1999. We then follow the same method as that used in Burchardi et al. (2019) to transform wages for county groups into 1990 US counties. Finally, we determine average wages in each county using the person weight (PERWT) for the selected sample and generate a variable for wage growth in each county that is the 10-year difference in average annual wages for native non-movers.

A.5 Construction of business dynamism data

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

Wages. The county-level average annual wage for every five years from 1975 to 2010 is taken from the Quarterly Census of Employment and Wages (BLS, 2018). The data for each period are then transformed from the US counties for that period to 1990 US counties using the transition matrices developed in Burchardi et al. (2019) and then converted to 2010 US dollars using the Personal Consumption Expenditures Price Index from the Bureau of Economic Analysis (BEA, 2018). We generate this county-level average annual wage for all industries as well as manufacturing (SIC 20-39 and NAICS 31-33) and services (SIC 60-67 and NAICS 52-53).

Growth Rate Skewness. The growth rate skewness variable for 2010 US counties for each five years from 1995 to 2010 is estimated using data from the Longitudinal Business Database (US Census Bureau, 2018b). We compute the Kelly Skewness of employment growth rates across 4-digit sectors, and then transition this measure from 2010 to 1990 US counties.

Job Creation and Destruction Rates. Job creation and destruction data are taken from the Business Dynamics Statistics (US Census Bureau, 2018a) for metropolitan statistical areas (MSAs) and transitioned to 1990 US counties based on weights derived from 1990 population data.

A.6 Construction of local output data

Local output data come from the BEA (2021)'s county-level GDP estimates for five-year periods for the available window from 2001 to 2019. These estimates are used to calculate the autocorrelation of county-level output per capita, a target moment of the structural model estimation.

A.7 Additional Tables and Figures

Appendix Table 1: Assignment of States to Census Divisions (US Census Bureau, 2013)

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

Appendix Table 2: Robustness - Alternative Share-Based Instruments and Rejection Rates

	Δ^{5yr} Patent Flows Per Capita		
Specification:	*Predicted Ancestry Shares (Baseline)*	*Realized Immigration Shares (Card, 2001)*	*Realized Ancestry Shares*
	(1)	(2)	(3)
Adão et al. (2019) First Stage False Rejection Rate (%)	3.8	27.4 *Overreject*	24.5 *Overreject*
$\text{Immigration}_{d,t}$	0.202^{**}	0.161	0.163
	(0.084)	(*0.075*)	(*0.071*)
N	18846	18846	18846
First Stage F-Stat	656	695	361
Instrument Functional Form:			
Instrumented Ancestry	Yes	No	No
Push Factor Leave-Out	Yes	No	No
Controls:			
Geography FE	State	State	State
Time FE	Yes	Yes	Yes

Notes: This table displays the results of estimating equation (1), where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) and the endogenous variable is non-European immigration (1,000s) to county d at time t. Column 1 uses our baseline instrument but with predicted ancestry shares, as opposed to predicted ancestry in levels. Column 2 is an instrument based on Card (2001) that utilizes realized immigration shares. Column 3 replaces the realized immigration shares in column 2 with realized ancestry shares. We report the first-stage F-statistic on the excluded instrument for each specification. For each instrument, we report the false rejection rate in the first-stage regression for a robustness test that follows the method proposed by Adão et al. (2019). See Appendix Table 3 for details. 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: Results from Placebo Analysis Based on Adão et al (2019)

	(1)	(2)	(3)	(4)
	Coefficient (Mean)	Coefficient (St. Dev.)	Standard Error (Median)	Rejection Rate (%)
Panel A: Realized Immigration Shares (Card, 2001)				
First Stage	-0.0038	0.0703	0.0377	24.5
Reduced Form	-0.0003	0.0182	0.0117	15.1
Panel B: Realized Ancestry Shares				
First Stage	-0.0040	0.0791	0.0398	27.4
Reduced Form	-0.0004	0.0200	0.0128	14.9
Panel C: Predicted Ancestry Shares (Baseline Instrument)				
First Stage	-0.0016	0.0397	0.0249	3.8
Reduced Form	0.0004	0.0946	0.0801	9.1

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 realized immigration shares (as in Card (2001)), realized ancestry shares, and our predicted baseline ancestry shares (as in the ancestry-share version of our baseline instrument). We then run 1,000 placebo regressions of the endogenous immigration variabale on the placebo variables for the Card (2001) instrument (Panel A), the Card-style instrument that uses ancestry shares (Panel B), and our ancestry-share instrument (Panel C); we also run the comparable reduced-form regressions where the dependent variable is our primary measure of patenting, the five-year difference in patenting flows per 100,000 people. Column 1 reports the mean value of the coefficient over all placebo regressions, whereas column 2 reports the standard deviation. Column 3 then reports the median standard error for the coefficient of interest over all placebo regressions, and column 4 reports the fraction of placebo regressions for which we reject the null hypothesis of no effect at the 5% statistical significance threshold. As shown, the traditional shift-share instrument suffers from the over-rejection identified in Adão et al. (2019) with false rejection rates of 24.6% in the first stage and 13.6% in the reduced-form specification. The ancestry-share version of our baseline instrument has false rejection rates of 4.7% (first stage) and 8.4% (reduced form). The latter is similar to the false rejection rates reported in Adão et al. (2019) when using their proposed standard error correction (labelled "AKM").

Appendix Table 4: Robustness - County Level vs. State Level Regressions

	Δ^{5yr} Patent Flows Per Capita			
	County Level		*State Level*	
	Predicted Ancestry Shares	*Realized Ancestry Shares*	*Predicted Ancestry Shares*	*Realized Ancestry Shares*
	(1)	(2)	(3)	(4)
$\text{Immigration}_{d,t}$	0.2021**	0.1626**	0.0005***	0.0005***
	(0.0841)	(0.0713)	(0.0002)	(0.0001)
N	18,846	18,846	306	306
First Stage F-Stat	656	361	97	1,154
Geography FE	State	State	Division	Division
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. Columns 1 and 3 use our baseline instrument but with predicted ancestry shares, as opposed to predicted ancestry in levels, and columns 2 and 4 use the comparable instrument but with realized ancestry shares. Columns 1 and 2 report a county level analysis while columns 3 and 4 repeat each regression at the state level. We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 5: Robustness - Alternative Instruments for Immigration

	Alternative Instrument Constructions			
	Leave-Out Correlated Counties	*Leave-Out Own Continent*	*Ancestry in 1975 only*	*Stop Push-Pull in 1960*
	(1)	(2)	(3)	(4)
Panel A	Δ^{5yr} Patent Flows Per Capita			
$\text{Immigration}_{d,t}$	0.096***	0.122***	0.111***	0.101***
	(0.035)	(0.045)	(0.040)	(0.038)
N	18,846	18,846	18,846	18,846
First Stage F-Stat	127	828	1,171	1,750
AR Wald F-Test p-value	0.002	0.016	0.011	0.012
Panel B	IHS of Patent Flows Per Capita			
$\text{IHS}(\text{Immigration}_{d,t})$	1.672***	1.649***	1.644***	1.725***
	(0.178)	(0.159)	(0.148)	(0.161)
N	21,987	21,987	21,987	21,987
First Stage F-Stat	63	56	109	54
AR Wald F-Test p-value	0.000	0.000	0.000	0.000
Geography FE	State	State	State	State
Time FE	Yes	Yes	Yes	Yes

Notes: Panel A of 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; Panel B reports the comparable regression where the dependent variable is the IHS of patenting per 100,000 people and the endogenous variable is the IHS of non-European immigration (1,000s). Column 1 takes the sum over push-pull interaction up to the year 1960 only in Step 1 to create an instrument for ancestry. Column 2 replaces predicted ancestry in $t-1$ with predicted ancestry in 1975 for all periods. Column 3 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 4 replaces the economic pull factor in Step 1 with the share of all migrants who settle in d but excluding migrants from the same continent as o (instead of using only European migrants). We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 6: Robustness - First Stage Controlling for Lagged Immigration Shocks

	$\text{Immigration}_{d,t}$	
	(1)	(2)
$\text{ImmigrationShock}_{d,t}$	1.580***	1.623***
	(0.196)	(0.222)
$\text{ImmigrationShock}_{d,t-1}$		-0.064
		(0.232)
N	21,987	18,846
R^2	0.495	0.572
Geography FE	County	County
Time FE	Yes	Yes

Notes: This table reports the results for the coefficient estimates for the first-stage specification for non-European immigration (1,000s) for the instrument described in equation (5). Column 1 provides our baseline first stage regression with county and time fixed effects while column 2 adds the lagged immigration shock as a control. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix Table 7: Panel Regression of 5-Year Difference in Patenting per 100,000 People on Immigration using Alternative Patent Counts

	Δ^{5yr} Patent Flows Per Capita			
	Assignee (Unweighted)	*Assignee (Cite Weight)*	*Inventors (Unweighted)*	*Inventors (Cite Weight)*
	(1)	(2)	(3)	(4)
$\text{Immigration}_{d,t}$	0.122***	0.150***	0.085**	0.137***
	(0.045)	(0.050)	(0.037)	(0.045)
N	18,846	18,846	18,846	18,846
First Stage F-Stat	911	911	911	911
AR Wald F-Test p-value	0.014	0.008	0.037	0.007
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. We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 8: Permutation Tests for Main Specification

	(1)	(2)	(3)	(4)
	Coefficient		Standard Error	RHS Rejection
	(Mean)	(St. Dev.)	(Mean)	Rate (%)
Panel A: First Stage				
Placebo 1	0.0007	0.018	0.008	0.40
Placebo 2	-0.0006	0.013	0.008	0.10
Placebo 3	-0.0127	0.031	0.021	1.90
Panel B: Reduced Form				
Placebo 1	0.0035	0.054	0.041	1.30
Placebo 2	0.0009	0.049	0.037	1.50
Placebo 3	0.0034	0.069	0.045	5.20

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

Appendix Table 9: County-Level Panel Regressions of Difference in Patenting on Population Growth

	Δ^{5yr} Patent Flows Per Capita		
	(1)	(2)	(3)
Panel A: OLS			
Δ Population$_{d,t}$	0.281***	0.279***	0.157*
	(0.086)	(0.087)	(0.079)
N	18,846	18,840	18,846
R^2	0.046	0.068	0.190
Panel B: IV			
Δ Population$_{d,t}$	0.136***	0.130***	0.140**
	(0.044)	(0.045)	(0.069)
N	18,846	18,840	18,846
First Stage F-Stat	110	103	63
AR Wald F-Test p-value	0.014	0.021	0.013
Panel C: First Stage	Δ $Population_{d,t}$		
Immigration Shock ($\hat{I}_{d,t}$)	1.897***	1.888***	2.081***
	(0.181)	(0.186)	(0.263)
N	18,846	18,840	18,846
R^2	0.324	0.340	0.804
Geography FE	State	State	County
Time FE	Yes	Yes	Yes
State-Time FE	No	Yes	No

Notes: Panels A and B of this table report the OLS and IV results, respectively, of the estimation of equation (1) where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) in county d in the five-year period ending in t and the endogenous variable is population growth (1,000s) in d and period t. Panel C reports the results for step 3 of instrument construction, or the coefficient estimates for the first-stage specification for population change (1,000s) for the instrument described in equation (5). The table includes the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test for each of the IV specifications. Standard errors are clustered by state for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix Table 10: Robustness - Additional Controls from Baseline Year (1970)

	Δ^{5yr} Patent Flows Per Capita				
	(1)	(2)	(3)	(4)	(5)
$\text{Immigration}_{d,t}$	0.122***	0.125**	0.125***	0.106**	0.090**
	(0.045)	(0.048)	(0.045)	(0.040)	(0.036)
Population Density (1970)		-0.001			
		(0.001)			
Patents per 1,000 People (1975)			-3.377		
			(2.313)		
Share High School Education (1970)				51.754***	
				(10.185)	
Share 4+ Years College (1970)					178.858***
					(25.374)
N	18,846	18,840	18,840	18,846	18,846
First Stage F-Stat	911	2,062	920	945	1,017
AR Wald F-Test p-value	0.014	0.016	0.014	0.018	0.021
Geography FE	State	State	State	State	State
Time FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), where the dependent variable is the change in patenting per 100,000 people (population is based on baseline 1970 levels) and the endogenous variable is non-European immigration (1,000s) to d in t. Column 1 repeats our main specification, whereas columns 2-5 add as a control county d's population density in 1970, patents filed in 1975 per 1,000 people (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. We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 11: Robustness - Alternative Samples

	Δ^{5yr} Patent Flows Per Capita				
	Mexico	*China*	*India*	*Philippines*	*Vietnam*
	(1)	(2)	(3)	(4)	(5)
Panel A: Excluding Given Country					
$\text{Immigration}_{d,t}$	0.091***	0.123***	0.122***	0.122***	0.122***
	(0.028)	(0.046)	(0.045)	(0.044)	(0.045)
N	18,846	18,846	18,846	18,846	18,846
First Stage F-Stat	666	1,576	1,267	1,261	1,179
AR Wald F-Test p-value	0.003	0.015	0.014	0.014	0.014
Panel B: Including Only Given Country					
$\text{Immigration}_{d,t}$	0.125***	0.089***	0.145***	0.140**	0.125*
	(0.047)	(0.028)	(0.039)	(0.054)	(0.069)
N	18,846	18,846	18,846	18,846	18,846
First Stage F-Stat	2,094	535	318	22	2
AR Wald F-Test p-value	0.015	0.003	0.001	0.000	0.148
Geography FE	State	State	State	State	State
Time FE	Yes	Yes	Yes	Yes	Yes

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

Appendix Table 12: Mechanisms: Patents by Inventor Type

	Δ^{5yr} Patent Flows Per Capita			
	All Inventors	*Domestic Inventors*	*Immigrant Inventors*	*Teams of Domestic & Immigrant Inventors*
	(1)	(2)	(3)	(4)
$\text{Immigration}_{d,t}$	0.085**	0.069**	0.003***	0.009**
	(0.037)	(0.030)	(0.001)	(0.004)
N	18,846	18,846	18,846	18,846
First Stage F-Stat	911	911	911	911
AR Wald F-Test p-value	0.037	0.038	0.004	0.027
Geography FE	State	State	State	State
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), for changes in patenting per 100,000 people with non-European immigration to d in t as the endogenous variable. Column 1 uses our baseline patenting variable but with a patent's county designated based on inventor location (as opposed to assignee location). Column 2 repeats this specification but limits to patents with only domestic inventors, defined as those whose first patent was filed in the US (92% of all patents). Column 3 limits patents in the dependent variable to those with only immigrant inventors, defined as those whose first patent was filed abroad and have at least one patent in the US (1% of all patents). Finally, Column 4 limits patents in the dependent variable to only those with domestic and immigrant inventor teams (4% of all patents). Patents with at least one foreign inventor, defined as those with all patents filed abroad, make up the remaining 3% of all patents. We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 13: Mechanisms: Domestic Non-Mover Innovators

	Δ^{5yr} Patent Flows Per Capita			
	All Inventors	*Only Domestic Inventors*	*Only Domestic Inventors with Prior US Patent*	*Only Domestic Inventors with Prior US Patent* in Same County
	(1)	(2)	(3)	(4)
Immigration$_{d,t}$	0.085**	0.069**	0.040**	0.033**
	(0.037)	(0.030)	(0.018)	(0.014)
N	18,846	18,846	18,846	18,846
First Stage F-Stat	911	911	911	911
AR Wald F-Test p-value	0.037	0.038	0.037	0.032
Geography FE	State	State	State	State
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), for changes in patenting per 100,000 people with non-European immigration to d in t as the endogenous variable. Column 1 uses our baseline patenting variable but with a patent's county designated based on inventor location (as opposed to assignee location). Column 2 repeats this specification but limits to patents with only domestic inventors, or inventors whose first patent was filed in the US (92% of all patents). Column 3 further limits patents in the dependent variable to those with only domestic inventors who have filed at least one patent in the US prior to the current period (40% of all patents). Finally, Column 4 further limits patents in the dependent variable to those with only domestic inventors who have filed at least one patent prior to the current period in the same US county or at least one of the same counties in the case that they file in multiple locations in period t (32% of all patents). We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 14: Spillovers Analysis

	(1)	(2)	(3)	(4)
Panel A	Δ^{5yr} Patent Flows Per Capita			
$\text{Immigration}_{d,t}$	0.137***	0.124***	0.076*	0.076
	(0.048)	(0.045)	(0.044)	(0.048)
$\text{Immigration}_{s(d),t}$		0.007***		
		(0.002)		
Neighbors' $\text{Immigration}_{n(d),t}$ (Inverse Distance Weight)			6.916***	
			(2.061)	
$\text{Immigration}_{100km(d),t}$				0.077***
				(0.026)
$\text{Immigration}_{250km(d),t}$				0.005
				(0.005)
$\text{Immigration}_{500km(d),t}$				0.004
				(0.004)
N	18,846	18,846	18,846	18,846
First Stage F-Stat (first coefficient)	876	1,129	2,175	6,065
First Stage F-Stat (second coefficient)		807	162	383
First Stage F-Stat (third coefficient)				150
First Stage F-Stat (fourth coefficient)				66
AR Wald F-Test p-value	0.011	0.000	0.000	0.000
Panel B	Δ^{5yr} Wages			
$\text{Immigration}_{d,t}$	0.178***	0.180***	0.094***	0.105***
	(0.036)	(0.051)	(0.022)	(0.033)
$\text{Immigration}_{s(d),t}$		-0.001		
		(0.014)		
Neighbors' $\text{Immigration}_{n(d),t}$ (Inverse Distance Weight)			9.924***	
			(3.309)	
$\text{Immigration}_{100km(d),t}$				0.104***
				(0.038)
$\text{Immigration}_{250km(d),t}$				-0.012
				(0.020)
$\text{Immigration}_{500km(d),t}$				-0.004
				(0.016)
N	21,977	21,977	21,977	21,977
First Stage F-Stat (first coefficient)	872	881	3,065	7,031
First Stage F-Stat (second coefficient)		840	175	437
First Stage F-Stat (third coefficient)				160
First Stage F-Stat (fourth coefficient)				66
AR Wald F-Test p-value	0.000	0.000	0.000	0.000
Geography FE	Division	Division	Division	Division
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of our IV specification (1) for the change in patenting per 100,000 people (population is based on baseline 1970 levels) (Panel A) and the change in the real average annual wage ($100s, at 2010 prices) (Panel B) with non-European immigration (Panel A) and immigration limited to those aged 25+ (Panel B) (1,000s) to d in t as the endogenous variable. The first column repeats our baseline specification but with census division fixed effects. Column 2 adds as a second endogenous variable: total non-European immigration to the state in which d is located, excluding own-immigration to d, in period t and a comparable instrument. Column 3 adds as a second endogenous variable the inverse-distance-weighted sum of non-European immigration to all counties in the US, excluding own-immigration, and an instrument constructed analogously. Column 4 includes variables, and appropriate instruments, for non-European immigration to counties within 100km (excluding d), 100km to 250km, and 250km to 500km of county d. For each specification we report the first-stage F-statistic(s), utilizing the F-statistic described in Angrist and Pischke (2009, p. 217-218) in the case of multiple endogenous variables. We report the p-value for the Anderson-Rubin Wald F test 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 15: Spillover Analysis – First Stage

	Column 1	Column 2		Column 3		Column 4			
	$Immigration_d^t$ (1)	$Immigration_{d,t}$ (2)	$Immigration_{s(d),t}$ (3)	$Immigration_{d,t}$ (4)	Neighbors' $Immigration_{n(d),t}$ (5)	$Immigration_{d,t}$ (6)	$Immigration_{100km(d),t}$ (7)	$Immigration_{250km(d),t}$ (8)	$Immigration_{500km(d),t}$ (9)
Immigration Shock ($\hat{I}_{d,t}$)	2.130***	2.126***	0.487**	2.093***	0.001	2.094***	-0.379	-0.080	0.345
	(0.072)	(0.072)	(0.232)	(0.055)	(0.002)	(0.058)	(0.257)	(0.264)	(0.450)
State Immigration Shock ($\hat{I}_{s(d),t}$)		0.003	2.778***						
		(0.002)	(0.125)						
Neighbors' Immigration Shock ($\hat{I}_{n(d),t}$)				4.938*	2.388***				
				(2.730)	(0.369)				
Immigration Shock 100km ($\hat{I}_{100km(d),t}$)						0.058	3.404***	-0.071	-1.264
						(0.040)	(0.993)	(0.322)	(0.764)
Immigration Shock 250km ($\hat{I}_{250km(d),t}$)						0.006	-0.047	2.623***	-0.617*
						(0.011)	(0.095)	(0.387)	(0.315)
Immigration Shock 500km ($\hat{I}_{500km(d),t}$)						-0.006	-0.201*	-0.339	2.030***
						(0.007)	(0.120)	(0.232)	(0.263)
N	18,846	18,846	18,846	18,846	18,846	18,846	18,846	18,846	18,846
First Stage F-Stat	876	1,129	807	2,175	162	6,065	383	150	66
Geogrpahy FE	Division	Division	Division	Division	Division	Division	Division	Division	Division
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of the first-stage regressions for the IV regressions shown in Table 14. Column 1 of this table provides the first stage regression for column 1 of Table 14. The first stages for column 2 of Table 14 are shown in columns 2 and 3 of this table while those for column 3 in Table 14 are shown in columns 4 and 5 of this table. Finally, columns 6-9 display the first-stage regressions for column 4 of Table 14. For each specification, we report the first-stage F-statistic for the IV estimation in Table 14, 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.

Appendix Table 16: Panel Regressions of Inflows of Native Migrants on Non-European Immigration

	Inflows of Internal Migrants	
	All Natives	*Non-Hispanic White Natives*
	(1)	(2)
$\text{Immigration}_{d,t}$	3.675***	2.100***
	(0.616)	(0.406)
N	9,415	9,415
First Stage F-Stat	3,484	3,484
AR Wald F-Test p-value	0.000	0.000
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. We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 17: Immigration and Economic Dynamism

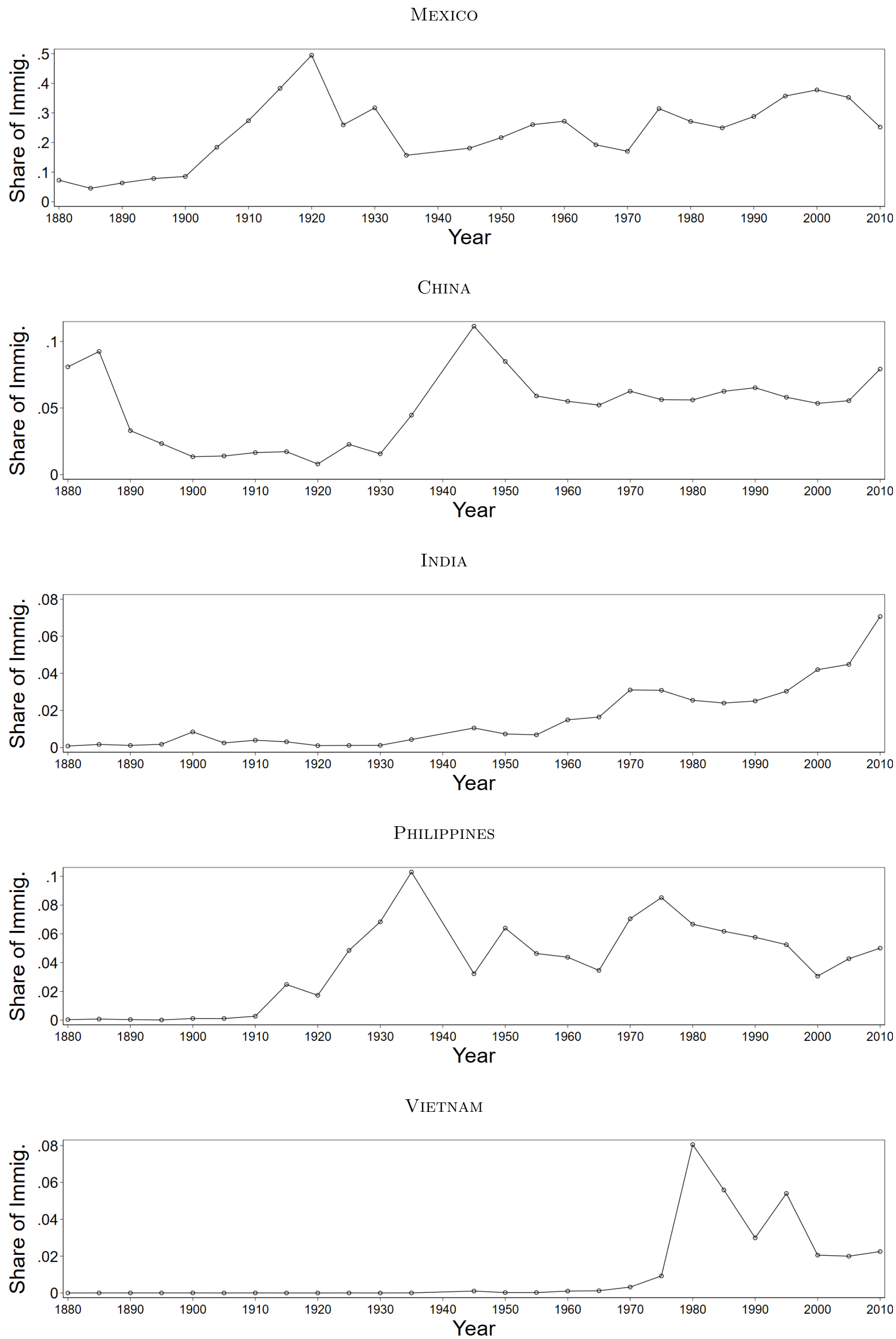
	Δ^{5yr} Job Creation Rate	Δ^{5yr} Job Destruction Rate	Δ^{5yr} Job Growth Rate Skewness
	(1)	(2)	(3)
$\text{Immigration}_{d,t}$	0.176***	0.152***	0.019***
	(0.033)	(0.035)	(0.004)
N	6,588	6,588	12,560
First Stage F-Stat	951	951	151
AR Wald F-Test p-value	0.000	0.000	0.000
Geography FE	State	State	State
Time FE	Yes	Yes	Yes

Notes: This table reports the results of our IV specification, described in equation (1), for each of our dependent variables with non-European immigration (1,000s) to d in t as the endogenous variable. Columns 1 and 2 report the results 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. We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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 18: Time Path of Innovation and Wage Elasticities

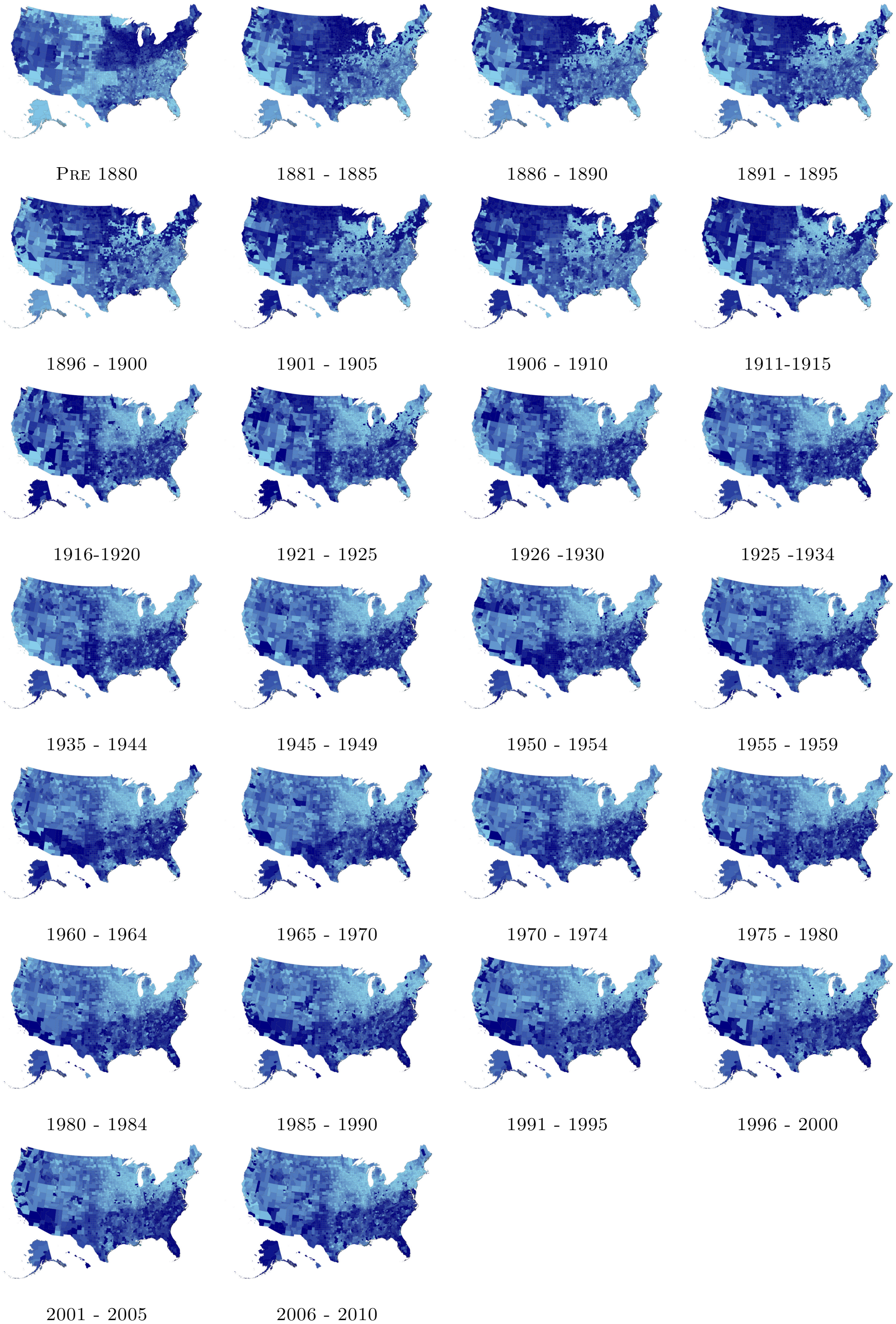
	IHS($Y_{d,t}$)	IHS($Y_{d,t+1}$)	IHS($Y_{d,t+2}$)
	(1)	(2)	(3)
Panel A:		Patent Flows	
IHS($Immigration_{d,t}$)	1.616***	1.712***	1.700***
	(0.156)	(0.274)	(0.351)
N	15,705	15,705	15,705
First Stage F-Stat	109	31	15
AR Wald F-Test p-value	0.000	0.000	0.001
Panel B:		Wages	
IHS($Immigration_{d,t}$)	0.128***	0.151***	0.191*
	(0.016)	(0.050)	(0.099)
N	15,695	15,695	15,695
First Stage F-Stat	152	26	11
AR Wald F-Test p-value	0.000	0.004	0.158
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 inverse hyperbolic sine (IHS) of patents (Panel A) or IHS of wages (Panel B) and the endogenous variable is the IHS of non-European immigration (1,000s) to d in t. Columns 1 through 3 report regression results where the outcome is measured in period t, $t+1$, and $t+2$, respectively; for the regressions in columns 2 and 3 we include controls for the immigration shock in $t+1$ and the immigration shocks in $t+1$ and $t+2$, respectively. We report the first-stage F-statistic on the excluded instrument and the p-value for the Anderson-Rubin Wald F test 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.



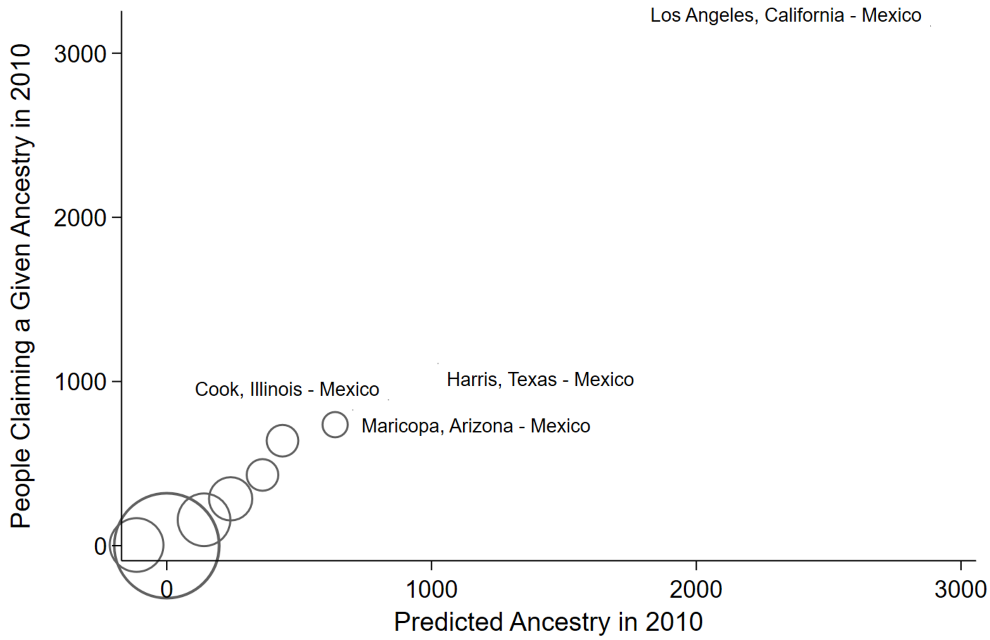
Appendix Figure 1: Share Non-European Immigrants to the US by Origin Country

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



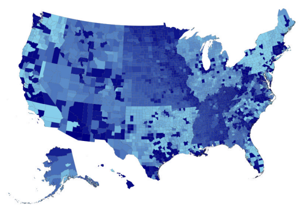
Appendix Figure 2: Destinations of European Immigrants to the US

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

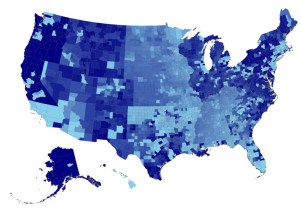


Appendix Figure 3: Step 1 – Predicting Ancestry (2010)

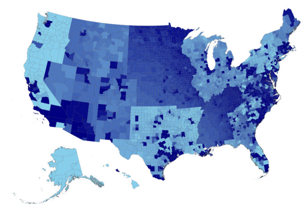
Notes: This figure is a binned scatter plot of actual ancestry in 2010 against predicted ancestry, as given in equation (3), where bins are fixed based on predicted ancestry and the size of each circle indicates the log number of observations in a given bin. The labeled counties are those with the highest number of individuals declaring a given ancestry in 2010. The corresponding regression of $A_{o,d,2010}$ on $\hat{A}_{o,d,2010}$, as defined in equation (3), yields an R^2 of 74.9%.



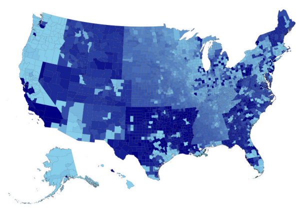
1975-1980



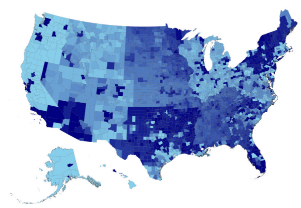
1980 - 1985



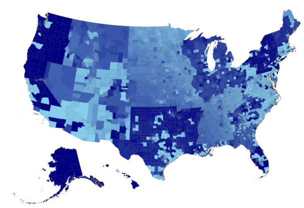
1985 - 1990



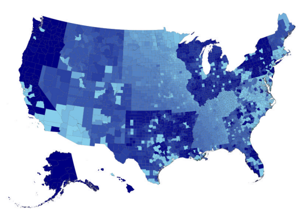
1990 - 1995



1995 - 2000



2000 - 2005



2005 - 2010

Appendix Figure 4: Immigration Shock Conditional on County and Time Fixed Effects

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 year fixed effects, and calculate the residuals. This figure provides a visualization for the immigration shocks used as instruments in the regression shown in column 3 of Table 3. The map's color coding depicts the 200 quantiles of the residuals across counties and within census periods. Darker colors indicate a higher quantile.

B Structural Model and Estimation Appendix

This appendix provides an equilibrium definition and balanced growth path analysis of the quantitative equilibrium regional endogenous growth model. This appendix also provides information on the solution and simulation of the model away from the balanced growth path, together with details of the structural estimation procedure as well as various model extensions.

B.1 Structural Model

B.1.1 Set-up

There are $d = 1, ..., D$ destination regions. There are $o = 1, ..., O$ origin nations. Time t is discrete.

Final Goods Production. A final good is produced by a firm in d with the technology

$$Y_t = Z_{d,t} Q_{d,t} L^{\alpha}_{Y,d,t}$$

where $Q_{d,t}$ is the number of ideas used by the firm in d at time t and $L_{Y,d,t}$ is labor used for production purposes by the firm in region d at time t. The elasticity of output to labor satisfies $0 < \alpha < 1$. The stationary exogenous shock $Z_{d,t}$ to production efficiency satisfies

$$\ln Z_{d,t} = \rho \ln Z_{d,t-1} + \epsilon_{d,t},$$

where the autocorrelation of the shock satisfies $0 < \rho < 1$. We have positive variance with $\epsilon_{d,t} \sim \mathcal{N}(0, \sigma^2_\epsilon)$, where $\sigma_\epsilon > 0$.

Idea Production. A mass $N_{d,t}$ of new ideas is produced each period by a research firm with an innovation or idea production technology given by

$$N_{d,t} = L^{\gamma}_{N,d,t} \bar{Q}^{\zeta}_{d,t-1},$$

where the elasticity of innovation to researchers satisfies $0 < \gamma < 1$ and the elasticity of innovation to ideas satisfies $0 < \zeta < 1$ where $0 < \gamma + \zeta \leq 1$. There are positive externalities in the growth process, through which past ideas aid in the production of new ideas. $\bar{Q}_{d,t-1}$ is a weighted average of varieties invented across regions at time $t - 1$ described further below. Note that we have that the mass of varieties invented in region d evolves according to

$$Q_{d,t} = N_{d,t} + Q_{d,t-1}.$$

In our baseline in the text we assume that $\zeta = 1 - \gamma$, but we quantitatively consider alternative cases allowing for $\gamma + \zeta \neq 1$ below in Figure 7.

Regional Idea Aggregates & Spillovers. The mass of ideas useful to researchers in region d at time is given by

$$\bar{Q}_{d,t} = \prod_{f=1}^{D} Q^{\alpha(d,f)}_{ft}.$$

The elasticity of region d's research-effective ideas to region f's invented ideas, is $\alpha(d,f)$. The elasticities sum to 1 for each region d, i.e.,

$$\sum_f \alpha(d,f) = 1,$$

and the elasticities are proportional to a term declining in the physical distance $\tilde{d}(d,f)$ between d and f with

$$\alpha(d,f) \propto 1 - \delta \tilde{d}(d,f)$$

for some value $\delta \geq 0$. Note that two extreme cases are nested: no idea spillovers ($\delta = \infty$) and full idea spillovers ($\delta = 0$). The baseline model described in the text imposes no spillovers with $\delta = \infty$, in which case $\bar{Q}_{d,t} = Q_{d,t}$. We quantitatively consider a case of full national idea spillovers with $\delta = 0$ below in Figure 5.

Population Structure: Residents, Immigrants, Domestic Migrants, and Ancestry. The population of region d is made up of current residents, domestic migrants, and immigrants, all of whom are members of the labor force. Growth in the population comes only from immigrants and domestic migrants, with accumulation of the labor force over time according to

$$L_{d,t+1} = \sum_{o=1}^{O} I_{o,d,t} + \sum_{d'=1}^{D} M_{d',d,t} + (1-\mu) L_{d,t}.$$

Above, $I_{o,d,t}$ is the mass of immigrants from origin o in destination d at time t. The sum of migrants across all destinations d from a given origin is

$$I_{o,t} = \sum_d I_{o,d,t},$$

and the sum of migrants across all origins o in a given destination is

$$I_{d,t} = \sum_o I_{o,d,t}.$$

Domestic migrants from origin county d' to destination county d at time t are given by $M_{d',d,t}$. Domestic migrants of ancestry o from county d' to county d at time t are indicated by $M_{o,d',d,t}$. A randomly selected fraction $\mu \in (0,1)$ of the domestic population receives a migration opportunity shock, in which case they are able to domestically migrate to any county including, potentially, their own according to the optimization problem laid out below. The total (gross) domestic outmigration from a county d' at time t is given by

$$M_{d',.,t} = \sum_{d=1}^{D} M_{d',d,t},$$

and we must have given the random assignment of the migration shock that $M_{d',.,t} = \mu L_{d',t}$. The total (gross) domestic inmigration to a county d at time t is given by

$$M_{.,d,t} = \sum_{d'=1}^{D} M_{d',d,t}.$$

The stock of residents in destination d with ancestry from origin o in time t is given by $A_{o,d,t}$ which evolves over time according to $A_{o,d,t+1} = I_{o,d,t} + \sum_{d'=1}^{D} M_{o,d',d,t} + (1-\mu)A_{o,d,t}$. Aggregates $A_{o,t}$ and $A_{d,t}$ are defined analogously to $I_{o,t}$ and $I_{d,t}$ above. Given random assignment of the migration shock we have that total (gross) outmigration of ancestry o from county d' at time t is

$$M_{o,d',\cdot,t} = \sum_{d=1}^{D} M_{o,d',d,t},$$

and given random migration shock assignment we also have that $M_{o,d',\cdot,t} = \mu A_{o,d',t}$.

Immigrant Population Dynamics & Destination Choices. The supply of migrants exogenously grows at rate n but is also subject to some stationary iid shocks, i.e.,

$$I_{o,t} = (1+n)^t e^{\nu_{o,t}},$$

where the shocks ν_t^o are normal with mean 0 and variance σ_ν^2. Individual migrants from origin o within the continuum of migrants $I_{o,t}$ statically optimize over destinations d according to a discrete choice framework, taking into account expectations of conditions in all possible destination counties in the following period. A migrant i's expected utility u from migrating to destination d in period t is

$$u_{o,d,t}(i) = e^{-\tau_{o,d,t}} \varepsilon_{d,t}(i) \mathbb{E}_t W_{d,t+1}^{\lambda} \left(\frac{A_{o,d,t+1}}{A_{o,t+1}} \right)^{1-\lambda}$$

where $\varepsilon_{d,t}(i)$ are iid extreme-value distributed shocks across migrants i with dispersion parameter θ, $\tau_{o,d,t} \sim \mathcal{N}(0, \sigma_\tau^2)$ are iid normal shocks representing bilateral costs, and the relative weight on wages versus ancestry composition satisfies $0 < \lambda < 1$. The expectations $\mathbb{E}_t$ in the immigration payoffs are rational and incorporate fully all information available within the model in period t.

Domestic Migrant Dynamics & Destination Choices. Any individual with ancestry o who is a domestic resident of county d' randomly receiving a migration shock statically optimizes their destination county d according to a discrete choice framework, taking into account expectations of conditions in all possible destination counties in the following period. Such a migrant ξ's expected utility $\tilde{u}$ from migrating to destination d in period t is

$$\tilde{u}_{o,d,t}(\xi) = \tilde{\varepsilon}_{d,t}(\xi) \mathbb{E}_t W_{d,t+1}^{\lambda} \left(\frac{A_{o,d,t+1}}{A_{o,t+1}} \right)^{1-\lambda}$$

where $\tilde{\varepsilon}_{d,t}(\xi)$ are iid extreme-value distributed shocks across domestic migrants ξ with dispersion parameter θ, and the relative weight on wages versus ancestry composition satisfies $0 < \lambda < 1$. The expectations $\mathbb{E}_t$ in the domestic migrant's payoffs are rational and incorporate fully all information available within the model in period t.

Resident Labor Supply. Each resident in the continuum of mass $L_{d,t}$ in destination d and time t supplies one unit of labor inelastically to only its local labor market and can choose whether to allocate this labor to the output sector "Y" or the innovation/new ideas sector "N," resulting in the following identity: $L_{Y,d,t} + L_{N,d,t} = L_{d,t}$.

B.1.2 Equilibrium Definition

An equilibrium in this economy is a sequence of local wages $\{W_{d,t}\}_d$, patent prices $\{p_{d,t}\}_d$, immigration flows $\{I_{o,d,t}\}_{o,d}$, $\{I_{d,t}\}_d$, $\{I_{o,t}\}_o$, domestic migration flows $\{M_{o,d',d,t}\}_{o,d',d}$, ancestry levels $\{A_{o,d,t}\}_{o,d}$, $\{A_{d,t}\}_d$, $\{A_{o,t}\}_o$, labor force levels $\{L_{d,t}\}_d$, labor force allocations $\{L_{N,d,t}, L_{Y,d,t}\}_d$, output levels $\{Y_{d,t}\}_d$, patent flows $\{N_{d,t}\}_d$, local idea levels $\{Q_{d,t}\}_d$, research knowledge levels $\{\bar{Q}_{d,t}\}$, and productivity levels $\{Z_{d,t}\}_d$ such that the following conditions hold.

Final Goods Producers Optimize. Taken as given the numeraire price of the nationally traded output good, local wages $W_{d,t}$, patent prices $p_{d,t}$, local idea levels $Q_{d,t-1}$, and local productivity levels $Z_{d,t}$ as given, the competitive local final goods producer in region d chooses patent demand $N_{d,t}$ and production labor demand $L_{Y,d,t}$ to maximize static profits

$$\max_{N_{d,t},L_{Y,d,t}} Z_{d,t}(N_{d,t} + Q_{d,t-1})L^{\alpha}_{Y,d,t} - W_{d,t}L_{Y,d,t} - p_{d,t}N_{d,t}.$$

This optimization leads to two input optimality conditions listed below.

Research Firms Optimize. Taking as given the price of new varieties or patents $p_{d,t}$ and the wage $W_{d,t}$, the research firm demands research labor $L_{N,d,t}$ to maximize flow profits according to

$$\max_{L_{N,d,t}} p_{d,t}L^{\gamma}_{N,d,t}\bar{Q}^{\zeta}_{d,t-1} - W_{d,t}L_{N,d,t}.$$

Note that the underlying timing here requires that the research firm only be paid for a single period's use of their new ideas or varieties, which are assumed to become freely available to all local firms after one period. This optimization leads to two input optimality conditions listed below, which we emphasize represent a static research choice given our assumption on the timing of expiration of protection of new ideas.

Immigrants Optimize. Taking as given wages $\{W_{d,t}\}_d$ in all regions, as well as ancestry levels $\{A_{o,d,t}\}_{o,d}$ and $\{A_{o,t}\}_o$, an individual immigrant i from origin o in period t optimally chooses their destination d to maximize their static expected utility

$$u_{o,d,t}(i) = e^{-\tau_{o,d,t}}\varepsilon_{d,t}(i)\mathbb{E}_t W^{\lambda}_{d,t+1}\left(\frac{A_{o,d,t+1}}{A_{o,t+1}}\right)^{1-\lambda}.$$

As usual, this structure together with the distributional assumption on $\varepsilon_{d,t}(i)$ leads via a discrete-choice law of large numbers across migrants to immigration shares given by

$$I_{o,d,t} = I_{o,t}\left(\frac{e^{-\theta\tau_{o,d,t}}\left(\mathbb{E}_t W^{\lambda}_{d,t+1}\left(\frac{A_{o,d,t+1}}{A_{o,t+1}}\right)^{(1-\lambda)}\right)^{\theta}}{\sum_{k=1}^{D} e^{-\theta\tau_{o,k,t}}\left(\mathbb{E}_t W^{\lambda}_{k,t+1}\left(\frac{A_{o,k,t+1}}{A_{o,t+1}}\right)^{(1-\lambda)}\right)^{\theta}}\right).$$

Domestic Migrants Optimize. Taking as given wages $\{W_{d,t}\}_d$ in all regions, as well as ancestry levels $\{A_{o,d,t}\}_{o,d}$ and $\{A_{o,t}\}_o$, an individual ξ of ancestry o from resident in county d' in period t and hit by the random migration opportunity shock optimally chooses their destination d to maximize their static expected utility

$$\tilde{u}_{o,d,t}(\xi) = \tilde{\varepsilon}_{d,t}(\xi)\mathbb{E}_t W_{d,t+1}^{\lambda} \left(\frac{A_{o,d,t+1}}{A_{o,t+1}}\right)^{1-\lambda}.$$

As usual, this structure together with the distributional assumption on $\tilde{\varepsilon}_{d,t}(\xi)$ leads via a discrete-choice law of large numbers across domestic migrants to shares given by

$$M_{o,d',d,t} = \mu A_{o,d',t} \left(\frac{\left(\mathbb{E}_t W_{d,t+1}^{\lambda} \left(\frac{A_{o,d,t+1}}{A_{o,t+1}}\right)^{(1-\lambda)}\right)^{\theta}}{\sum_{k=1}^{D} \left(\mathbb{E}_t W_{k,t+1}^{\lambda} \left(\frac{A_{o,k,t+1}}{A_{o,t+1}}\right)^{(1-\lambda)}\right)^{\theta}} \right).$$

Note that the share of domestic migrants of ancestry o from county d' to county d is a function of only the ancestry o and destination county d, allowing us to define domestic share variables

$$s_{o,d,t} = \left(\frac{\left(\mathbb{E}_t W_{d,t+1}^{\lambda} \left(\frac{A_{o,d,t+1}}{A_{o,t+1}}\right)^{(1-\lambda)}\right)^{\theta}}{\sum_{k=1}^{D} \left(\mathbb{E}_t W_{k,t+1}^{\lambda} \left(\frac{A_{o,k,t+1}}{A_{o,t+1}}\right)^{(1-\lambda)}\right)^{\theta}} \right)$$

where $M_{o,d',d,t} = \mu A_{o,d',t} s_{o,d,t}$.

Residents Optimize. Individual residents from the labor force of mass $L_{d,t}$ optimally choose whether to supply labor in the ideas sector or output sector in their local region. This optimization requires, if labor used in both sectors is positive in equilibrium, that the workers be indifferent across sectors and face a common wage $W_{d,t}$ in both final goods and idea production.

Note that the assumptions we have made, which constrain all profit maximization problems by firms and labor decisions by immigrants and residents to be static, do not require us to specify further the nature of household preferences, the details of the nationally traded goods market, nor the intertemporal prices of any assets or savings. To characterize the joint equilibrium dynamics of innovation, immigration, wages, and output, these supplemental details can remain unrestricted.

Labor Markets Clear. The total labor demanded in final goods and ideas production in region d equals the labor force $L_{Y,d,t} + L_{N,d,t} = L_{d,t}$, and the labor force evolves dynamically according to the optimal location decisions of immigrants and domestic inmigrants $L_{d,t+1} = I_{d,t} + (1-\mu)L_{d,t} + M_{.,d,t}$.

Ideas Markets Clear. The patent or variety flows demanded by the final goods firm in region d equal the patent or varieties produced by the research firms in region d at the value $N_{d,t}$.

Productivity Levels Evolve Exogenously. Productivity $Z_{d,t}$ in region d evolves stochastically and exogenously according to

$$\ln Z_{d,t} = \rho \ln Z_{d,t-1} + \epsilon_{d,t}$$

where shocks are iid according to $\epsilon_{d,t} \sim \mathcal{N}(0, \sigma_d^2)$.

B.1.3 Equilibrium Solution

The equilibrium of the economy can be summarized as a system of $2O + D \times (3O + 10)$ nonlinear equations in $2O + D \times (3O + 10)$ endogenous and exogenous variables. These equations are listed below.

- O equations characterizing the immigration push process $I_{o,t}$ from origin o to all destinations at time t
$$I_{o,t} = (1+n)^t e^{\nu_{o,t}}, \quad \nu_{ot} \sim N(0, \sigma_\nu^2)$$

- $O \times D$ equations characterizing $I_{o,d,t}$, the immigration flows from origin o to destination d at time t
$$I_{o,d,t} = I_{o,t} \left(\frac{e^{-\theta \tau_{o,d,t}} \left(\mathbb{E}_t W_{d,t+1}^{\lambda} \left(\frac{A_{o,d,t+1}}{A_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}}{\sum_{k=1}^{D} e^{-\theta \tau_{o,k,t}} \left(\mathbb{E}_t W_{k,t+1}^{\lambda} \left(\frac{A_{o,k,t+1}}{A_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}} \right).$$
$$\tau_{okt} \sim N(0, \sigma_\tau^2)$$
Note that $I_{o,t} = \sum_{d=1}^{D} I_{o,d,t}$ is redundant based on the equations above.

- OD equations characterizing $s_{o,d,t}$, the domestic migration shares of flows of ancestry o potential migrants to destination d at time t
$$s_{o,d,t} = \left(\frac{\left(\mathbb{E}_t W_{d,t+1}^{\lambda} \left(\frac{A_{o,d,t+1}}{A_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}}{\sum_{k=1}^{D} \left(\mathbb{E}_t W_{k,t+1}^{\lambda} \left(\frac{A_{o,k,t+1}}{A_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}} \right).$$
Note that $M_{o,d',d,t} = \mu A_{o,d',t} s_{o,d,t}$ are redundant given $s_{o,d,t}$ and $A_{o,d',t}$.

- $O \times D$ equations linking ancestry shares to immigration and domestic migration flows
$$A_{o,d,t+1} = (1-\mu) A_{o,d,t} + I_{o,d,t} + \sum_{d'=1}^{D} \mu A_{o,d',t} s_{o,d,t}$$

- O equations linking total ancestry stocks to regional ancestry stocks
$$A_{o,t} = \sum_{d=1}^{D} A_{o,d,t}$$

- D equations linking population dynamics to immigration and domestic migration flows
$$L_{d,t+1} = (1-\mu) L_{d,t} + \sum_{o=1}^{O} I_{o,d,t} + \sum_{o=1}^{O} \sum_{d'=1}^{D} \mu A_{o,d',t} s_{o,d,t}$$

- D equations with the final goods production function for $Y_{d,t}$

$$Y_{d,t} = Z_{d,t} Q_{d,t} L_{Y,d,t}^{\alpha}$$

- D equations characterizing exogenous local productivity dynamics

$$\ln Z_{d,t} = \rho \ln Z_{d,t-1} + \epsilon_{d,t}, \quad \epsilon_{d,t} \sim \mathcal{N}(0, \sigma_d^2).$$

- D equations with the idea production function for the mass of new ideas $N_{d,t}$

$$N_{d,t} = L_{N,d,t}^{\gamma} \bar{Q}_{d,t-1}^{\zeta}$$

- D equations characterizing idea dynamics in each region

$$Q_{d,t} = N_{d,t} + Q_{d,t-1}.$$

- D equations summarizing spillovers through the effective ideas available to researchers in d at time t, $\bar{Q}_{d,t}$:

$$\bar{Q}_{d,t} = \prod_{f=1}^{D} Q_{ft}^{\alpha(d,f)}.$$

- D equations linking labor used in production $L_{Y,d,t}$ inversely to the wage

$$\alpha Z_{d,t} Q_{d,t} L_{Y,d,t}^{\alpha-1} = W_{d,t}$$

- D equations linking labor used in research $L_{N,d,t}$ inversely to the wage

$$\gamma p_{d,t} L_{N,d,t}^{\gamma-1} \bar{Q}_{d,t-1}^{\zeta} = W_{d,t}.$$

- D equations linking the price of new ideas positively to the local productivity shock and labor used in production.

$$Z_{d,t} L_{Y,d,t}^{\alpha} = p_{d,t}.$$

- D equations for labor market clearing

$$L_{N,d,t} + L_{Y,d,t} = L_{d,t}$$

Balanced Growth Path Growth Rates. We say that variable X grows at rate g_X if $X_t = (1 + g_X) X_{t-1}$ or equivalently if $X_t \propto (1 + g_X)^t$. We guess and verify that the equilibrium conditions are satisfied with a steady state growth path structure. Assume that all shocks are equal to 0, i.e., $\tau_{o,d,t} = 0$, $\epsilon_{d,t} = 0$, and $\nu_{o,t} = 0$. Note that if g_Q is the growth rate of $Q_{d,t}$ for each region, then $\bar{Q}_{d,t}$ also trivially grows at rate g_Q, since $\sum_f \alpha(d,f) = 1$ for all d. Then, note that the growth rate of new ideas in each region is given by

$$\frac{N_{d,t}}{Q_{d,t-1}} = \frac{L_{N,d,t}^{\gamma} \bar{Q}_{d,t-1}^{\zeta}}{Q_{d,t-1}}$$

$$= \left(\frac{L_{N,d,t}^{\gamma}}{Q_{d,t-1}^{1-\zeta}}\right)\left(\frac{\bar{Q}_{d,t-1}}{Q_{d,t-1}}\right)^{\zeta}$$

On a balanced growth path we have

$$g_Q = \frac{N_{d,t}}{Q_{d,t-1}}$$

$$g_Q = \left(\frac{L_{N,d,t}^{\gamma}}{Q_{d,t-1}^{1-\zeta}}\right)\left(\frac{\bar{Q}_{d,t-1}}{Q_{d,t-1}}\right)^{\zeta}$$

$$g_Q \propto \frac{L_{N,d,t}^{\gamma}}{Q_{d,t-1}^{1-\zeta}}$$

which implies that

$$L_{N,d,t}^{\gamma} \propto Q_{d,t-1}^{1-\zeta} \propto Q_{d,t}^{1-\zeta}.$$

$$Q_{d,t} \propto L_{N,d,t}^{\frac{\gamma}{1-\zeta}}.$$

So we conclude that

$$1 + g_N = 1 + g_{\bar{Q}} = 1 + g_Q = (1 + g_{L_N})^{\frac{\gamma}{1-\zeta}}.$$

Intuitively, this means that in the long run, only increased research labor input drives growth, a direct implication of the fact that this model falls into the class of semi-endogenous growth models from Jones (1995). Recall that the optimality condition for ideas in goods production is

$$p_{d,t} = Z_{d,t} L_{Y,d,t}^{\alpha},$$

so on a steady state growth path we have

$$1 + g_P = (1 + g_{L_Y})^{\alpha}.$$

The optimality condition for labor demand in innovation is

$$W_{d,t} = \gamma p_{d,t} L_{N,d,t}^{\gamma-1} \bar{Q}_{d,t-1}^{\zeta}.$$

But we have that $L_{N,d,t}^{\gamma} \propto \bar{Q}_{d,t-1}^{1-\zeta}$ on a steady state growth path (from the patenting equation arguments above) so that

$$W_{d,t} \propto p_{d,t} \frac{\bar{Q}_{d,t-1}^{1-\zeta} \bar{Q}_{d,t-1}^{\zeta}}{L_{N,d,t}} = p_{d,t} \frac{\bar{Q}_{d,t-1}}{L_{N,d,t}}$$

so that

$$1 + g_W = (1 + g_P)\frac{(1 + g_{\bar{Q}})}{(1 + g_{L_N})} = (1 + g_{L_Y})^{\alpha} \frac{(1 + g_{L_N})^{\frac{\gamma}{1-\zeta}}}{(1 + g_{L_N})} = (1 + g_{L_Y})^{\alpha}(1 + g_{L_N})^{\frac{\gamma}{1-\zeta}-1}.$$

But then from the optimality condition for labor demand in production we also have that

$$W_{d,t} = \alpha Z_{d,t} Q_{d,t} L_{Y,d,t}^{\alpha-1},$$

so that on a steady-state growth path

$$1 + g_W = (1 + g_Q)\,(1 + g_{L_Y})^{\alpha-1}$$

$$1 + g_W = (1 + g_{L_N})^{\frac{\gamma}{1-\zeta}}\,(1 + g_{L_Y})^{\alpha-1}.$$

Equalizing the expressions for $1 + g_W$ from the production and innovation labor demand optimality conditions yields

$$(1 + g_{L_N})^{\frac{\gamma}{1-\zeta}}\,(1 + g_{L_Y})^{\alpha-1} = (1 + g_{L_Y})^{\alpha}(1 + g_{L_N})^{\frac{\gamma}{1-\zeta}-1},$$

$$1 + g_{L_N} = 1 + g_{L_Y}.$$

But then since $L_{d,t} = L_{Y,d,t} + L_{N,d,t}$, we also have $1 + g_L = 1 + g_{L_Y} = 1 + g_{L_N}$. Now, also note that since

$$I_{o,t} = (1 + n)^t e^{\nu_{o,t}},$$

we immediately see that on a growth path $1 + g_{I_o} = 1 + n$. Since the endogenous immigration flows $I_{o,d,t}$ are proportional to $I_{o,t}$ on a growth path, we also have $1 + g_{I_o} = 1 + g_{I_{o,d}}$. Since the ancestry accumulation equations imply (once the stationarity of the domestic migration shares $s_{o,d,t}$ on a steady state growth path is noted) that ancestry is proportional to immigration flows, we also have that $1 + g_{A_{o,d}} = 1 + g_{A_o} = 1 + n$. And therefore, since the labor accumulation equations also imply proportionality between total labor and immigration and ancestry, we have $1 + g_L = 1 + n$. In other words, all labor or population outcomes (total and disaggregated immigration, total and disaggregated domestic migration, total and disaggregated ancestry, total and disaggregated labor forces) all grow at rate $1 + n$ on a steady state growth path, driven by the growth in immigration flows at rate $1 + n$. But then at this point we can write several growth rates from above more explicitly, i.e.,

$$1 + g_Q = 1 + g_{\bar{Q}} = 1 + g_N = (1 + n)^{\frac{\gamma}{1-\zeta}}$$

$$1 + g_p = (1 + n)^{\alpha}$$

$$1 + g_W = (1 + n)^{\frac{\gamma}{1-\zeta}+\alpha-1}.$$

Now from the goods production function we also have

$$Y_{d,t} = Z_{d,t} Q_{d,t} L^{\alpha}_{Y,d,t}$$

implying

$$1 + g_Y = (1 + g_Q)(1 + g_{L_Y})^{\alpha} = (1 + n)^{\frac{\gamma}{1-\zeta}}(1 + n)^{\alpha} = (1 + n)^{\frac{\gamma}{1-\zeta}+\alpha}.$$

Therefore we immediately have that the growth rate of per capita output is

$$1 + g_{Y/L} = (1 + g_Y)/(1 + n) = (1 + n)^{\frac{\gamma}{1-\zeta}+\alpha-1} = 1 + g_W,$$

i.e., wages and per capita output grow at the same rate. Note that for wages and per-capita output to grow at a positive rate we must have the following parametric restriction

$$\frac{\gamma}{1-\zeta} + \alpha - 1 > 0.$$

In the constant returns innovation function case, our baseline model with $\gamma = 1 - \zeta$, we have that this restriction is always satisfied since $\alpha > 0$.[34] But with weaker long-run externalities from idea stocks if $\zeta < 1 - \gamma$ the condition is needed to ensure that ideas have a large enough influence on the marginal product of labor to overcome the long-run neoclassical impact of growing labor supply.

[34] The condition is also satisfied in our robustness checks to multiples cases with $\zeta \neq 1 - \gamma$ in Figure 7 below.

Balanced Growth Path Equilibrium Conditions. Given the derivations above of BGP growth rates, we can scale or detrend the variables and equations above to express them in stationary form away from the BGP. The number of variables is again $2O + D \times (3O + 10)$, denoted with lowercase labels:

1. $i_{o,t} = \frac{I_{o,t}}{(1+n)^t}$, O immigration supply shocks
2. $i_{o,d,t} = \frac{I_{o,d,t}}{(1+n)^t}$, $O \times D$ immigration flows to region d from o at time t
3. $s_{o,d,t}$, $O \times D$ domestic migration flow shares of ancestry o domestic migrants to destination d at time t are already stationary
4. $a_{o,d,t} = \frac{A_{o,d,t}}{(1+n)^t}$, $O \times D$ ancestry stocks from o in region d in time t
5. $a_{o,t} = \frac{A_{o,t}}{(1+n)^t}$, O ancestry stocks from o in total in time t
6. $l_{d,t} = \frac{L_{d,t}}{(1+n)^t}$, D total labor stocks
7. $l_{N,d,t} = \frac{L_{N,d,t}}{(1+n)^t}$, D labor inputs used in innovation
8. $l_{Y,d,t} = \frac{L_{Y,d,t}}{(1+n)^t}$, D labor inputs used in production
9. $y_{d,t} = \frac{Y_{d,t}}{(1+n)^{\left(\frac{\gamma}{1-\zeta}+\alpha\right)t}}$, D outputs
10. D values of $z_{d,t}$, which is already stationary productivity $z_{d,t} = Z_{d,t}$
11. $n_{d,t} = \frac{N_{d,t}}{(1+n)^t}$, D masses of new ideas
12. $q_{d,t} = \frac{Q_{d,t}}{(1+n)^{\left(\frac{\gamma}{1-\zeta}\right)t}}$, D masses of ideas invented locally
13. $\bar{q}_{d,t} = \frac{\bar{Q}_{d,t}}{(1+n)^{\left(\frac{\gamma}{1-\zeta}\right)t}}$, D aggregates of ideas useful for innovation locally
14. $w_{d,t} = \frac{W_{d,t}}{(1+n)^{\left(\frac{\gamma}{1-\zeta}+\alpha-1\right)t}}$, D wages
15. $p_{d,t} = \frac{p_{d,t}}{(1+n)^{\alpha t}}$, D prices of new ideas

These stationary variables are pinned down by the same number of nonlinear equations in stationary form, which are equivalent to the raw equilibrium conditions above but simply rescaled. The equations are:

1. Immigration push shock distributions

$$i_{o,t} = e^{\nu_{o,t}}$$

2. Endogenous immigration flows

$$i_{o,d,t} = i_{o,t} \left(\frac{e^{-\theta\tau_{o,d,t}} \left(\mathbb{E}_t w_{d,t+1}^{\lambda} \left(\frac{a_{o,d,t+1}}{a_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}}{\sum_{k=1}^{D} e^{-\theta\tau_{o,k,t}} \left(\mathbb{E}_t w_{k,t+1}^{\lambda} \left(\frac{a_{o,k,t+1}}{a_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}} \right).$$

3. Endogenous domestic migration shares

$$s_{o,d,t} = \left(\frac{\left(\mathbb{E}_t w_{d,t+1}^{\lambda} \left(\frac{a_{o,d,t+1}}{a_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}}{\sum_{k=1}^{D} \left(\mathbb{E}_t w_{k,t+1}^{\lambda} \left(\frac{a_{o,k,t+1}}{a_{o,t+1}} \right)^{(1-\lambda)} \right)^{\theta}} \right).$$

4. Ancestry accumulation equations

$$a_{o,d,t+1} = \frac{1}{1+n} \left((1-\mu) a_{o,d,t} + i_{o,d,t} + \sum_{d'=1}^{D} s_{o,d,t} \mu a_{o,d',t} \right)$$

$$a_{o,d,t+1} = \frac{1}{1+n} \left((1-\mu) a_{o,d,t} + i_{o,d,t} + s_{o,d,t} \mu \sum_{d'=1}^{D} a_{o,d',t} \right)$$

$$a_{o,d,t+1} = \frac{1}{1+n} \left((1-\mu) a_{o,d,t} + i_{o,d,t} + s_{o,d,t} \mu a_{o,t} \right)$$

5. Ancestry across regions identity

$$a_{o,t} = \sum_{d=1}^{D} a_{o,d,t}$$

6. Labor force accumulation equations

$$l_{d,t+1} = \frac{1}{1+n} \left((1-\mu) l_{d,t} + \sum_{o=1}^{O} i_{o,d,t} + \sum_{d'=1}^{D} \sum_{o=1}^{O} s_{o,d,t} \mu a_{o,d',t} \right)$$

$$= \frac{1}{1+n} \left((1-\mu) l_{d,t} + \sum_{o=1}^{O} i_{o,d,t} + \sum_{o=1}^{O} s_{o,d,t} \mu \sum_{d'=1}^{D} a_{o,d',t} \right)$$

$$= \frac{1}{1+n} \left((1-\mu) l_{d,t} + \sum_{o=1}^{O} i_{o,d,t} + \sum_{o=1}^{O} s_{o,d,t} \mu a_{o,t} \right)$$

7. Labor market clearing equations

$$l_{d,t} = l_{Y,d,t} + l_{N,d,t}$$

8. Output production functions

$$y_{d,t} = z_{d,t} q_{d,t} l_{Y,d,t}^{\alpha}$$

9. Regional productivity shocks stochastic processes

$$\ln z_{d,t} = \rho \ln z_{d,t-1} + \epsilon_{d,t}$$

10. Idea production functions

$$n_{d,t} = l_{N,d,t}{}^{\gamma} \bar{q}_{d,t-1}^{\zeta} (1+n)^{\frac{\gamma\zeta}{\zeta-1}}$$

11. Idea accumulation equations

$$q_{d,t} = n_{d,t} + \left(\frac{1}{1+n}\right) q_{d,t-1}$$

12. Regional research knowledge aggregators

$$\bar{q}_{d,t} = \prod_{f=1}^{D} q_{ft}^{\alpha(d,f)}$$

13. Labor demand optimality for final goods producers

$$w_{d,t} = \alpha z_{d,t} q_{d,t} l_{Y,d,t}^{\alpha-1}$$

14. Labor demand optimality from research firms

$$w_{d,t} = \gamma p_{d,t} l_{n,d,t}^{\gamma-1} \bar{q}_{d,t-1}^{\zeta} (1+n)^{\frac{\gamma\zeta}{\zeta-1}}$$

15. Idea demand optimality from final goods producers

$$p_{d,t} = z_{d,t} l_{Y,d,t}^{\alpha}$$

B.1.4 Numerical Solution and Simulation

We solve the stationary system of equations from section B.1.3 above using second-order perturbation around the nonstochastic BGP of the economy. This nonlinear solution approach is crucial for accounting for the nonlinear mapping from shocks to immigration supply ν_{ot} to immigration flows at the regional level which is state-dependent, varying with predetermined ancestry levels and current wages.

To simulate the model, we draw immigration supply shocks ν_{ot}, regional productivity shocks $\epsilon_{d,t}$, and bilateral immigration cost shocks $\tau_{o,d,t}$ for a large number of periods $T = 1000$, $O = 10$ origins, and $D = 9$ destination regions. Given a parametrization of the model, the exogenous shock draws together with the nonlinear policy functions obtained in our solution step allow for unconditional simulation of the model. This unconditionally simulated data can be processed to produce a range of moments for structural estimation of the model, which is detailed below. Given that we only compute local and symmetric national responses, and given that model nonlinearities relate primarily to asymmetric histories across locations, we compute impulse repsonses using a linearized version of the model.

We implement all of these numerical model steps, i.e., solution, unconditional simulation, and impulse response calculations, using Dynare within a MATLAB environment. Given the smooth nature of our equilibrium conditions, the well behaved non-stochastic BGP, and the large number of equilibrium conditions, the Dynare package is a natural choice for numerical analysis in this context.

B.2 Structural Estimation

To parameterize our model, we first externally calibrate or fix the magnitudes of various parameters to values commonly employed in the literature. To match our empirical approach, we solve and simulate the model in five-year periods, and we choose the value n of exogenous population growth n, equal to the BGP growth rate of knowledge or ideas in our economy, to be 2% on an annualized basis. Note that the value α has a literal labor share in output interpretation, but α also has an interpretation in many tightly related growth models as the parameter governing markups for intermediate goods firms. So we choose α to imply a round value of a 20% markup, i.e., $\alpha = 1/1.2 \approx 0.8$ in our baseline. Based on the analysis in Caliendo et al. (2019), we match an elasticity of immigration shares to local wages of 0.5 through the choices $\lambda = 0.5$ and $\theta = 1$. We also choose the domestic migration shock probability μ to guarantee a steady-state mobility rate of around 5.6% on an annualized basis from CPS data in the 1980-2010 period (US Census Bureau, 2022).

Estimated Parameters. After external calibration of the parameters noted above, there are seven remaining parameters in our model:

1. Elasticity of local innovation to researchers γ
2. Autocorrelation of regional productivity shocks ρ
3. Volatility of regional productivity shocks σ_ϵ
4. Volatility of immigration supply shocks σ_ν
5. Volatility of bilateral immigration cost shocks σ_τ
6. Linear decline in research knowledge spillovers with distance δ
7. Elasticity of local innovation to idea stocks ζ

We structurally estimate the values of the first five parameters above using an overidentified simulated method of moments (SMM) procedure outlined below.[35] For the sixth parameter, related to idea spillovers, we explore the implications of varying the parameter to extreme values implying full, frictionless spillovers of ideas across regions ($\delta = 0$) versus no idea spillovers across idea-autarkic regions ($\delta = \infty$, our baseline described in the main text). For each of these alternative cases for idea spillovers, we implement the full SMM estimation procedure below conditional upon the appropriate value of δ. Note also that the values we choose for δ, 0 vs ∞, imply spillovers that are either non-existent or independent of distance, implying that we do not need to explicitly specify the geographic structure of the model. Finally, for the seventh parameter ζ, we make the baseline assumption $\zeta = 1 - \gamma$ of constant returns in innovation, relaxing this assumption in robustness checks in Figure 7 below.

[35] Note that one of our target moments is an IV regression coefficient, leading us to sometimes interchangeably refer to this approach as an indirect inference procedure in the main text.

Target Moments. To discipline the values of the five estimated parameters, we target the value of six related moments:

1. IV coefficient estimating the elasticity of patenting to immigration at the county d level
2. Standard deviation of origin-level immigration flows at the origin o level
3. Standard deviation of destination d-level immigration flows
4. Standard deviation of origin $o \times$ destination d-level immigration flows
5. Autocorrelation of output per capita at the county d level
6. Autocorrelation of patenting at the county d level

Moments 1-4 and 6 are directly computable within the 1975-2010 sample used for the main reduced-form empirical results in the paper. We compute the fifth moment based on the BEA (2021)'s county level GDP estimates for five-year periods within the available 2001-19 window.

Although the mapping from parameters to moments in our model is nonlinear and joint in nature, there are certain parameters particularly influential for determining the value of individual moments in our simulation. In particular, the IV-estimated elasticity of patenting to immigration moves directly in the model with the underlying local elasticity of innovation to researchers γ. The volatilities of origin-, destination-, and origin $\times$ destination-level immigration flows depend upon the volatilities of origin-, destination-, and origin $\times$ destination-level exogenous shocks σ_ν, σ_ϵ, and σ_τ, respectively. The autocorrelations of per-capita output and patenting increase with the autocorrelation of underlying regional productivity shocks ρ.

SMM Objective and Standard Errors. First, we collect the five estimated parameters into the vector $\theta = (\gamma, \rho, \sigma_\epsilon, \sigma_\nu, \sigma_\tau)'$. We similarly collect the values of the six target moments m into vectors, denoting by $m(X)$ the value of these moments in the empirical data X, denoting by $m^S(\theta)$ the values of these moments based on our unconditionally simulated data in the model, and denoting by $m(\theta)$ the population values of these moments. Our SMM estimation procedure generates point estimates $\hat{\theta}$ as the solution to the minimization problem

$$\min_{\theta}(m(X) - m^S(\theta))'W(m(X) - m^S(\theta)),$$

where W is a symmetric weighting matrix for the simulated moment deviations. If the moment vector behaves in an asymptotically normal fashion according to

$$\sqrt{N}\,(m(X) - m(\theta)) \to_d \mathcal{N}(0, V),$$

then standard SMM derivations yield asymptotic normality for the parameter estimates

$$\sqrt{N}\left(\hat{\theta} - \theta\right) \to_d \mathcal{N}(0, \Sigma),$$

where the asymptotic variance Σ is given by the sandwich formula

$$\Sigma = \left(1 + \frac{1}{S}\right)\left(\frac{\partial m'}{\partial \theta} W \frac{\partial m}{\partial \theta}\right)^{-1} \frac{\partial m'}{\partial \theta} WVW \frac{\partial m}{\partial \theta} \left(\frac{\partial m'}{\partial \theta} W \frac{\partial m}{\partial \theta}\right)^{-1}$$

Above, $\frac{\partial m}{\partial \theta}$ is the Jacobian of model moments to parameters and S is the ratio of the simulated to empirical sample sizes.

Some practical decisions must be made to compute point estimates $\hat{\theta}$ as well as a feasible estimate $\hat{\Sigma}$ of the asymptotic variance above. First, we use the identity weighting matrix $W = I$. We compute the Jacobian numerically using finite differences relative to our point estimates. To compute an estimate of the moment covariance matrix $\hat{V}$, we first impose diagonality across the moments which all differ by aggregation level and sample size. We then compute asymptotic variances for each moment using a combination of standard analytic formulas, clustering by state, and the Delta method. The resulting standard errors reported are given by $\left(\frac{\text{diag}\hat{\Sigma}}{N}\right)^{0.5}$. The main text's Table 6 reports parameter estimates, standard errors, and model vs data moments for our baseline case with no idea spillovers, and Table 19 reports the same information for the alternative case with full idea spillovers.

B.3 Immigration and Naturalization Act Accounting Exercise

In order to model a scenario that mimics a hypothetical failure of the Immigration and Naturalization Act (INA) to pass in 1965, we compute the counterfactual evolution over time of macroeconomic aggregates in a version of our model in which we feed a string of negative exogenous shocks to immigration supply, symmetric across origins, which reduce the US population growth rate by 16% relative to our baseline calibrated model.

To compute this 16% value, we proceed as follows. We first extract overall population counts and counts of the population of the foreign born from decadal US Census tabulations in the 1860-2010 time period (US Census Bureau, 2014). The total population in Census year t, P_t, is made up of native, N_t, and foreign-born individuals, F_t,

$$P_t = N_t + F_t.$$

We can then decompose the growth rate of the US population as a whole into a fraction accounted for by natives ($\Delta N_t / \Delta P_t$) and the remaining fraction accounted for by the foreign born ($\Delta F_t / \Delta P_t$),

$$\frac{\Delta P_t}{P_t} = \frac{\Delta P_t}{P_t} \left(\frac{\Delta N_t}{\Delta P_t} + \frac{\Delta F_t}{\Delta P_t} \right).$$

The share of the US population growth rate accounted for by natives fell from 95% in the decades before the INA (1860-1960) to 80% in the decades after (1970-2010).

We then assume that the *only* exogenous change in a world with the INA compared to a world without is the process for immigration. The growth of the native population, as a share of the US population, remains constant, ${}^{\Delta N_t}/_{P_t}|_{no\,INA} = {}^{\Delta N_t}/_{P_t}|_{INA}$. To match the decrease of the share of the population growth rate coming from natives (${}^{\Delta N_t}/_{\Delta P_t}$) from 95% in the pre-INA period to 80% in the post-INA period, we calibrate a world without the INA by feeding a string of negative exogenous shocks to immigration supply, symmetric across regions, such that the US population growth rate declines by 16%.[36]

[36] The population growth rate is x% lower in a world without versus with the INA, $g|_{no\,INA} = (1-x) g|_{INA}$. Imposing that the contribution of natives to the population remains constant, $g|_{no\,INA} {}^{\Delta N_t}/_{\Delta P_t}|_{no\,INA} = g|_{INA} {}^{\Delta N_t}/_{\Delta P_t}|_{INA}$, and plugging in the empirical contribution of natives to population growth pre- and post-INA (95% and 80% respectively), we get $x = 1 - {}^{0.8}/_{0.95} \approx 0.16$.

We can compute the total difference in immigration over 45 years (corresponding to 1965-2010) in our two scenarios, with and without the INA. Given that the reduction in the annual population growth rate from 1.03% to 0.87% is solely due to a reduction in immigration, there would have been 21 million fewer migrants.[37]

Appendix Table 19: Parameters and Model Fit, Full Idea Spillovers

Panel A: Moments	Data	Model
IV coeff., patenting$_{d,t}$ on immigration $I_{d,t}$	1.6519	1.6418
	(0.1500)	
Std. deviation, o immigration $I_{o,t}$	0.4061	0.3931
	(0.0284)	
Std. deviation d immigration $I_{d,t}$	0.1794	0.1815
	(0.0110)	
Std. deviation, o-d immigration $I_{o,d,t}$	0.0716	0.1188
	(0.0117)	
Autocorrelation, output per capita $Y_{d,t}/L_{d,t}$	0.9611	0.9518
	(0.0057)	
Autocorrelation, patenting$_{d,t}$	0.9309	0.8745
	(0.0065)	
Panel B: Estimated Parameters	**Symbol**	**Value**
Elasticity, patenting to labor	γ	0.7674
		(0.1392)
Autocorrelation, county TFP	ρ	0.8681
		(0.0366)
Std. deviation, county TFP shocks	σ_ϵ	0.0239
		(0.0130)
Std. deviation, immigration push shocks	σ_ν	0.5864
		(0.0797)
Std. deviation, bilateral immigration shocks	σ_τ	0.5780
		(0.0714)

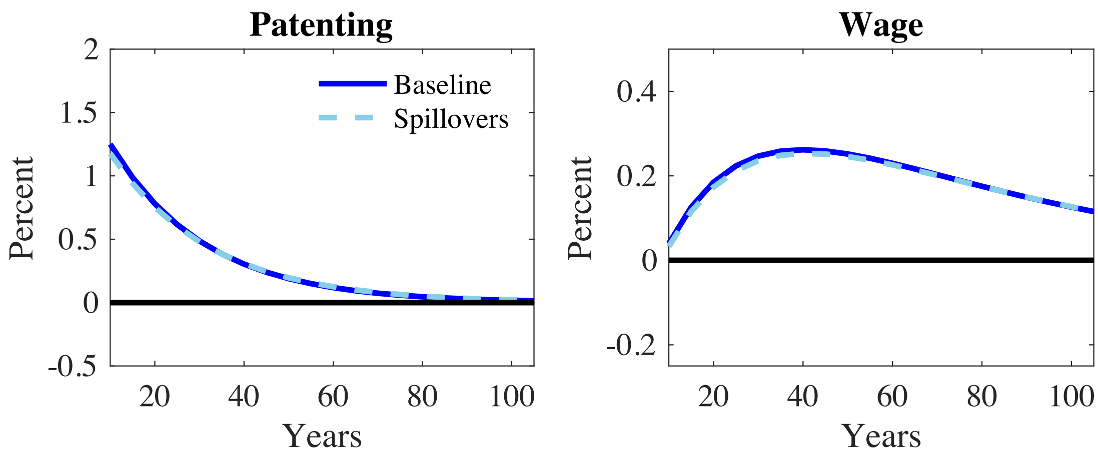
Notes: The table reports the model fit and estimated parameters in the alternative model with full idea spillovers across counties. The top Panel A reports targeted data moments vs simulated model moments. The bottom Panel B reports the estimated parameters. The standard errors, in parentheses beneath moments and estimates, are clustered by state.

[37] The realized US population growth is 1.03% per year, from 195 to 309 million. In a counterfactual scenario without the INA, with an annual population growth rate falling by 16% to 0.87% per year, the total population would grow from 195 to 289 million, 21 million fewer than in the scenario with the INA, entirely attributable to missing migrants.

Appendix Table 20: Target IV Regression in Simulated Model Data

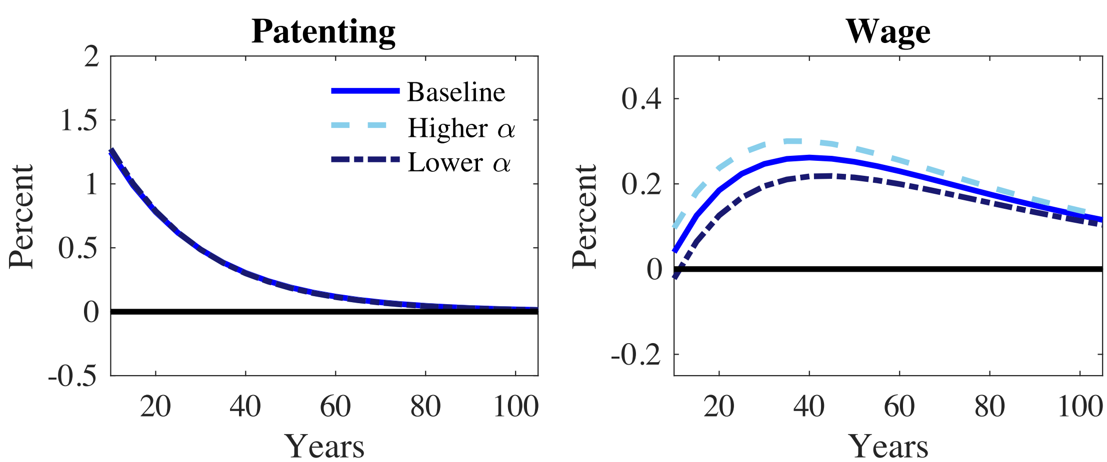
	IHS(Patenting)
IHS(Immigration)	1.641***
	(0.176)
N	8991
First Stage F-Stat	51
AR Wald F-Test p-value	0.000
Geography FE	Yes
Time FE	Yes

Notes: The table reports the coefficient β_{IHS} from (7) estimated using $IHS(\hat{I}_{d,t})$ as an instrumental variable in simulated model data. The simulated panel dataset at the county by time level is constructed using our baseline parameter estimates from Table 6. We unconditionally simulate D = 9 destination counties d with O=10 origin countries o for total of $T = 1000$ time periods t, with timing conventions yielding a total of D $\times$ T - D = 8991 observations in our simulated dataset. Standard errors are clustered by destination. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.



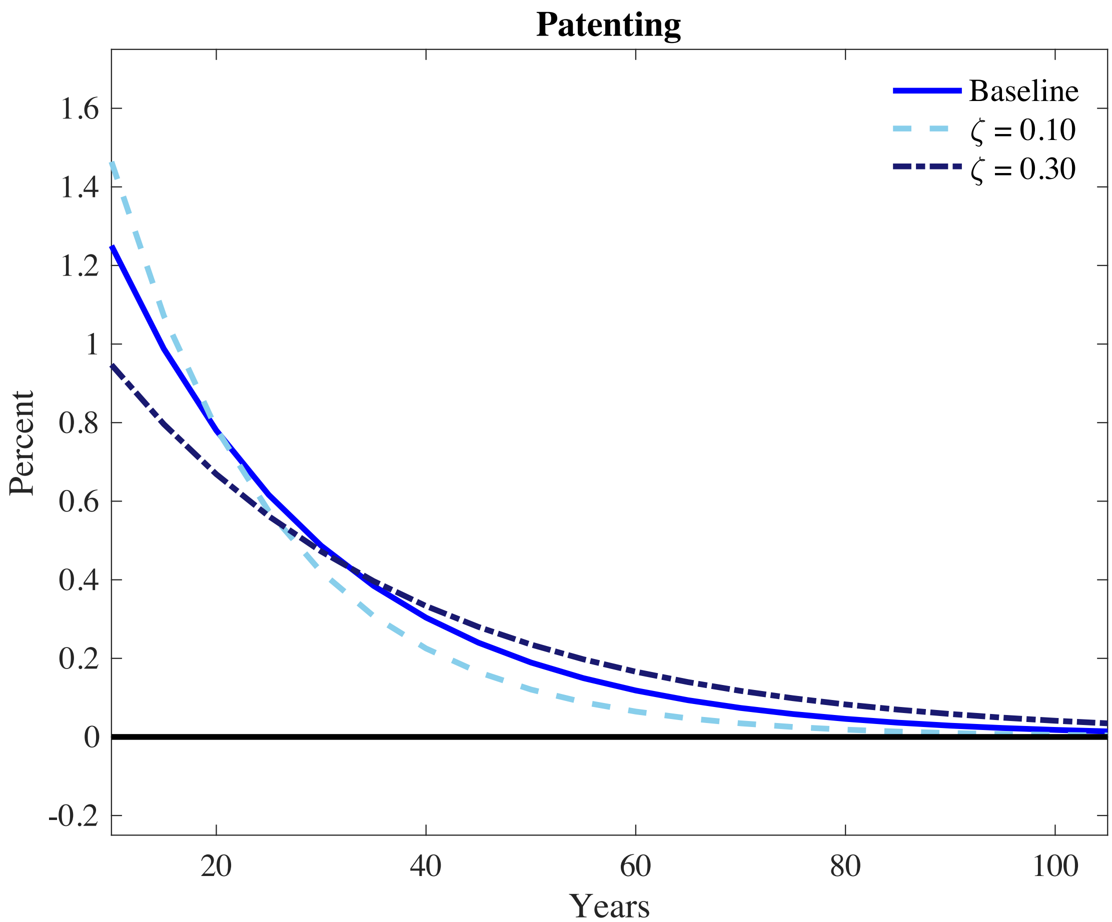
Appendix Figure 5: Full Idea Spillovers in the Model

Notes: The figure plots impulse response functions to a one-standard deviation immigration shock in period 1. The left panel plots patenting $n_{d,t}$. The right panel plots the response of the wage $w_{d,t}$. The immigration shock is from a single origin o, and the responses of the labor force, patenting, and the wage are local responses for a county d. The solid blue line labelled Baseline traces the impact of the immigration shock in our baseline estimated model with no cross-region spillovers. The dashed light blue line labelled Spillovers reports the impact of an immigration shock in our alternative estimated model allowing for full idea spillovers. The responses are in percentage deviations from the balanced growth path.



Appendix Figure 6: Alternative Returns to Scale in the Model

Notes: The figure plots impulse response functions to a one-standard deviation immigration shock in period 1. The left panel plots patenting $n_{d,t}$. The right panel plots the response of the wage $w_{d,t}$. The immigration shock is from a single origin o, and the responses of the labor force, patenting, and the wage are local responses for a county d. The solid blue line labelled Baseline traces the impact of the immigration shock in our baseline estimated model with $\alpha \approx 0.8$. The dashed light blue line labelled Higher α reports the case of $\alpha = 0.95$, while the dashed dot darker blue line labelled Lower α uses $\alpha = 0.7$. The responses are in percentage deviations from the balanced growth path.



Appendix Figure 7: Alternative Returns to Scale in Innovation

Notes: The figure plots impulse response functions of patenting $n_{d,t}$ to a one-standard deviation immigration shock in period 1. The immigration shock is from a single origin o, and the responses of the labor force, patenting, and the wage are local responses for a county d. The solid blue line labelled Baseline traces the impact of the immigration shock in our baseline estimated model with an elasticity of innovation past ideas of $1-\hat{\gamma} \approx 0.2$. The dashed light blue line instead considers an elasticity of innovation to past ideas of $\zeta = 0.1$ in an extended, re-estimated semiendogenous growth model with innovation function $N_{d,t} = L_{N,d,t}^{\gamma} Q_{d,t-1}^{\zeta}$. The dashed dot darker blue line considers an elasticity of innovation to past ideas of $\zeta = 0.30$ in the extended model, again re-estimated. The responses are in percentage deviations from the balanced growth path.

References

Adão, R., M. Kolesár, and E. Morales (2019): "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 134, 1949–2010.

Angrist, J. D. and J.-S. Pischke (2009): *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

BEA (2018): "Personal Consumption Expenditures Price Index," data, US Bureau of Economic Analysis.

——— (2021): "GDP by County, Metro, and Other Areas," data, US Bureau of Economic Analysis.

BLS (2018): "Quarterly Census of Employment and Wages," data, US Bureau of Labor Statistics.

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

Caliendo, L., M. Dvorkin, and F. Parro (2019): "Trade and labor market dynamics: General equilibrium analysis of the china trade shock," *Econometrica*, 87, 741–835.

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

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.

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

Manson, S., J. Schroeder, D. V. Riper, K. Knowles, T. Kugler, F. Roberts, and S. Ruggles (2023): "IPUMS National Historical Geographic Information System: Version 18.0 [dataset]," data, Integrated Public Use Microdata (IPUMS).

NBER (2018): "US Intercensal Population by County and State 1970-on," data, National Bureau of Economic Research.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2018-2020): "Integrated Public Use Microdata Series: Version 8.0, 9.0, 10.0 [dataset]," data, University of Minnesota.

US Census Bureau (2013): "Census Regions and Divisions of the United States," data, US Census Bureau.

——— (2014): "The "Second Great Wave" of Immigration: Growth of the Foreign-Born Population Since 1970," Web posting.

——— (2018a): "Business Dynamics Statistics," data, US Census Bureau.

——— (2018b): "Longitudinal Business Database," data, US Census Bureau.

——— (2022): "CPS Historical Migration/Geographic Mobility Tables," data, US Census Bureau.

USPTO (Accessed: Mar. 28-29, 2022): "Data Download Tables," data, U.S. Patent and Trademark Office.