Talents and Cultures: Immigrant Inventors and Ethnic Diversity in the Age of Mass Migration

Francesco Campo, U Milan Bicocca

Mariapia Mendola, U Milan Bicocca and IZA

Andrea Morrison, U Pavia, Utrecht and ICRIOS-Bocconi

Gianmarco Ottaviano, U Bocconi, Baffi-CAREFIN, CEP and CEPR

Immigrant Inventors

*“**The global race for talent is on**, with countries and businesses competing for the best and brightest. Talented individuals migrate much more frequently than the general population, and the **United States has received exceptional inflows of human capital. This foreign talent has transformed U.S. science and engineering, reshaped the economy,** and influenced society at large” (W. R. Kerr, 2021)*

Immigrant Inventors (Cont.)

- **Charles Steinmetz**, born in Breslau (Germany), studied at Zurich Polytechnic. He migrated in 1889 to US fleeing persecution in Germany due to its socialist ideas. He soon became chief consulting engineer at GE
- In the words of the historian T. Hughes, he "***introduced American engineers to advance mathematical modes*** *of analyzing alternative current light and power systems. These modes* ***greatly enhanced the problem solving abilities*** *of engineering colleagues at GE*" (Hughes 2004; 161)

Immigrant Inventors (Cont.)

- Age of Mass Migration: Between 1840 and 1930 about 30 millions Europeans migrated to the US
- 40% of today US population descends from those immigrants
- Different waves:
  - 1st 1830s/40s: northern Europeans (e.g. Ireland, Germany and England)
  - 2nd 1850s/80s: German and Scandinavian
  - 3rd >1880 till 1924: South and Eastern European (e.g. Italian, Russian), 40% of total foreign born
- Foreign born population from 5% to 14% (1840-1924)
- Starting from 1922, quota restrictions are introduced and strengthened overtime
- Quotas affected mainly recent immigrants groups (Southern and Eastern Europeans)

This Paper

- Leverages data from the Age of Mass Migration to address the following research questions:
  - Where do immigrant inventors go?
  - Do they follow their co-ethnic network?
  - Does ‘cultural diversity’ attract them?
  - If so, is it because of productive or consumption amenities?

- Exploits a unique historical dataset of immigrant inventors over a century of US history (1840-1940)
- Characterizes the geographical patterns of immigrant inventors’ location and their knowledge creation (patenting)
- Identifies the drivers of immigrant inventors’ location choices

This Paper (Cont.)

- To guide the empirical analysis develops a simple model of inventors' location choices
- Identifies the effects of productive vs consumption amenities linked to ethnic networks and cultural diversity
- Isolates a causal impact by exploiting exogenous variation in diversity by:
  - Adopting a classic shift-share methodology
  - Exploiting the policy change introduced by the US Immigration Acts during the 1920s and 1930s

This Paper (Cont.)

- Finds that:
  - Cultural diversity is a significant *pull factor* for immigrant inventors, over and above co-ethnic network and immigration size effects
  - The dominant driver is productivity rather than consumption amenities
  - We rule out alternative mechanisms such as inter-group connections (proxied by inter-ethnic marriages and residential contact), cultural proximity and natives' attitudes (proxied by migrant ethnic groups' salience in newspapers).

Related Literature

- Benefits from immigration from complementarities with natives (Peri & Sparber, 2009; Ottaviano & Peri, 2012)
- Due to different skills, backgrounds, values, norms, ideas (Giuliano 2007; Algan & Cahuc, 2010)
- Both share and composition of foreign labor force are important (Ottaviano & Peri, 2005, 2006; Ager & Bruckner, 2013; Alesina & Rapoport, 2016; Docquier et al., 2018)
- Role of inventors and scientists arrived during 20th century for subsequent growth of US economy (Moser et al., 2014; Ganguli, 2015; Akcigit et al, 2017)
- Importance of top inventors (Azoulay et al, 2010; Moretti, 2019)

Related Literature (Cont.)

- Growing body of work on the Age of Mass Migration (Abramitzky & Boustan, 2017; Ager & Hansen, 2017; Hatton & Ward, 2018; Sequeira et al, 2018; Tabellini, 2018):
  - Waves of immigration vary across countries of origin and over time
  - Large variation in both size and composition (diversity) of immigrants
  - Introduction of yearly quotas in 1922 ended an era of (almost) unconstrained immigration
  - Main focus has been on low-skilled immigrants, less evidence on the arrivals of top-end talents and inventors (upper-tale of skill distribution)
- **Our contribution**: relations among local diversity, talent attraction and innovation

Patent Data: US Patent and Trademark Office

- Original dataset on immigrant inventors from historical records of the US Patent and Trademark Office (Diodato, Morrison and Petralia, 2022)
- Text-mining, semi-automated procedures to extract detailed information on immigrant inventors' country of origin and US county of residence between 1870 and 1940 (Petralia et al, 2016)
- Dataset identifies:
  - Patent documents belonging to immigrants
  - Immigrant patent holders' nationality and county of residence
- Given the fast naturalization process in the Age of Mass Migration, these immigrant patent holders are recently-arrived (adult) immigrants at the top end of the skill distribution

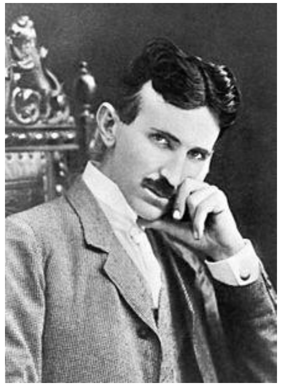
County Data: US Census

- Federal Census data from 1870 to 1930 to derive the following county-level variables:
  - Total immigrants by country of origin (birthplace)
  - Total resident population

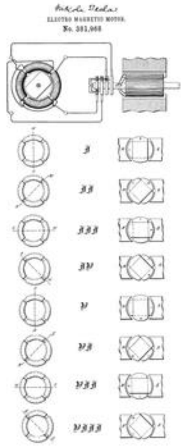
- Combining USPTO and Census data gives a county-by-ethnicity panel over 7 decades with two cross-sectional dimensions:
  - Geographic variation across US counties (about 2700 US 1990-boundaries counties)
  - Immigrants' birthplace variation (15 countries/areas of birth identified by both datasets)

Patent Data

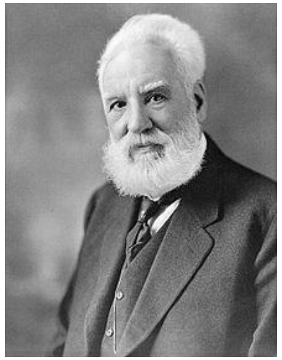
1890 U.S. Patent 381,968, alternating induction motor



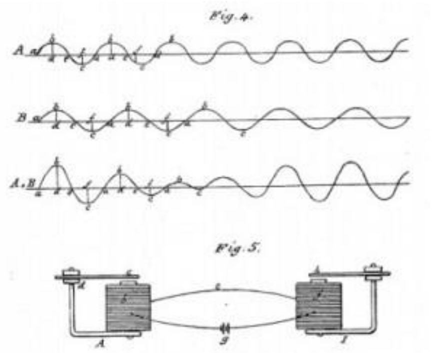
Nikola Tesla, Serbian



1876 U.S. Patent No. 174465, Improvement in Telegraphy

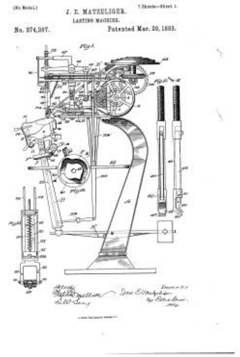


Alexander Graham Bell, Scottish



1883 US patent No. 274207 ‘lasting’ machine

Jan Ernst Matzelige, Dutch



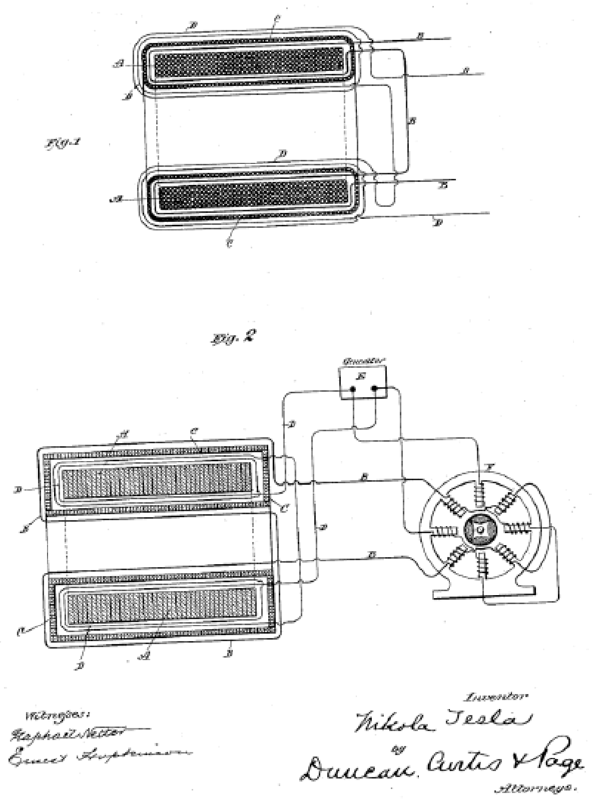
Patent Data (Cont.)

(No Model.)

N. TESLA.

ELECTRICAL TRANSFORMER OR INDUCTION DEVICE.

No. 433,702. Patented Aug. 5, 1890.



UNITED STATES PATENT OFFICE.

NIKOLA TESLA, OF NEW YORK, N. Y., ASSIGNOR TO THE TESLA ELECTRIC COMPANY, OF SAME PLACE.

ELECTRICAL TRANSFORMER OR INDUCTION DEVICE.

SPECIFICATION forming part of Letters Patent No. 433,702, dated August 5, 1890.

Application filed March 26, 1890. Serial No. 345,390. (No model.)

To all whom it may concern:

Be it known that I, NIKOLA TESLA, a subject of the Emperor of Austria-Hungary, from Smiljan, Lika, border country of Austria-Hungary, residing at New York, in the county and State of New York, have invented certain new and useful Improvements in Electrical Transformers or Induction Devices, of which the following is a specification, reference being had to the drawings accompanying and forming a part of the same.

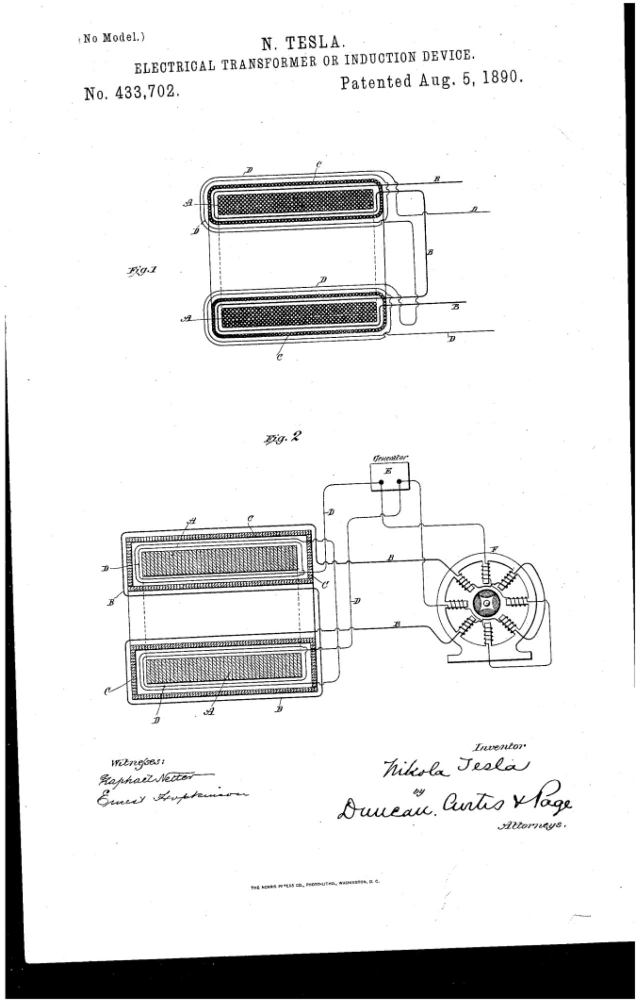
This invention is an improvement in electrical transformers or converters, and has for its main objects the provision of means for securing, first, a phase difference between the primary and secondary currents adapted to the operation of my alternating-current motors and other like purposes, and, second, a constant current for all loads imposed upon the secondary.

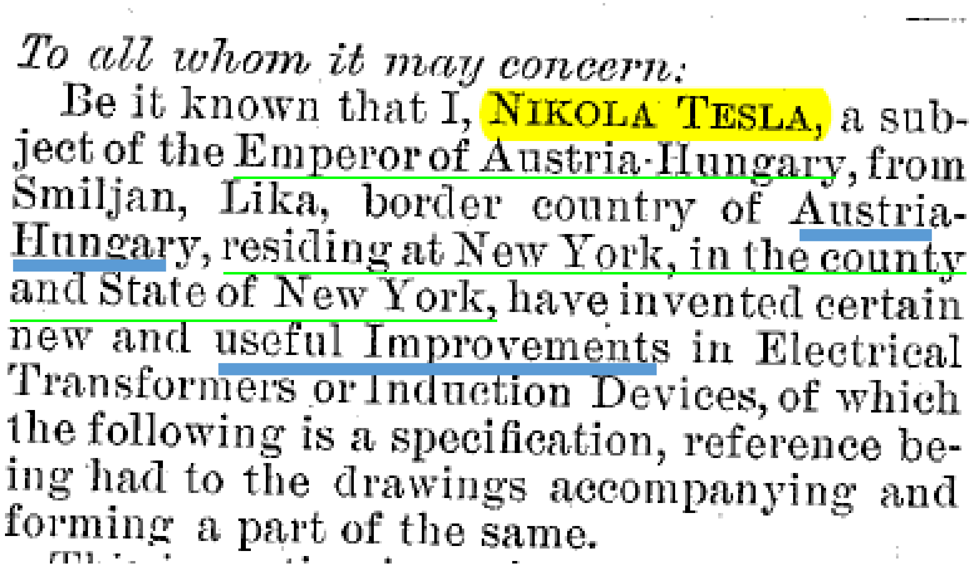
In transformers as constructed now and heretofore it will be found that the electro-motive force of the secondary very nearly coincides with that of the primary, being, however, of opposite sign. At the same time the currents, both primary and secondary, lag behind their respective electro-motive forces; but as this lag is practically or nearly the same in the case of each it follows that the maximum and minimum of the primary and secondary currents will nearly coincide, but differ in sign or direction, provided the secondary be not loaded or if it contain devices having the property of self-induction. On the other hand, the lag of the primary behind the impressed electro-motive force may be diminished by loading the secondary with a non-inductive or dead resistance—such as incandescent lamps—whereby the time interval between the maximum or the minimum periods of the primary and secondary currents is increased. This time interval, however, is limited, and the results obtained by phase difference in the operation of such devices as my alternating-current motors can only be approximately realized by such means of producing or securing this difference, as above indicated, for it is desirable in such cases that there should exist between the primary and secondary currents, or those which, however produced, pass through the two circuits of the motor, a difference of phase of ninety degrees; or, in other words, the current in one circuit should be maximum when that in the other circuit is minimum. To more perfectly attain to this condition I obtain or secure an increased retardation of the secondary current in the following manner: Instead of bringing the primary and secondary coils or circuits of a transformer into the closest possible relations, as has hitherto been done, I protect in a measure the secondary from the inductive action or effect of the primary by surrounding either the primary or the secondary with a comparatively-thin magnetic shield or screen. Under these conditions or circumstances, as long as the primary current has a small value, the shield protects the secondary; but as soon as the primary current has reached a certain strength, which is arbitrarily determined, the protecting magnetic shield becomes saturated and the inductive action upon the secondary begins. It results, therefore, that the secondary current begins to flow at a certain fraction of a period later than it would without the interposed shield, and since this retardation may be obtained without necessarily retarding the primary current also, an additional lag is secured, and the time interval between the maximum or minimum periods of the primary and secondary currents is increased. I have further discovered that such a transformer may, by properly proportioning its several elements and determining in a manner well understood the proper relations between the primary and secondary windings, the thickness of the magnetic shield, and other conditions, be constructed to yield a constant current at all loads. No precise rules can be given for the specific construction and proportions for securing the best results, as this is a matter determined by experiment and calculation in particular cases; but the general plan of construction which I have described will be found under all conditions to conduce to the attainment of this result.

In the accompanying drawings I have illustrated the construction above set forth.

Figure 1 is a cross-section of a transformer embodying my improvement. Fig. 2 is a simi-

Patent Data (Cont.)





To all whom it may concern:

Be it known that I, NIKOLA TESLA, a subject of the Emperor of Austria-Hungary, from Smiljan, Lika, border country of Austria-Hungary, residing at New York, in the county and State of New York, have invented certain new and useful Improvements in Electrical Transformers or Induction Devices, of which the following is a specification, reference being had to the drawings accompanying and forming a part of the same.

We identify:

- Patent files belonging to immigrant inventor
- Immigrant inventor's nationality
- Immigrants inventor's residency (county/state)

Descriptive Stats: USPTO

Table 1: Patents and number of migrant inventors in US by nationality. 1880-1930

	1880-90		1890-00		1900-1910		1910-1920		1920-30		1930-1940		1880-1940	
Nationality	Pat.	Inv.	Pat.	Inv.	Pat.	Inv.	Pat.	Inv.	Pat.	Inv.	Pat.	Inv.	Pat.	Inv.
Asia	0	0	7	5	59	39	285	185	245	144	21	14	621	390
Australia and New Zealand	0	0	1	1	6	4	9	8	18	11	16	3	52	28
Austro-Hungarian Emp.	25	3	91	41	396	257	1,363	896	1,017	532	285	99	3,240	1,855
Benelux	8	5	19	9	133	71	184	98	86	47	29	6	461	238
Canada	27	20	108	54	405	216	541	256	572	242	229	76	1,912	877
Eastern Europe	16	8	62	45	393	268	1,377	811	1,528	898	502	143	3,996	2,213
France	26	11	56	29	278	130	281	143	257	118	85	22	994	459
Germany	124	60	305	171	1,325	699	2,065	927	1,014	431	316	108	5,203	2,420
Great Britain and Ireland	876	313	1,422	699	3,537	1,721	4,431	2,019	3,795	1,345	1,871	416	16,271	6,656
Greece	0	0	3	2	25	14	77	59	118	94	15	9	240	179
Italy	9	6	51	25	289	195	743	510	751	428	312	66	2,195	1,244
Portugal	0	0	0	0	3	3	13	9	26	22	1	1	43	35
Rest Of America	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Scandinavia	65	46	340	203	1,601	741	2,311	1,140	1,479	678	700	180	6,623	3,046
Spain	5	5	9	5	39	19	54	35	86	48	5	5	198	117
Switzerland	47	17	45	26	277	142	385	183	286	128	205	40	1,318	546
Total	0	0	0	0	0	0	0	0	0	0	0	0	43,367	20,303

[1] Data source: Diodato et al. (2021). Each row displays the number of patents and inventors by nationality and decade from 1880 until 1940. Last two columns report the same information for the whole period under consideration. Bottom row aggregates data across all ethnicities.

Descriptive Stats: Census

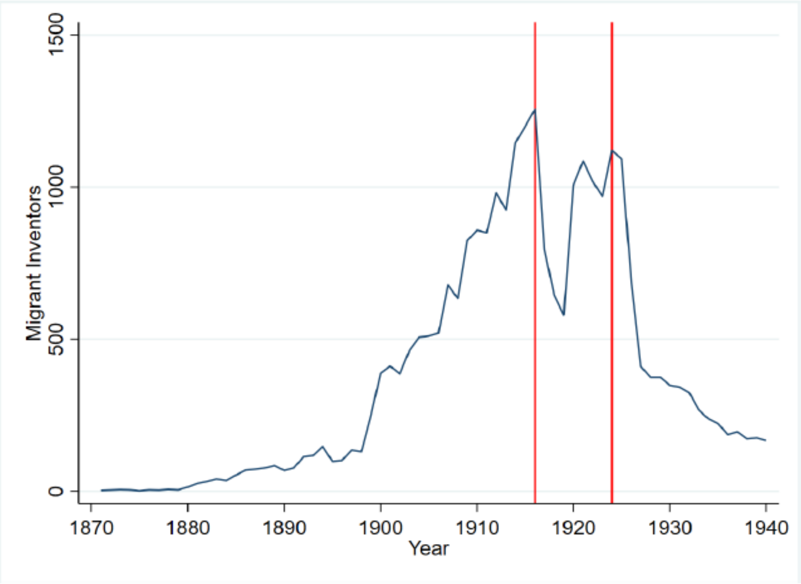
Table 2: Immigration shares (%) and within–diversity in US Census data 1870-1930

Birthplace	1870	1880	1890	1900	1910	1920	1930
Asia	0.16	0.20	0.17	0.26	0.00	0.00	0.00
Australia and New Zealand	0.00	0.00	0.01	0.00	0.00	0.00	0.00
Austro-Hungarian Emp.	0.14	0.14	0.48	0.76	1.81	1.41	1.10
Benelux	0.11	0.05	0.17	0.17	0.17	0.17	0.15
Canada	1.28	1.44	1.57	1.55	2.56	1.99	2.06
Eastern Europe	0.02	0.07	0.52	1.07	1.80	2.64	2.28
France	0.30	0.21	0.18	0.13	0.13	0.14	0.11
Germany	4.40	3.95	4.45	3.50	2.70	1.59	1.32
Great Britain and Ireland	6.83	5.56	4.99	3.66	2.78	2.04	1.76
Greece	0.00	0.00	0.00	0.01	0.11	0.17	0.13
Italy	0.02	0.05	0.29	0.64	1.45	1.52	1.47
Portugal	0.00	0.00	0.02	0.05	0.07	0.07	0.05
Rest Of America	0.10	0.14	0.16	0.17	0.26	0.48	0.06
Scandinavia	0.61	0.83	1.49	1.48	1.49	1.24	1.04
Spain	0.00	0.00	0.01	0.00	0.02	0.04	0.04
Switzerland	0.20	0.10	0.16	0.15	0.13	0.11	0.09
All migrants	14.18	12.74	14.67	13.60	15.48	13.62	11.68
Within migrants diversity (Theil)	1.40	1.49	1.78	2.00	2.13	2.19	2.12

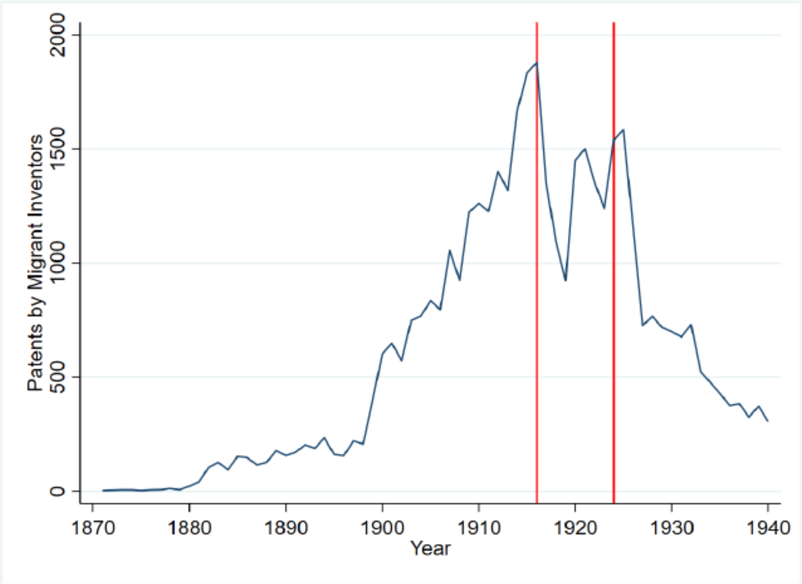
[1] Data source: NHGIS IPUMS county-level decennial census files (Manson et al., 2019). Each row indicates the (%) share, out of U.S. total population, of immigrants by foreign birthplace and decade from 1880 until 1930. Last two rows report, respectively, the (%) share of foreign-born population and the Theil index of diversity within the foreign-born population.

Descriptive Stats (Cont.)

Pantents and migrant inventors over time

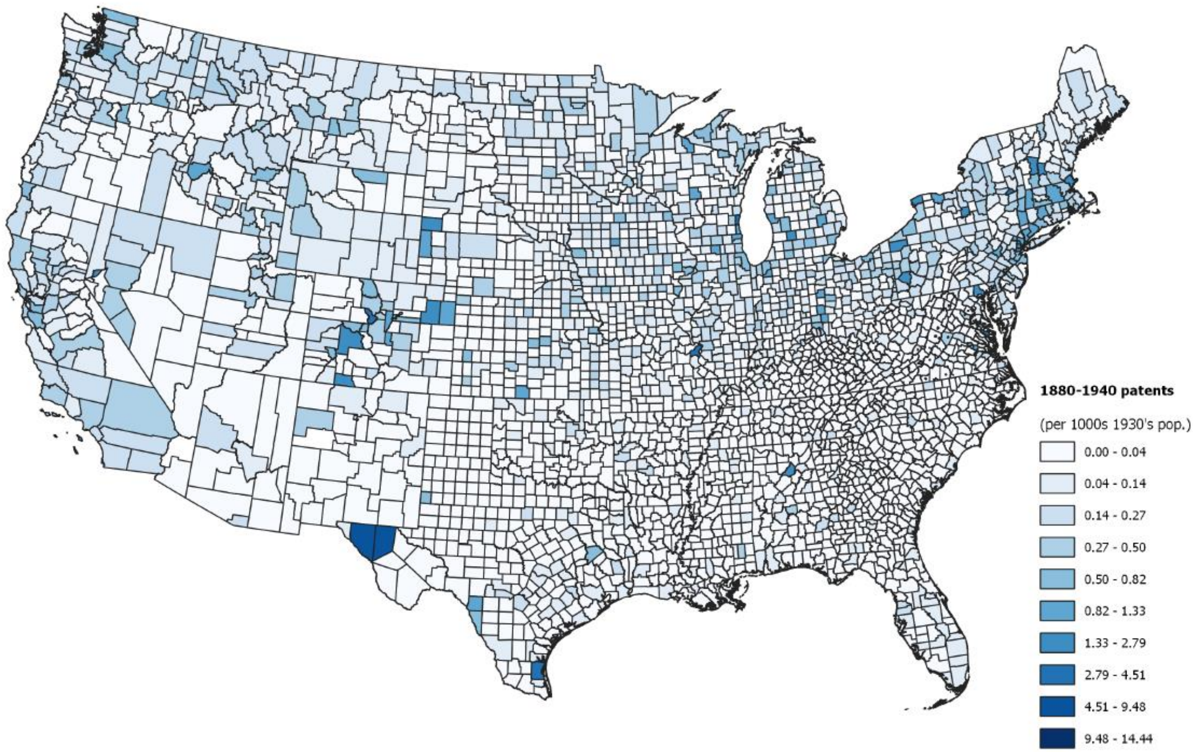


(a) Migrant inventors



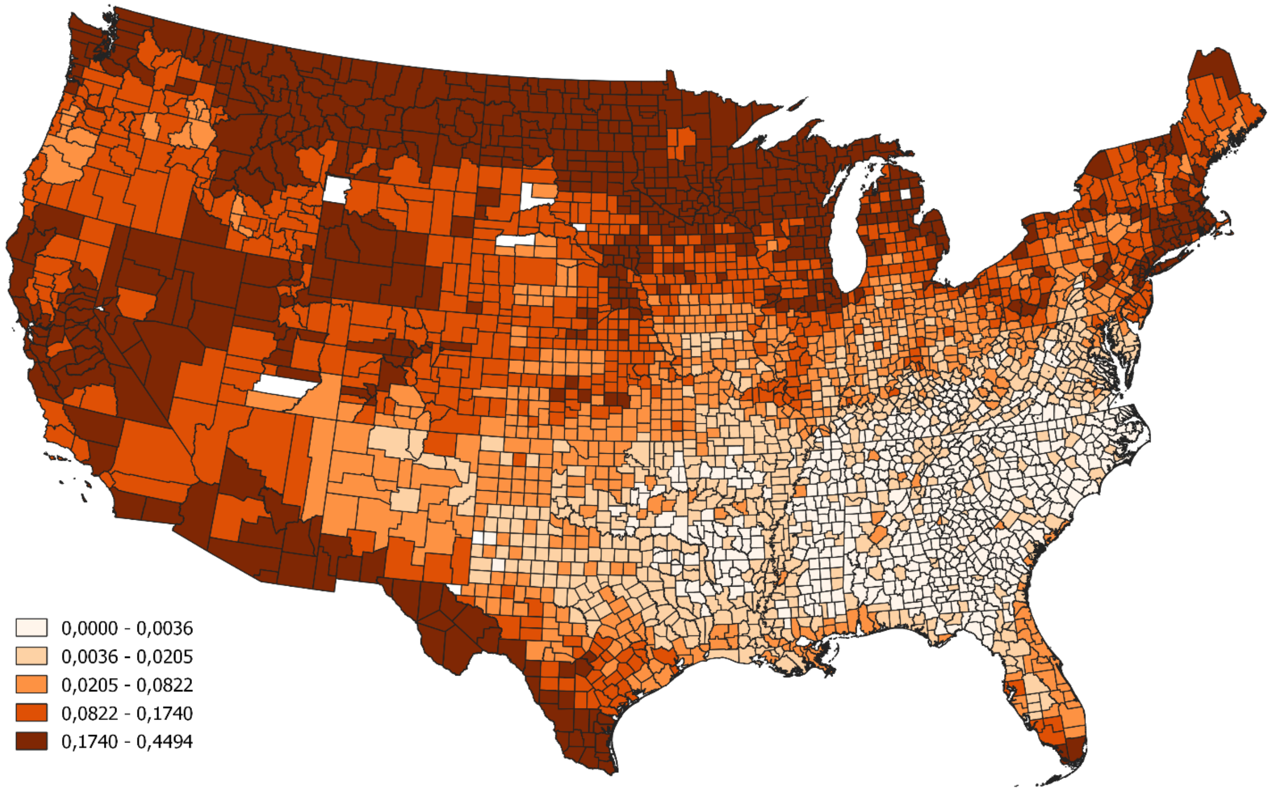
(b) Patents by migrant inventors

Descriptive Stats (Cont.)



Migrant inventors' patents by county
(on 1000s 1930 population)

Descriptive Stats (Cont.)



Share of foreign-born population by county
(average 1870-1940)

A Simple Spatial Economy

- Leaving aside causation (more on this below) observed patterns cannot be interpreted without a conceptual framework (Roback, 1982; Ottaviano & Peri, 2005, 2006)
- Open system of a large number N of counties indexed $c=1,\ldots,N$
- Two factors of production, mobile inventors L and immobile land H_c owned by locally resident landlords
- Inventors are differentiated in M groups indexed $i=1,\ldots,M$ in terms of non-market attributes ('cultural traits'): $\sum_i L_{ic}=L_c$, $\sum_c L_c=L$
- Inter-county commuting costs are prohibitive: for all inventors the counties of work and residence coincide
- No intra-county commuting costs: focus on inventors' inter-county location choices

A Simple Spatial Economy (Cont.)

- Local 'cultural diversity' d_c is measured in terms of the composition of resident groups and enters both production and consumption as a localized external effect
- Inventors' preferences are defined over the consumption of land H and a freely traded homogeneous good Y

$$U_{ic} = A_U(d_c)\, H_{ic}^{1-\mu} Y_{ic}^{\mu}$$

- $A_U(d_c)$ captures the 'utility effect' associated with local diversity d_c
- $A_U'(d_c)>0$ if diversity is 'consumption amenity'
- $A_U'(d_c)<0$ if diversity is 'consumption disamenity'

A Simple Spatial Economy (Cont.)

- Good Y is supplied by perfectly competitive firms exploiting inventions through a linear technology $Y_{jc} = N_{jc}$
- Inventions are themselves supplied by perfectly competitive firms employing both land and inventors as inputs with CRS technology

$$N_{jc} = A_Y(d_c)\, H_{jc}^{1-\alpha} L_{jc}^{\alpha}$$

- $A_Y(d_c)$ captures the 'productivity effect' associated with local diversity d_c
- $A_Y'(d_c)>0$ if diversity is 'production amenity'
- $A_Y'(d_c)<0$ if diversity is 'production disamenity'

A Simple Spatial Economy (Cont.)

- Taking good Y as numeraire (p_c=1), firms' free entry and exit imply that profits are zero in all counties

$$r_c^{1-\alpha} w_c^{\alpha} = (1-\alpha)^{1-\alpha} \alpha^{\alpha} A_Y(d_c)$$

- And inventors' free mobility implies that indirect utility V_{ic} is equalized across all counties

$$V_{ic} = (1-\mu)^{1-\mu} \mu^{\mu} A_u(d_c) \frac{E_{ic}}{r_c^{1-\mu}}$$

- $E_{ic}= w_c$ is inventor wage, r_c is land rent

A Simple Spatial Economy (Cont.)

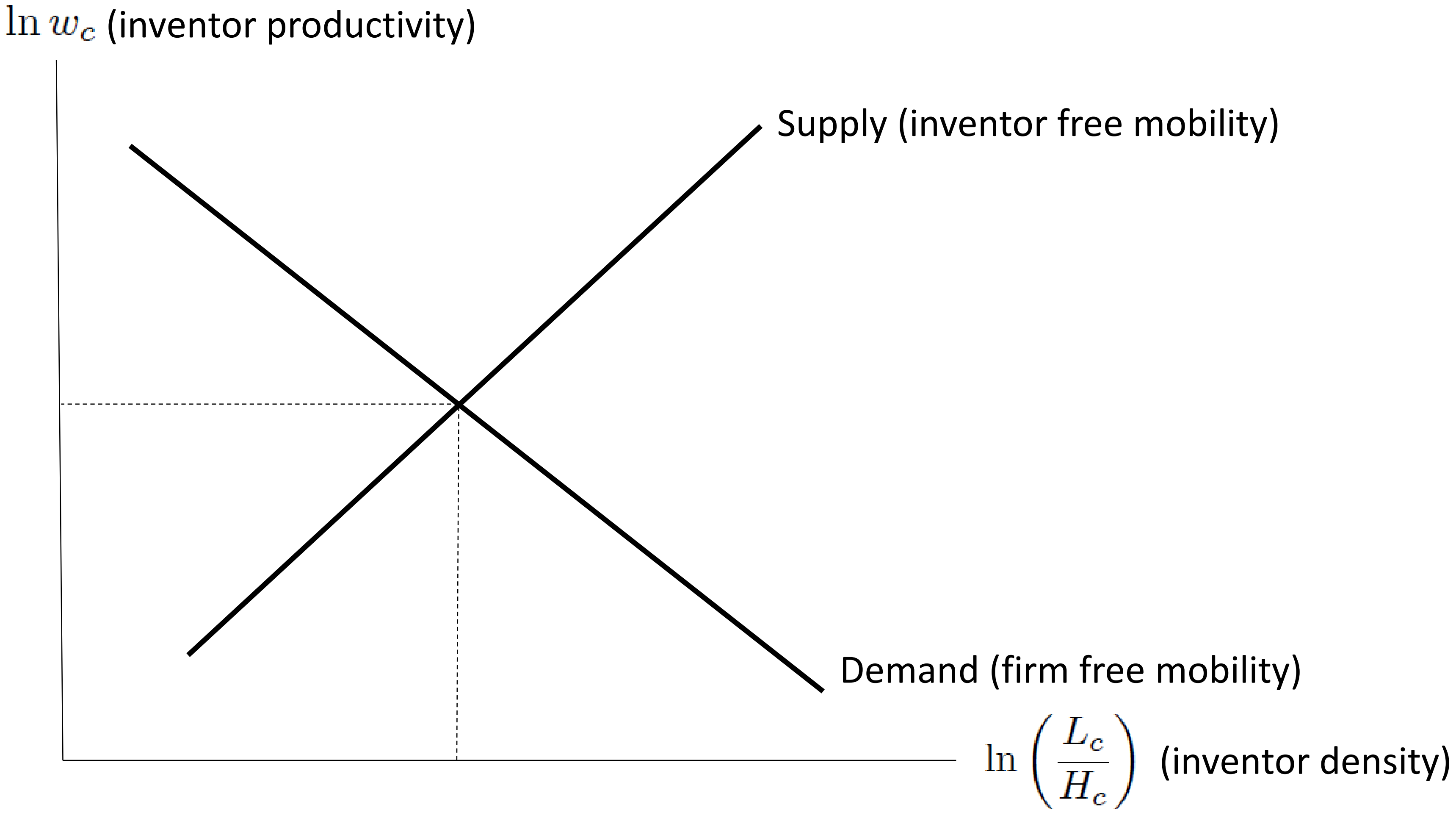
- The market for inventors is in equilibrium when demand

$$\ln w_c = \ln \Theta + \ln A_Y(d_c) - (1-\alpha) \ln \left(\frac{L_c}{H_c} \right)$$

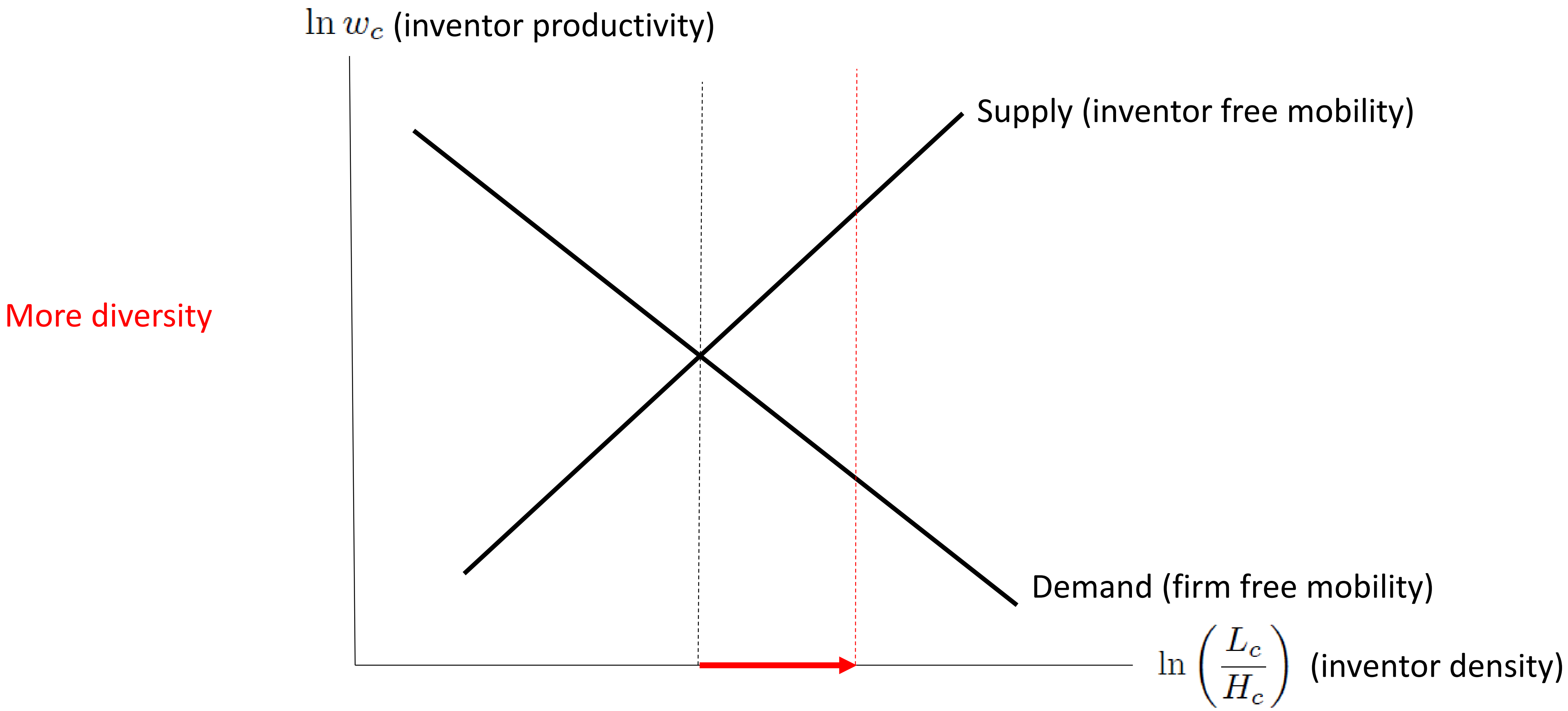
meets supply

$$\ln w_c = \ln \Omega - \frac{1}{\mu} \ln A_u(d_c) + \frac{1-\mu}{\mu} \ln \left(\frac{L_c}{H_c} \right)$$

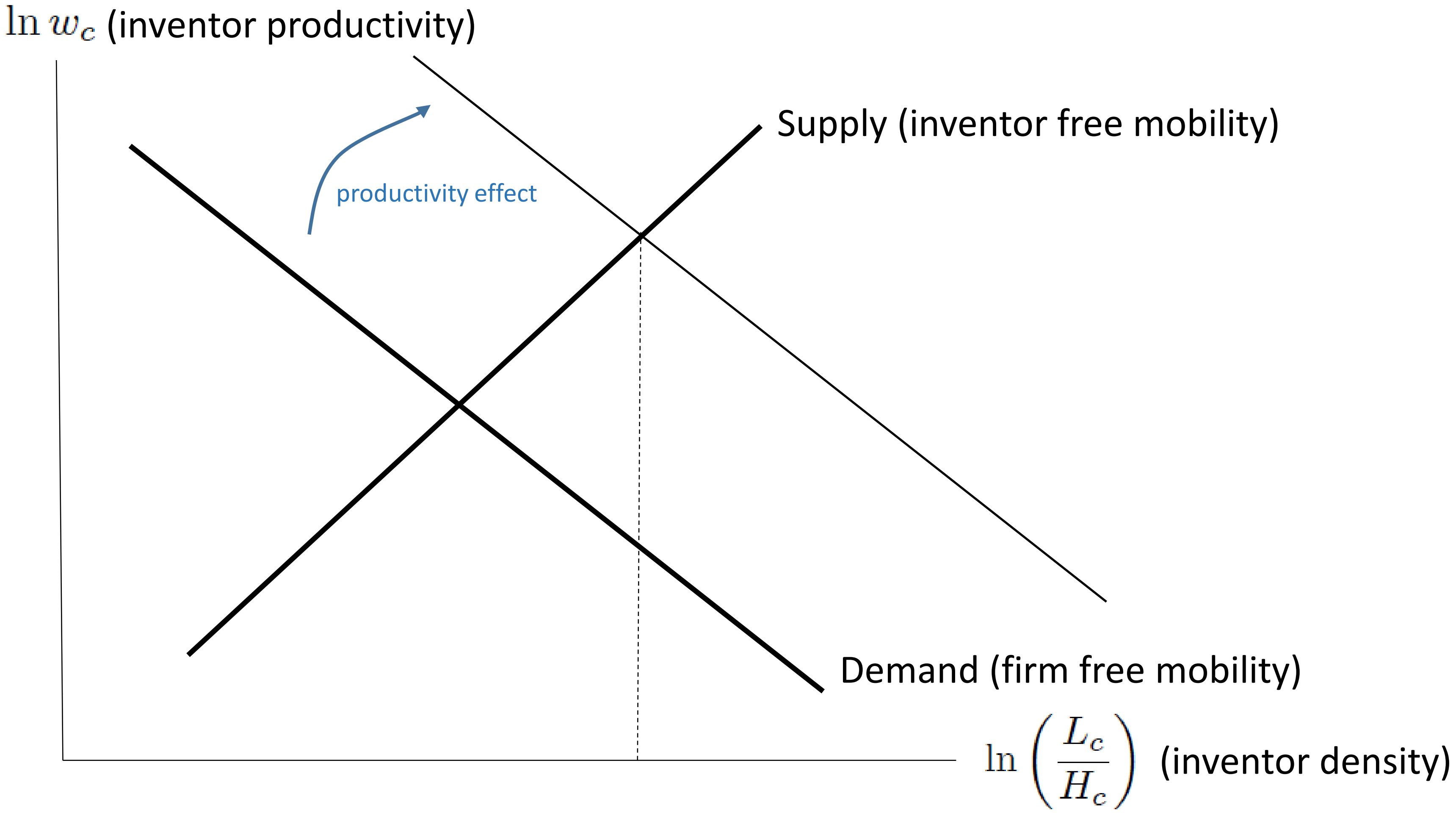
A Simple Spatial Economy (Cont.)



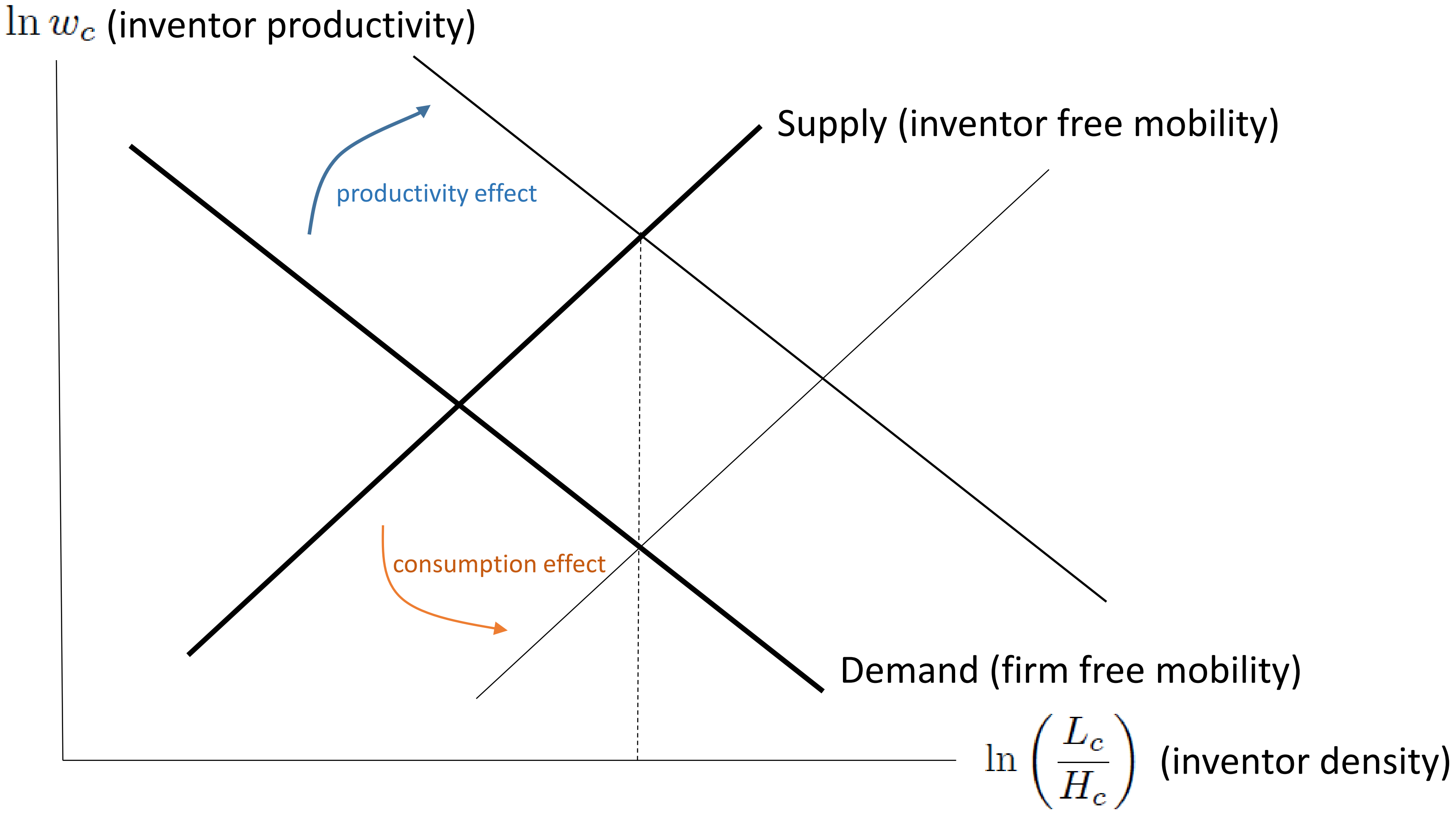
A Simple Spatial Economy (Cont.)



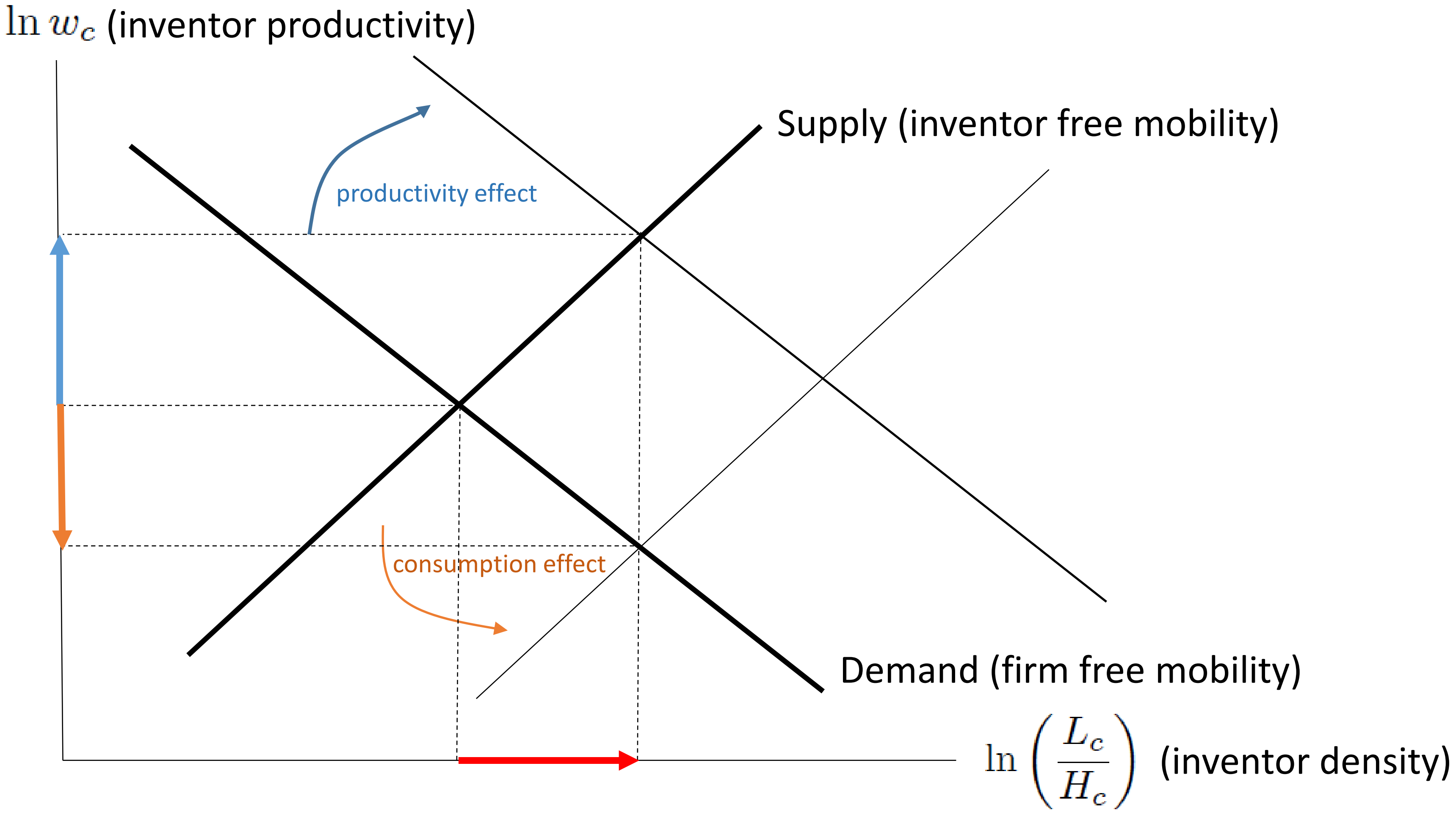
A Simple Spatial Economy (Cont.)



A Simple Spatial Economy (Cont.)



A Simple Spatial Economy (Cont.)



Empirical Strategy

- Unit of analysis is the sub-population ‘cell’ as defined by US county of residence and birthplace (i.e. ethnicity)
- Study how within-cell changes in ‘cultural diversity’ affect within-cell changes in immigrant inventors’ outcomes
- Exploit decennial variation within ethnicity-county cells, while controlling for a set of fixed factors and time-varying control variables
- All explanatory variables standardized: coefficients reflect how a standard deviation change in the explanatory variables is associated on average with changes in the dependent variable

Empirical Strategy (Cont.)

Benchmark specification:

$$\ln(Y_{ecst}) = \alpha_0^y + \beta_1^y s_{ecst} + \beta_2^y s_{-ecst} + \beta_3^y Theil_{-ecst} + \delta_{st}^y + \mu_{ec}^y + t\pi_{ec}^y + \epsilon_{ecst}^y$$

- where Y_{ecst} is the outcome for inventors in ethnic group (country of origin) e, resident in county c in state s at time t.
- Y_{ecst} is either the (log) number of inventors in ethnic group e or their (log) average patenting productivity.
- Variables of interest are s_{ecst} that is a meaure of co-ethnic network, while s_{-ecst} and $Theil_{-ecst}$ measure between-group and within-group diversity respectively.
- We control for ethnicity-by-county fixed effects μ_{ec}, state-by-year fe δ_{st} and ethnicity-by-country time-linear trends $t\pi_{ec}$

Identification: shift share IV

We consider 1870 as reference year and define the predicted change in the stock of ethnic group e in county c between census year t – 1 and t .

$$\Delta\widehat{N}_{ecst} = s^{US}_{ecs,1870} \times \Delta N_{e,-s,[t-1;t]} \qquad t = 1880, .., 1930$$

- where $\Delta N_{e,-s,[t-1;t]}$ is the (leave-out versione of the) aggregate shift component (i.e. the change in the stock of immigrants from group e btw t – 1 and t in US)
- $s^{US}{}_{ecs,1870}$ is the share of immigrants from e in s and c (out of all migrants from e in the US)

We then compute the predicted stock of immigrants from e in county c for year t as

$$\widehat{N}_{ecst} = N_{ecs,1870} + \sum_{\tau \le t} \Delta\widehat{N}_{ecs\tau} \qquad t = 1880, .., 1930.$$

We finally use the predicted stocks to compute the shift share IVS for between and within diversity variables

Identification: Quasi-Experiment from the '1920 quota-system'

In 1922 the 'quota system' is introduced:

- set the yearly inflow from a given country to be equal to a small percentage (3%) of the stock of co–nationals living in the US in 1900. (from 1924 on, the reference year for the quota calculation switched to 1890);
- the quota regime generated an asymmetric (negative) shock on immigration from different nationalities:
- Southern and Eastern Europeans mostly affected, with a substantial reversion of the trend in immigration during 1920s and 1930s, if compared to the period 1900-1914;
- immigration from Northern Europe barely affected: i) in 1890 they were the majority of the foreign born population in the US $\rightarrow$higher quota; ii) immigration from Northern Europe significantly slowed down from 1900 onwards

Quota exposure by foreign nationality

Ethnicity	(1) Avg yearly inflow 1900-1914	(2) Avg yearly quota 1922-1930	(3) Quota exposure
Asia	9,243	2,022	0.78
Australia and New Zealand	454	537	0
Austro-Hungarian Emp.	75,026	14,571	0.81
Benelux	6,546	3,419	0.48
Canada	26,253	Unrestricted	0
Eastern Europe	139,383	29,762	0.79
France	4,093	4,449	0
Germany	23,976	54,086	0
Great Britain and Ireland	52,498	69,830	0
Greece	8,186	1,162	0.86
Italy	78,037	16,823	0.78
Portugal	3,882	1,156	0.70
Rest Of America	18,720	0	0
Scandinavia	34,956	25,471	0.27
Spain	1,718	405	0.76
Switzerland	2,537	2,596	0

[1] Column 1 indicates the average number of arrivals by birthplace between 1900 and 1914 (source: 1920 IPUMS Full-Count Census micro-data (Ruggles et al., 2003)). Column 2 reports the average quota by nationality between 1922 and 1930, i.e. the maximum number of new arrivals to US allowed by 1921 and 1924's Immigration Acts (source: Census Statistical Abstract 1931). Column 3 displays the values of aggregate quota exposure by ethnicity as defined in (15) (Ager and Hansen, 2017).

Quasi-Experiment: 'quota-exposure'

We define an ethnicity-by-county measure of quota-exposure during the 1920s (Ager and Hansen, 2017)

$$Q.exp._{ecs,1930} = s^{US}_{ecs,1920} \times \max\left(\frac{Imm_{e,00-14} - Q_e}{Imm_{e,00-14}}, 0\right)$$

- where $s^{US}{}_{ecs,1920}$ is the county share of migrants from e alreay in the US in 1920.
- $Imm_{e,00-14}$ is the yearly migation inflow from county c to the US from 1900 and 1914
- Q_e is the yearly number of immigrants from e allowed to enter the US by the corresponding quota btw 1922 and 1930.
- The ratio in the *max (.)* measures the quota exposure for foreign group e in the US and ranges btw 0 and 1.

Quasi-Experiment: 'WWI-exposure'

We construct an ethnicity-by-county measure of WWI-exposure during the 1910s (Tabellini, 2020)

$$WWI.exp._{ecs,1920} = s^{US}_{ecs,1910} \times Enemy_e \times Imm_{e,00-10}$$

- where $Enemy_e$ is a dummy equal to 1 for enemy countries (Germany and Austro-Hungarian Empire).
- $Imm_{e,00-10}$ is the average yearly migation inflow from country e to the US from 1900 and 1910
- $s^{US}{}_{ecs,1910}$ is the county share of total migrants from e already in the US in 1910.

The rationale for using $WWI.exp_{ecs,1920}$ and $Q.exp_{ecs,1930}$ to build instruments for s_{ecst}, s_{-ecst} and $Theil_{-ecst}$… is that counties w/ higher shares of WWI- and quota-affected ethinic groups are expected to experience less immigration growth from those ethinic groups.

Results- First stage shift share instrument

	(1) Network s_{ecst}	(2) Within diversity $Theil_{-ecst}$	(3) Between diversity s_{-ecst}
shift-share $\widehat{s}_{ecst}$	0.3456***	-0.0501***	-0.0214***
	(0.0277)	(0.0069)	(0.0055)
shift-share $\widehat{Theil}_{-ecst}$	0.0050	0.0948***	0.0301***
	(0.0057)	(0.0105)	(0.0056)
shift-share $\widehat{s}_{-ecst}$	-0.0549***	-0.4340***	0.3051***
	(0.0104)	(0.0232)	(0.0188)
Observations	171,990	171,990	171,990
Ethnicity by County FE	Yes	Yes	Yes
Year by State FE	Yes	Yes	Yes
Ethn. by County time-linear trends	Yes	Yes	Yes
S&W Weak identification test	203.7	168.6	138.8

Results

A) Dep. var: log(number of immigrant inventors)				
	(1)	(2)	(3)	(4)
	OLS		Shift-Share IV	
	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$
Within Diversity: $Theil_{-ecst}$	0.0184***	0.0266***	0.0288**	0.4242***
	(0.0019)	(0.0026)	(0.0140)	(0.0716)
Between Diversity: s_{-ecst}	0.0012***	0.0052***	0.0111***	0.0518***
	(0.0002)	(0.0005)	(0.0022)	(0.0065)
Network: s_{ecst}	0.0101***	0.0320***	0.0235***	0.0908***
	(0.0015)	(0.0039)	(0.0047)	(0.0111)
Observations	171,990	171,990	171,990	171,990
R-squared	0.6482	0.7195		
B) Dep. var: log(immigrant inventors' productivity)				
	(1)	(2)	(3)	(4)
	OLS		Shift-Share IV	
	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$
Within Diversity: $Theil_{-ecst}$	0.0139***	0.0157***	0.0032	0.1923***
	(0.0019)	(0.0023)	(0.0152)	(0.0556)
Between Diversity: s_{-ecst}	0.0007***	0.0027***	0.0016	0.0237***
	(0.0002)	(0.0004)	(0.0018)	(0.0044)
Network: s_{ecst}	0.0046***	0.0146***	0.0088***	0.0529***
	(0.0011)	(0.0024)	(0.0033)	(0.0079)
Observations	171,990	171,990	171,990	171,990
R-squared	0.5011	0.6302		
Ethnicity by County FE	Yes	Yes	Yes	Yes
Year by State FE	Yes	Yes	Yes	Yes
Ethn. by County time-linear trends		Yes		Yes

Results (w/ 1870 pop control)

A) Dep. var: log(number of immigrant inventors)				
	(1)	(2)	(3)	(4)
	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$
	Baseline			
Within Diversity: $Theil_{-ecst}$	0.4242***	0.3056***	0.2187***	0.2547***
	(0.0716)	(0.0633)	(0.0624)	(0.0630)
Between Diversity: s_{-ecst}	0.0518***	0.0395***	0.0310***	0.0241***
	(0.0065)	(0.0055)	(0.0056)	(0.0054)
Network: s_{ecst}	0.0908***	0.0788***	0.0709***	0.0597***
	(0.0111)	(0.0103)	(0.0102)	(0.0090)
Observations	171,990	170,820	170,820	171,990
B) Dep. var: log(immigrant inventors productivity)				
	(1)	(2)	(3)	(4)
	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$
	Baseline			
Within Diversity: $Theil_{-ecst}$	0.1923***	0.1370***	0.0913*	0.1377**
	(0.0556)	(0.0523)	(0.0538)	(0.0573)
Between Diversity: s_{-ecst}	0.0237***	0.0179***	0.0136***	0.0181***
	(0.0044)	(0.0040)	(0.0043)	(0.0047)
Network: s_{ecst}	0.0529***	0.0472***	0.0434***	0.0474***
	(0.0079)	(0.0076)	(0.0078)	(0.0081)
Observations	171,990	170,820	170,820	171,990
Ethnicity by County FE	Yes	Yes	Yes	Yes
Year by State FE	Yes	Yes	Yes	Yes
Ethn. by County time-linear trends	Yes	Yes	Yes	Yes
1870 (log) pop × Year		Yes	Yes	
1870 controls × Year			Yes	
1870 ethnicities shares × Year				Yes

Results (w/ pop size and frontier exposure control)

A) Dep. var: log(number of immigrant inventors)

	(1)	(2)	(3)	(4)
	OLS		2SLS	
	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$
Within Diversity: $Theil_{-ecst}$	0.0247***	0.0245***	0.3863***	0.3858***
	(0.0025)	(0.0025)	(0.0659)	(0.0660)
Between Diversity: s_{-ecst}	0.0046***	0.0046***	0.0527***	0.0526***
	(0.0005)	(0.0005)	(0.0066)	(0.0066)
Network: s_{ecst}	0.0315***	0.0315***	0.0919***	0.0916***
	(0.0039)	(0.0039)	(0.0111)	(0.0111)
$log(pop)_{cst}$	0.0261***	0.0265***	-0.1523***	-0.1517***
	(0.0075)	(0.0075)	(0.0244)	(0.0244)
Years since exposure to frontier		0.0005***		0.0003
		(0.0002)		(0.0002)
Observations	171,990	171,990	171,990	171,990

B) Dep. var: log(immigrant inventors productivity)

	(1)	(2)	(3)	(4)
	OLS		2SLS	
	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$
Within Diversity: $Theil_{-ecst}$	0.0152***	0.0152***	0.1738***	0.1739***
	(0.0023)	(0.0023)	(0.0515)	(0.0515)
Between Diversity: s_{-ecst}	0.0025***	0.0025***	0.0241***	0.0242***
	(0.0004)	(0.0004)	(0.0045)	(0.0045)
Network: s_{ecst}	0.0145***	0.0145***	0.0534***	0.0534***
	(0.0024)	(0.0024)	(0.0079)	(0.0079)
$log(pop)_{cst}$	0.0070	0.0070	-0.0745***	-0.0746***
	(0.0048)	(0.0047)	(0.0182)	(0.0182)
Years since exposure to frontier		0.0001		-0.0000
		(0.0002)		(0.0002)
Observations	171,990	171,990	171,990	171,990
Ethnicity by County FE	Yes	Yes	Yes	Yes
Year by State FE	Yes	Yes	Yes	Yes
Ethn. by County time-linear trends	Yes	Yes	Yes	Yes

Heterogenous effects by 1880 county pop.size

	(1)	(2)	(3)	(4)	(5)	(6)
A) Dep. var: log(number of immigrant inventors)						
	1st tercile $pop_{c1880} <= 9798$		2nd tercile $9806 >= pop_{c1880} <= 18831$		3rd tercile $pop_{c1880} >= 18854$	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$	$log(L)_{ecst}$
Within Diversity: $Theil_{-ecst}$	0.0067***	-0.6576	0.0041*	0.1048	0.0572***	0.3761***
	(0.0021)	(0.9864)	(0.0021)	(0.0942)	(0.0089)	(0.0689)
Between Diversity: s_{-ecst}	0.0007**	-0.0520	0.0020***	0.0182	0.0179***	0.0915***
	(0.0003)	(0.0819)	(0.0007)	(0.0189)	(0.0018)	(0.0078)
Network: s_{ecst}	0.0044*	-0.0286	0.0175***	0.0500**	0.0774***	0.1394***
	(0.0026)	(0.0861)	(0.0061)	(0.0219)	(0.0087)	(0.0186)
Observations	48,690	48,690	54,900	54,900	68,400	68,400
B) Dep. var: log(immigrant inventors productivity)						
	(1)	(2)	(3)	(4)	(5)	(6)
	1st tercile $pop_{c1880} <= 9798$		2nd tercile $9806 >= pop_{c1880} <= 18831$		3rd tercile $pop_{c1880} >= 18854$	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$	$log(T)_{ecst}$
Within Diversity: $Theil_{-ecst}$	0.0061**	-1.0565	0.0062**	0.2122	0.0294***	0.1688***
	(0.0024)	(1.4558)	(0.0027)	(0.1446)	(0.0079)	(0.0534)
Between Diversity: s_{-ecst}	0.0005	-0.0834	0.0012	0.0330	0.0080***	0.0360***
	(0.0003)	(0.1213)	(0.0009)	(0.0270)	(0.0012)	(0.0060)
Network: s_{ecst}	0.0017	-0.0663	0.0143***	0.0655**	0.0316***	0.0686***
	(0.0020)	(0.1260)	(0.0052)	(0.0304)	(0.0051)	(0.0119)
Observations	48,690	48,690	54,900	54,900	68,400	68,400
Ethnicity by County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year by State FE	Yes	Yes	Yes	Yes	Yes	Yes
Ethn. by County time-linear trends	Yes	Yes	Yes	Yes	Yes	Yes

Robustness- First-stage quota and WWI instruments

	(1) Stage-zero	(2)	(3) 1st stage regressions	(4)
	ΔN_{ecst}	$\Delta Theil_{-ecst}$	Δs_{-ecst}	Δs_{ecst}
$1920 \times WWI.exp._{ecs,1920}$	-1.1664***			
	(0.3214)			
$1930 \times Q.exp._{ecs,1930}$	-270614.5328**			
	(126,261.5296)			
$1920 \times WWI\text{-}\Delta\widehat{Theil}_{-ecs1920}$		0.0354***	0.0135***	0.0032
		(0.0036)	(0.0016)	(0.0022)
$1930 \times Q\text{-}\Delta\widehat{Theil}_{-ecs1930}$		0.0615***	-0.0150***	0.0090***
		(0.0039)	(0.0036)	(0.0026)
$1920 \times WWI\text{-}\Delta\widehat{s}_{-ecs1920}$		0.0153***	0.1092***	0.0203***
		(0.0029)	(0.0033)	(0.0039)
$1930 \times Q\text{-}\Delta\widehat{s}_{-ecs1930}$		0.0274***	0.1083***	0.0203***
		(0.0021)	(0.0064)	(0.0037)
$1920 \times WWI\text{-}\Delta\widehat{s}_{ecs1920}$		0.0084***	0.0272***	0.0741***
		(0.0022)	(0.0030)	(0.0081)
$1930 \times Q\text{-}\Delta\widehat{s}_{ecs1930}$		0.0077***	0.0287***	0.0632***
		(0.0014)	(0.0037)	(0.0111)
Observations	171,795	171,795	171,795	171,795
Ethnicity by County FE	Yes	Yes	Yes	Yes
Year by State FE	Yes	Yes	Yes	Yes
First differences model Yes	Yes	Yes	Yes	Yes
S&W Weak identification test		101	252.7	30.43

Robustness- Results (quota and WWI instruments)

	(1)	(2)	(3)	(4)
	Immigrant inventors location choice		Immigrant inventors productivity	
	$\Delta log(L)_{ecst}$	$\Delta log(L)_{ecst}$	$\Delta log(T)_{ecst}$	$\Delta log(T)_{ecst}$
Within Diversity: $\Delta Theil\_{ecst}$	0.7004***	0.4200***	0.4343***	0.2599***
	(0.1436)	(0.0546)	(0.1288)	(0.0508)
Between Diversity: Δs_{ecst}	0.0118***	0.0239***	0.0047*	0.0134***
	(0.0029)	(0.0027)	(0.0025)	(0.0022)
Network: Δs_{ecst}	0.1025***	0.0927***	0.0183	0.0294***
	(0.0310)	(0.0206)	(0.0147)	(0.0098)
Observations	171,795	171,795	171,795	171,795
Ethnicity by County FE		Yes		Yes
Year by State FE	Yes	Yes	Yes	Yes
First differences model Yes	Yes	Yes	Yes	Yes

[1] All columns present 2SLS estimates employing WWI and Quota instruments as defined in Section 7.3. Columns 1 and 2 consider as outcome variable the 10 years-difference in (log) number of inventors from ethnicity e, living in county c, who are granted at least one patent between t and $t+1$. The outcome variable in Columns 3 and 4 is the 10 years-difference in (log) number of patents per inventor from ethnicity e and living in county c.

[2] All specifications correspond to a first-differenced version of the baseline model in (11), and include state by year fixed effects. Estimates in Columns 2 and 4 also adjust for ethnicity by county fixed effects. Standard errors clustered at ethnicity-by-county level in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Conclusion

- Using a shift-share approach and the quasi-experimental variation induced by US immigration quotas in the Age of Mass Migration, we find that:
  - Immigrant inventors are attracted to counties with more diversity
  - This is mainly due to the fact that diversity promotes their productivity
  - Diversity positively affects immigrant inventors' location choices on top of co-national networks
- Moreover:
  - We rule out alternative mechanisms such as inter-group connections (proxied by inter-ethnic marriages and residential contact), cultural proximity and natives' attitudes (proxied by migrant ethnic groups' salience in newspapers)– See Online Appendix.