Announcements

- Grades and comments on your Du Bois figures are up on Blackboard
- Grades for the first referee report will be up before Spring Break, email me if you did not get a confirmation
- Remember that the second referee report is due March 15th on Miller (2008) "Women's Suffrage, Political Responsiveness, and Child Survival in American History"
- Also think about making progress on your data projects
- No class on the Friday before Spring Break

Last Names in the US Context

- For the Clark and Cummins paper, check out the Stata and Excel examples on Blackboard for more on the issue of underlying social capital
- Now back to Clark's approach to the US
- Clark's sources for elite groups will be:
  - Descendants of Ashkenazi and Sephardic Jews
  - Descendants of wealthy individuals in 1923-24 with rare surnames
  - Descendants of individuals with rare surnames graduating from Ivy League schools in and before 1850
- Note how much more limited the time range needs to be

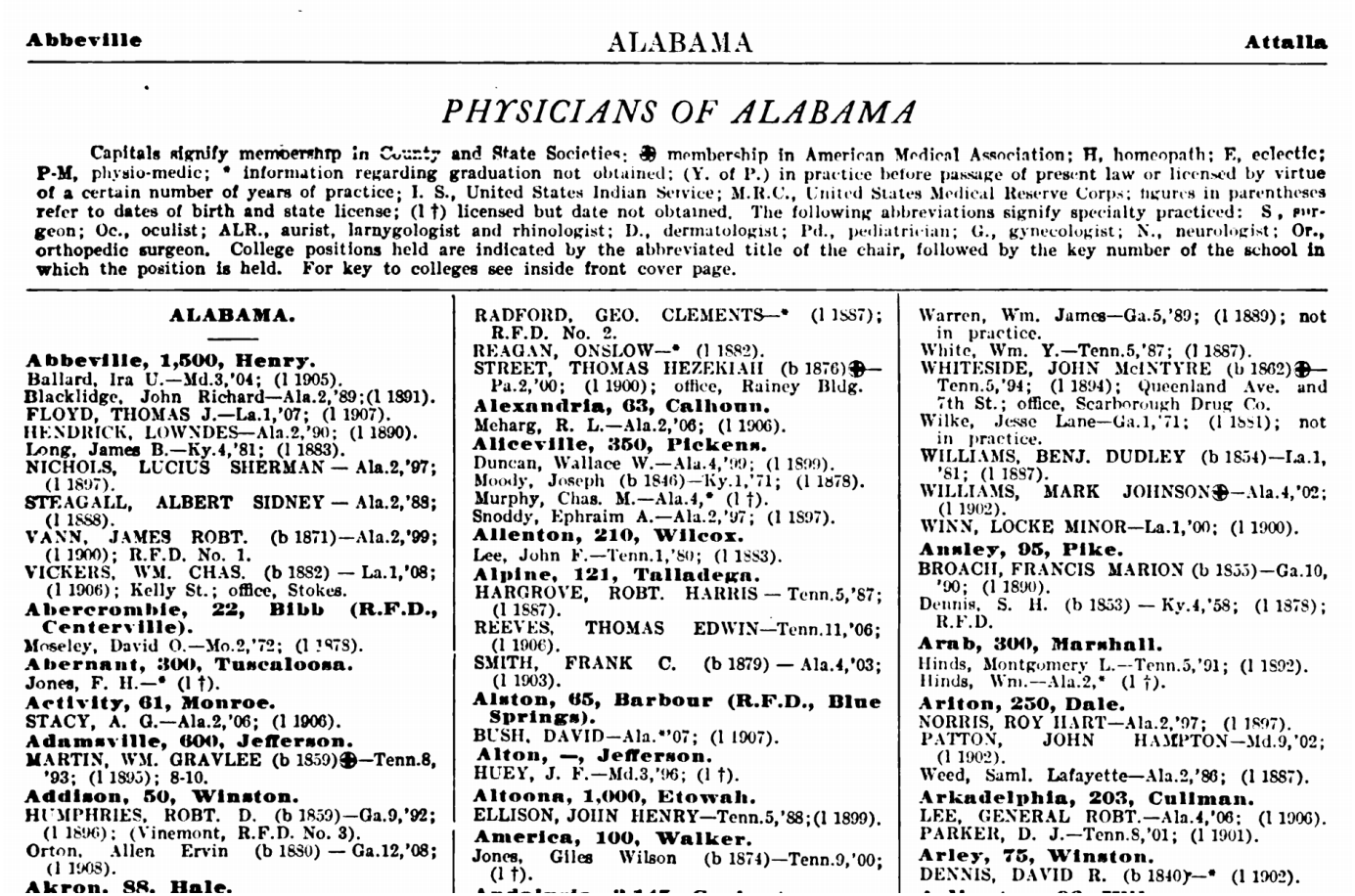
Last Names in the US Context

- Clark's sources for underclass groups will be:
  - Native Americans
  - Black Americans whose ancestors came to the United States before the Civil War
  - Descendants of the French settlers who came to the colonies between 1604 and 1759
- Think about how these groups differ from those used by Clark for Britain

Last Names in the US Context

- Measuring outcomes requires a different approach as well for the US
- Probate records are not as easily accessible (you would have to do a lot of work requesting one record at a time from many different locations)
- Instead, Clark is going to take an approach similar to looking at Cambridge and Oxford graduates
- He'll take advantage of the public directories of doctors and lawyers

Last Names in the US Context



Abbeville ALABAMA Attalla

PHYSICIANS OF ALABAMA

Capitals signify membership in County and State Societies; ⊕ membership in American Medical Association; H, homeopath; E, eclectic; P-M, physio-medic; * information regarding graduation not obtained; (Y. of P.) in practice before passage of present law or licensed by virtue of a certain number of years of practice; I. S., United States Indian Service; M.R.C., United States Medical Reserve Corps; figures in parentheses refer to dates of birth and state license; (l †) licensed but date not obtained. The following abbreviations signify specialty practiced: S., surgeon; Oc., oculist; ALR., aurist, larnygologist and rhinologist; D., dermatologist; Pd., pediatrician; G., gynecologist; N., neurologist; Or., orthopedic surgeon. College positions held are indicated by the abbreviated title of the chair, followed by the key number of the school in which the position is held. For key to colleges see inside front cover page.

ALABAMA.

Abbeville, 1,500, Henry.
Ballard, Ira U.—Md.3,'04; (l 1905).
Blacklidge, John Richard—Ala.2,'89;(l 1891).
FLOYD, THOMAS J.—La.1,'07; (l 1907).
HENDRICK, LOWNDES—Ala.2,'90; (l 1890).
Long, James B.—Ky.4,'81; (l 1883).
NICHOLS, LUCIUS SHERMAN — Ala.2,'97; (l 1897).
STEAGALL, ALBERT SIDNEY — Ala.2,'88; (l 1888).
VANN, JAMES ROBT. (b 1871)—Ala.2,'99; (l 1900); R.F.D. No. 1.
VICKERS, WM. CHAS. (b 1882) — La.1,'08; (l 1906); Kelly St.; office, Stokes.
Abercrombie, 22, Bibb (R.F.D., Centerville).
Moseley, David O.—Mo.2,'72; (l 1878).
Abernant, 300, Tuscaloosa.
Jones, F. H.—* (l †).
Activity, 61, Monroe.
STACY, A. G.—Ala.2,'06; (l 1906).
Adamsville, 600, Jefferson.
MARTIN, WM. GRAVLEE (b 1859)⊕—Tenn.8, '93; (l 1895); 8-10.
Addison, 50, Winston.
HUMPHRIES, ROBT. D. (b 1859)—Ga.9,'92; (l 1896); (Vinemont, R.F.D. No. 3).
Orton, Allen Ervin (b 1880) — Ga.12,'08; (l 1908).
Akron, 88, Hale.
RADFORD, GEO. CLEMENTS—* (l 1887); R.F.D. No. 2.
REAGAN, ONSLOW—* (l 1882).
STREET, THOMAS HEZEKIAH (b 1876)⊕—Pa.2,'00; (l 1900); office, Rainey Bldg.
Alexandria, 63, Calhoun.
Meharg, R. L.—Ala.2,'06; (l 1906).
Aliceville, 350, Pickens.
Duncan, Wallace W.—Ala.4,'99; (l 1899).
Moody, Joseph (b 1846)—Ky.1,'71; (l 1878).
Murphy, Chas. M.—Ala.4,* (l †).
Snoddy, Ephraim A.—Ala.2,'97; (l 1897).
Allenton, 210, Wilcox.
Lee, John F.—Tenn.1,'80; (l 1883).
Alpine, 121, Talladega.
HARGROVE, ROBT. HARRIS — Tenn.5,'87; (l 1887).
REEVES, THOMAS EDWIN—Tenn.11,'06; (l 1906).
SMITH, FRANK C. (b 1879) — Ala.4,'03; (l 1903).
Alston, 65, Barbour (R.F.D., Blue Springs).
BUSH, DAVID—Ala.*'07; (l 1907).
Alton, —, Jefferson.
HUEY, J. F.—Md.3,'96; (l †).
Altoona, 1,000, Etowah.
ELLISON, JOHN HENRY—Tenn.5,'88;(l 1899).
America, 100, Walker.
Jones, Giles Wilson (b 1874)—Tenn.9,'00; (l †).
Warren, Wm. James—Ga.5,'89; (l 1889); not in practice.
White, Wm. Y.—Tenn.5,'87; (l 1887).
WHITESIDE, JOHN McINTYRE (b 1862)⊕—Tenn.5,'94; (l 1894); Queenland Ave. and 7th St.; office, Scarborough Drug Co.
Wilke, Jesse Lane—Ga.1,'71; (l 1881); not in practice.
WILLIAMS, BENJ. DUDLEY (b 1854)—La.1, '81; (l 1887).
WILLIAMS, MARK JOHNSON⊕—Ala.4,'02; (l 1902).
WINN, LOCKE MINOR—La.1,'00; (l 1900).
Ansley, 95, Pike.
BROACH, FRANCIS MARION (b 1855)—Ga.10, '90; (l 1890).
Dennis, S. H. (b 1853) — Ky.4,'58; (l 1878); R.F.D.
Arab, 300, Marshall.
Hinds, Montgomery L.—Tenn.5,'91; (l 1892).
Hinds, Wm.—Ala.2,* (l †).
Arlton, 250, Dale.
NORRIS, ROY HART—Ala.2,'97; (l 1897).
PATTON, JOHN HAMPTON—Md.9,'02; (l 1902).
Weed, Saml. Lafayette—Ala.2,'86; (l 1887).
Arkadelphia, 203, Cullman.
LEE, GENERAL ROBT.—Ala.4,'06; (l 1906).
PARKER, D. J.—Tenn.8,'01; (l 1901).
Arley, 75, Winston.
DENNIS, DAVID R. (b 1840)—* (l 1902).

Last Names in the US Context

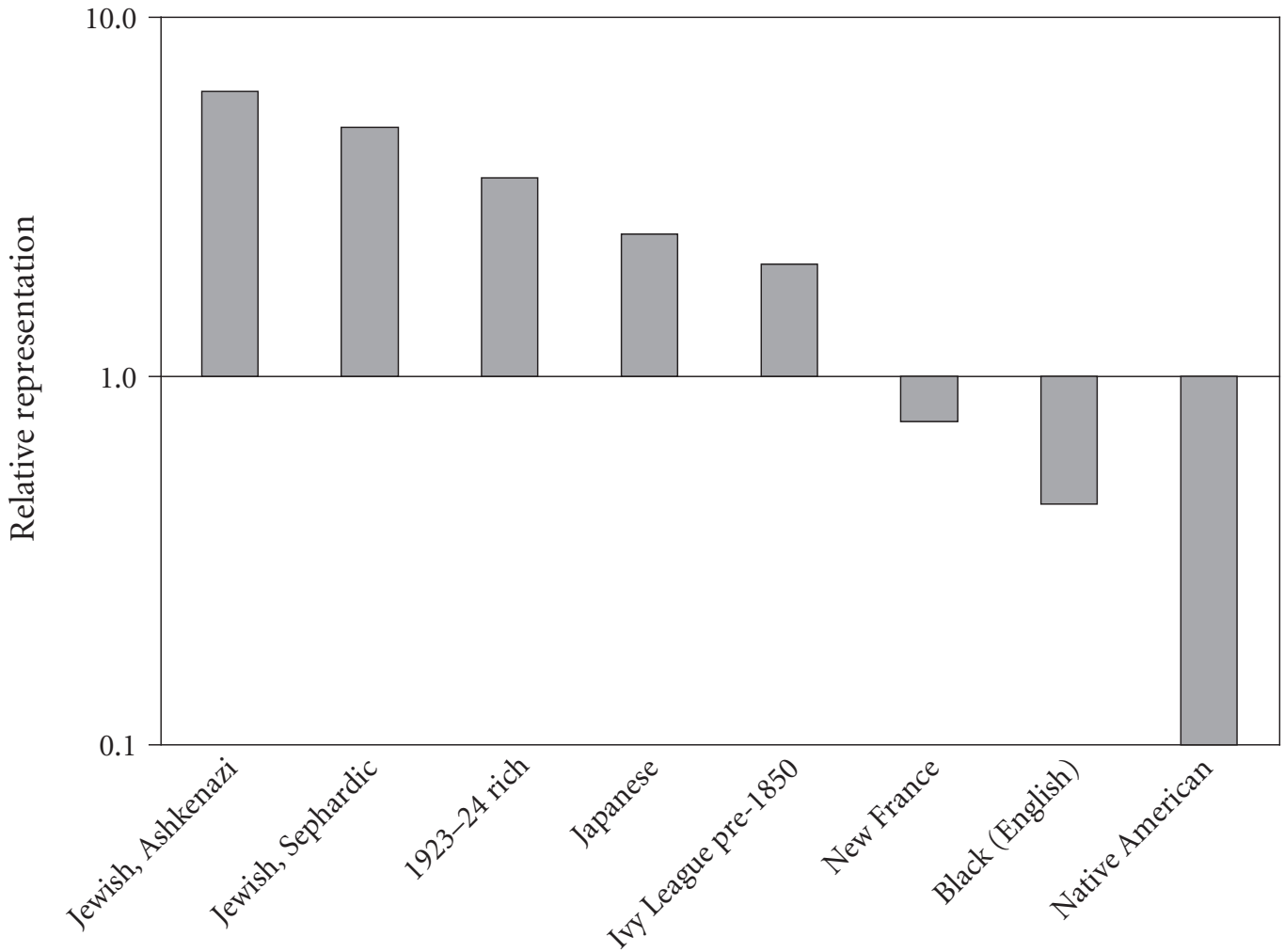


FIGURE 3.1. Relative representation of surname types among physicians.

Last Names in the US Context

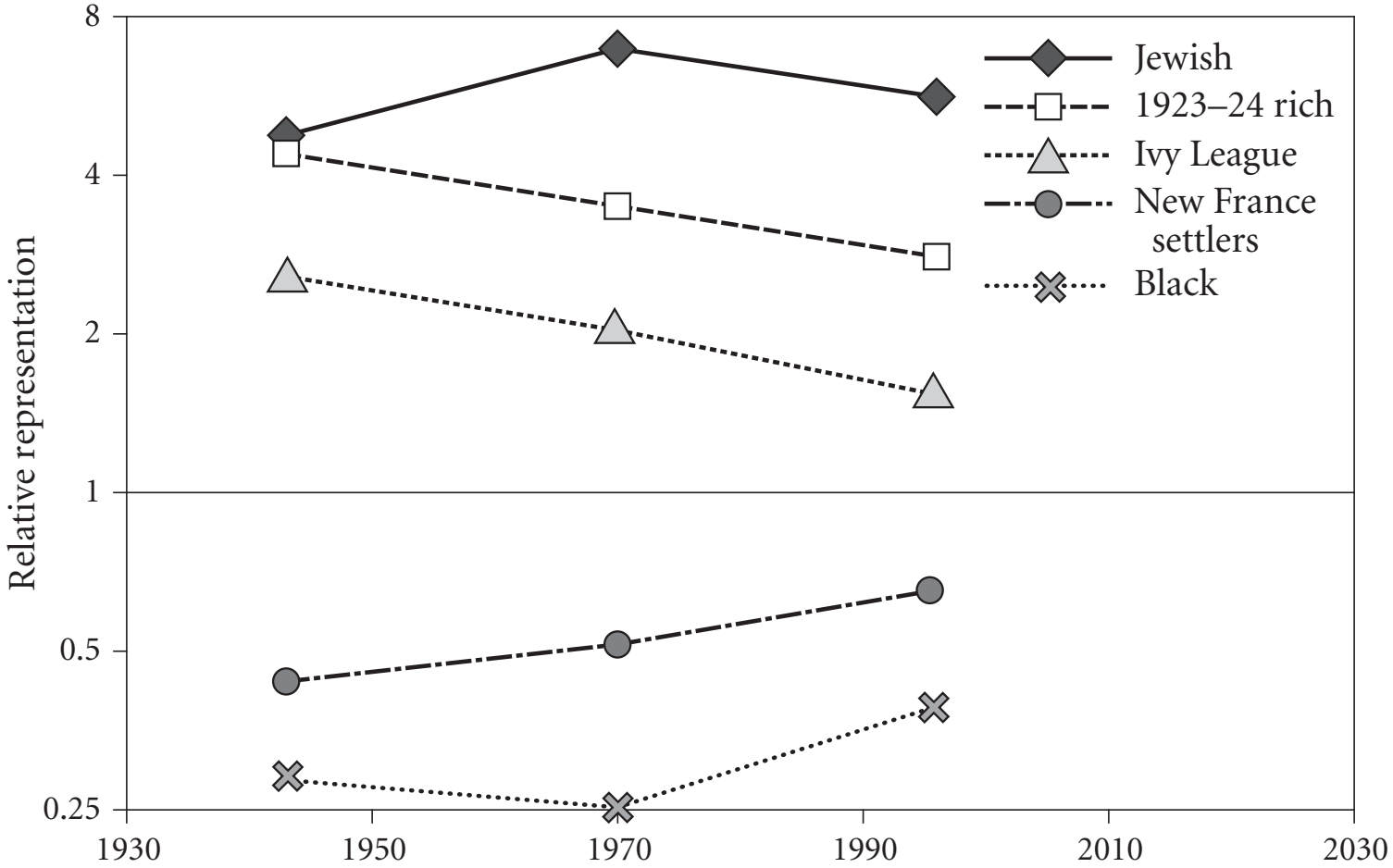


FIGURE 3.4. Relative representation of surname types among physicians, by generation.

Last Names in the US Context

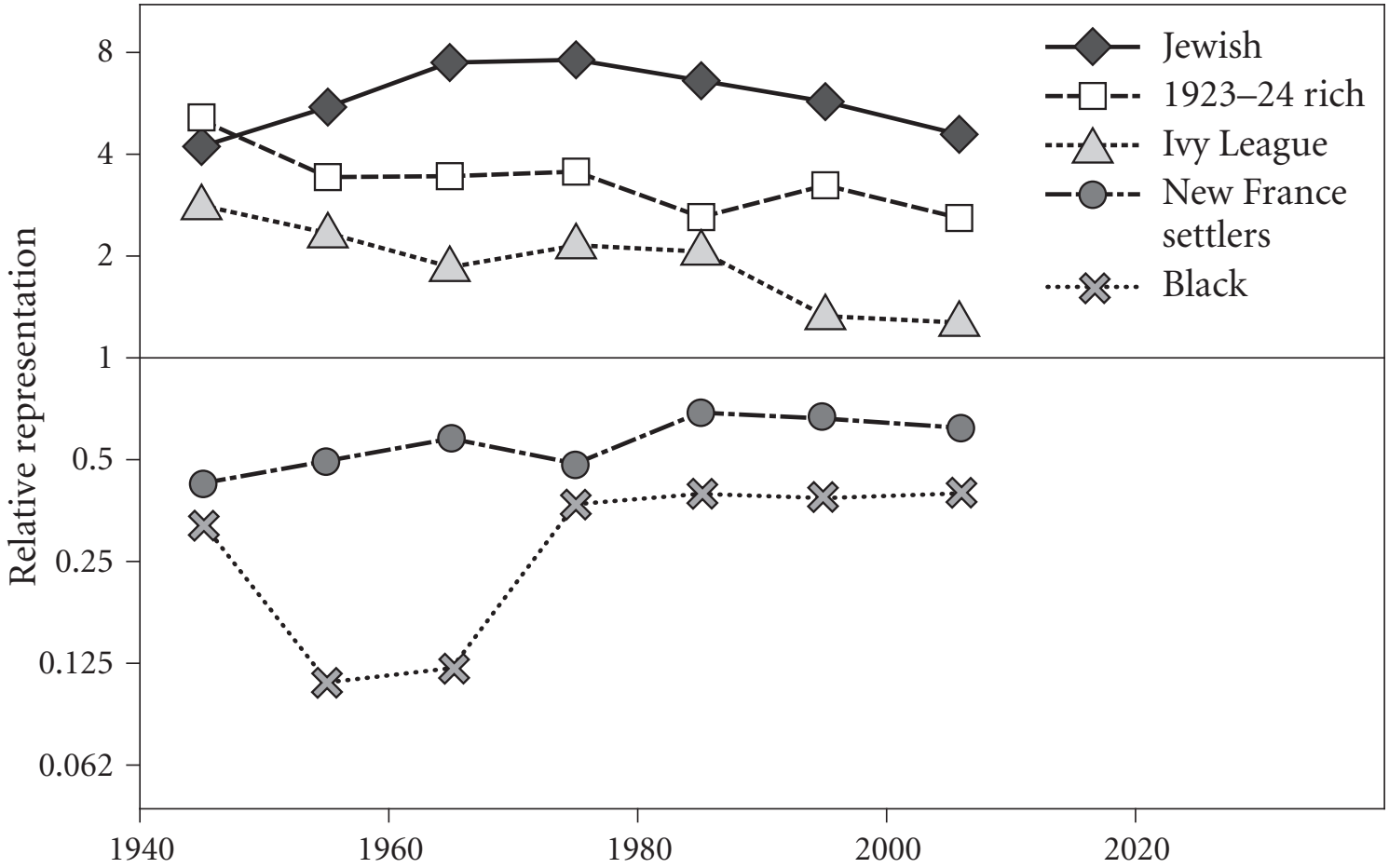


FIGURE 3.8. Relative representation of surname types among physicians, by decade.

Last Names in the US Context

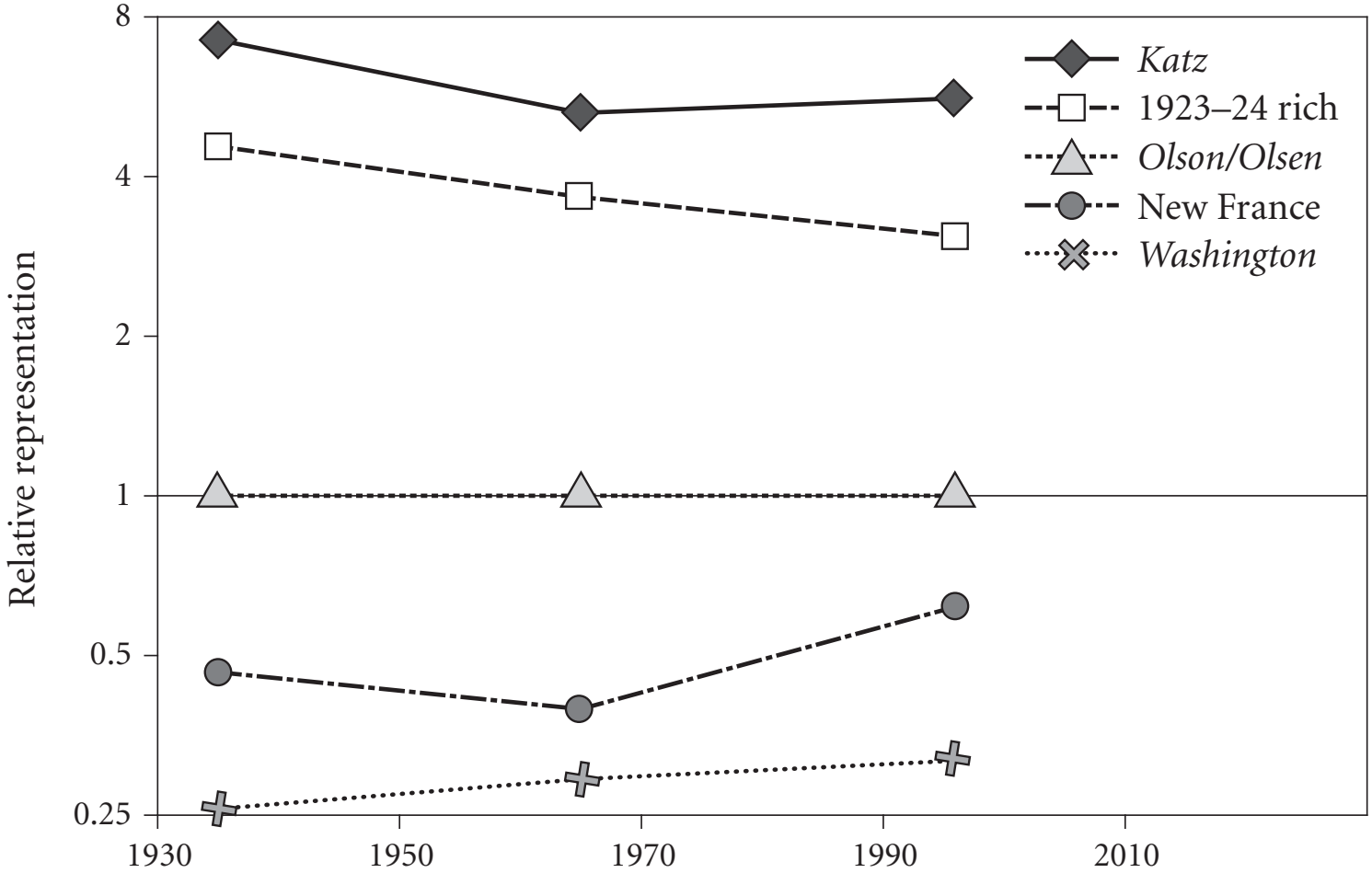


FIGURE 3.10. Relative representation of surname type among attorneys, by generation.

Last Names in the US Context

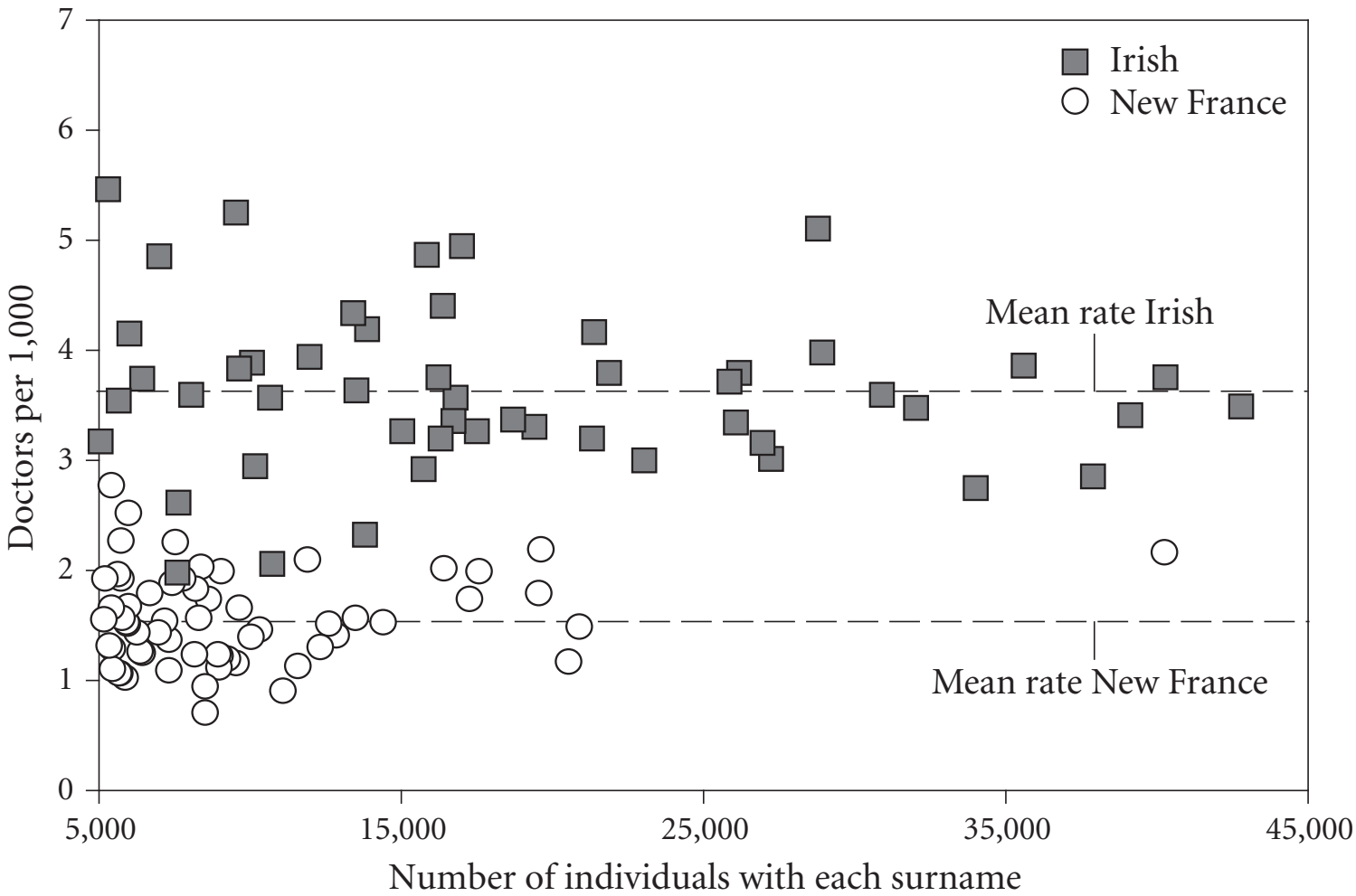


FIGURE 3.11. Physicians per thousand surname holders, most common Irish and New France surnames.

Last Names in the US Context

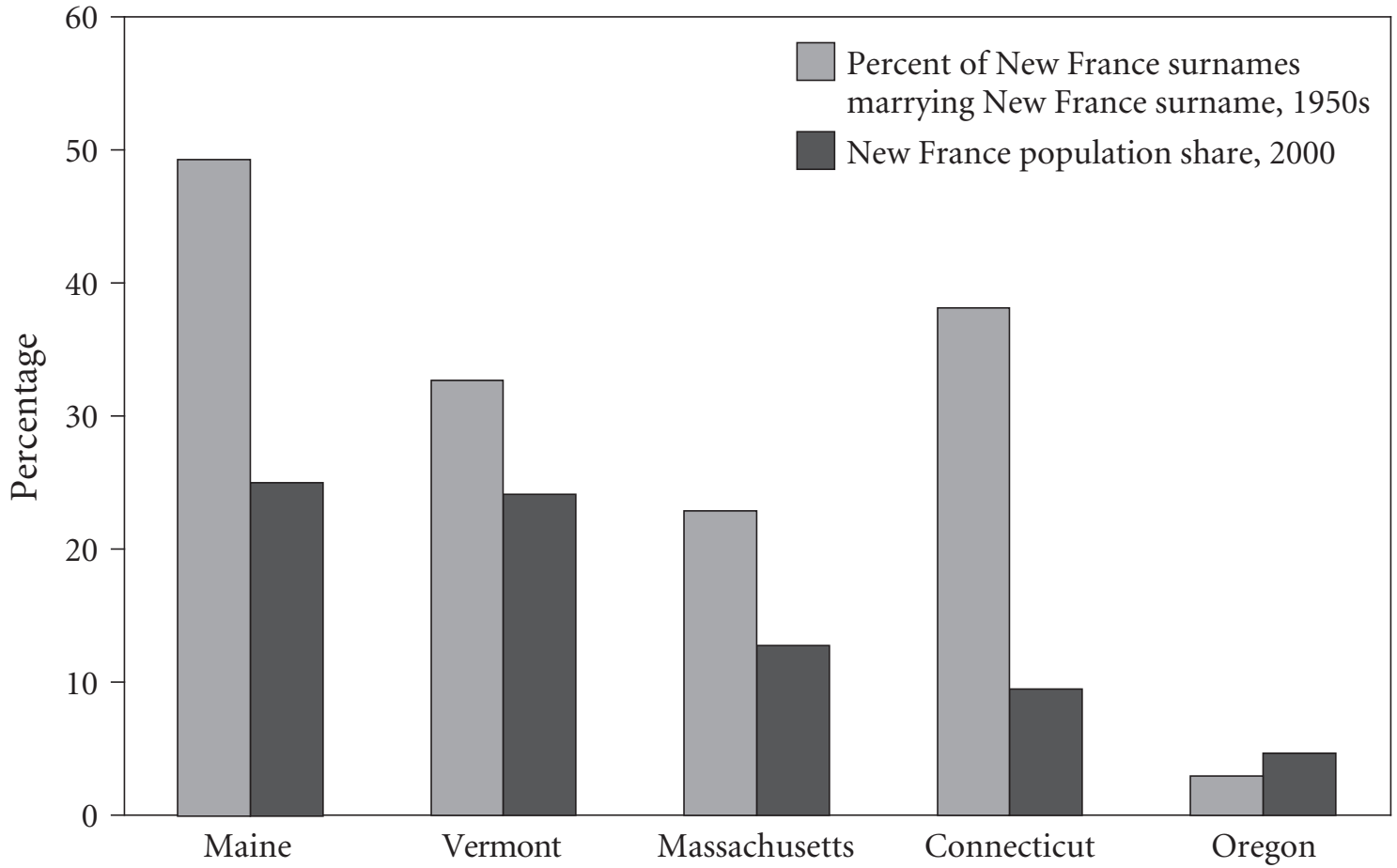
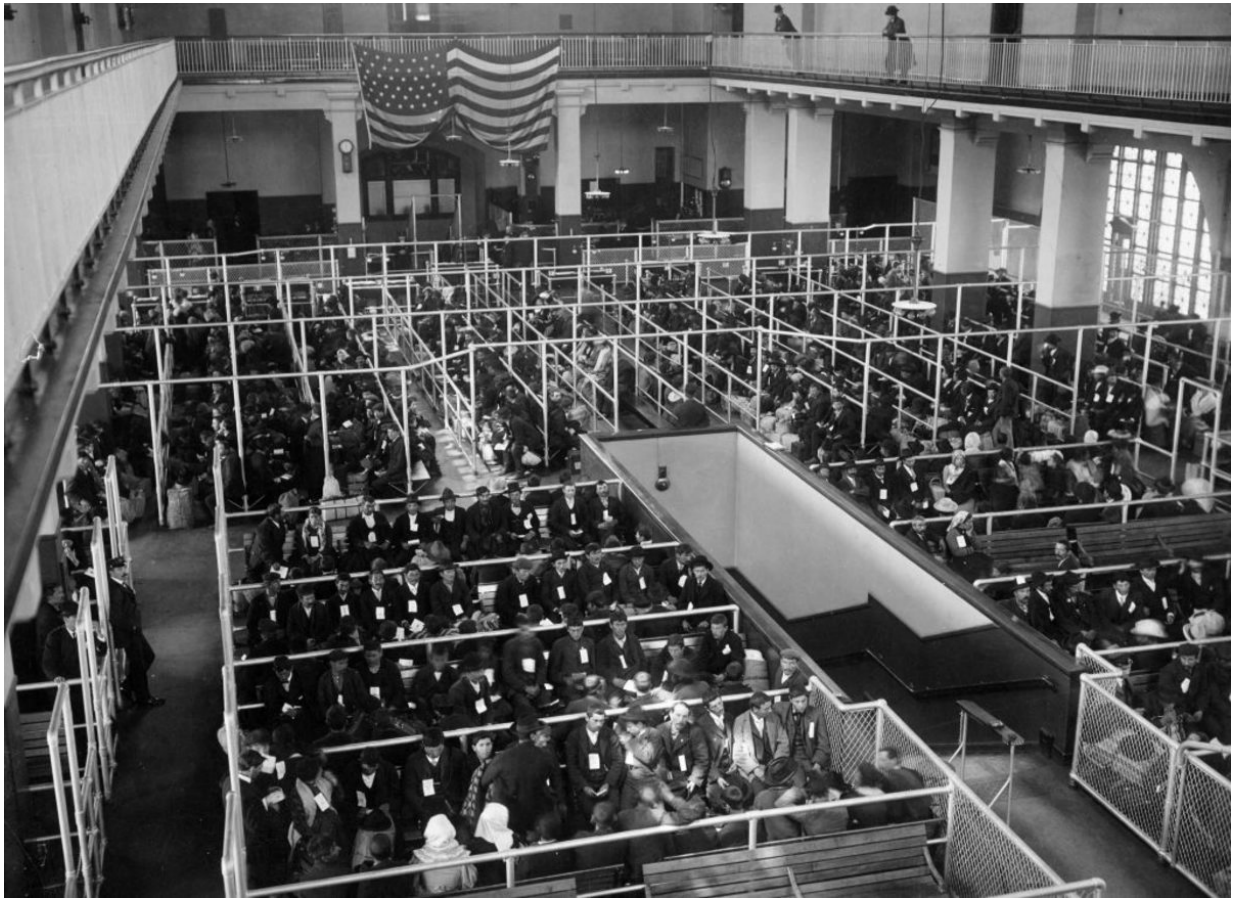


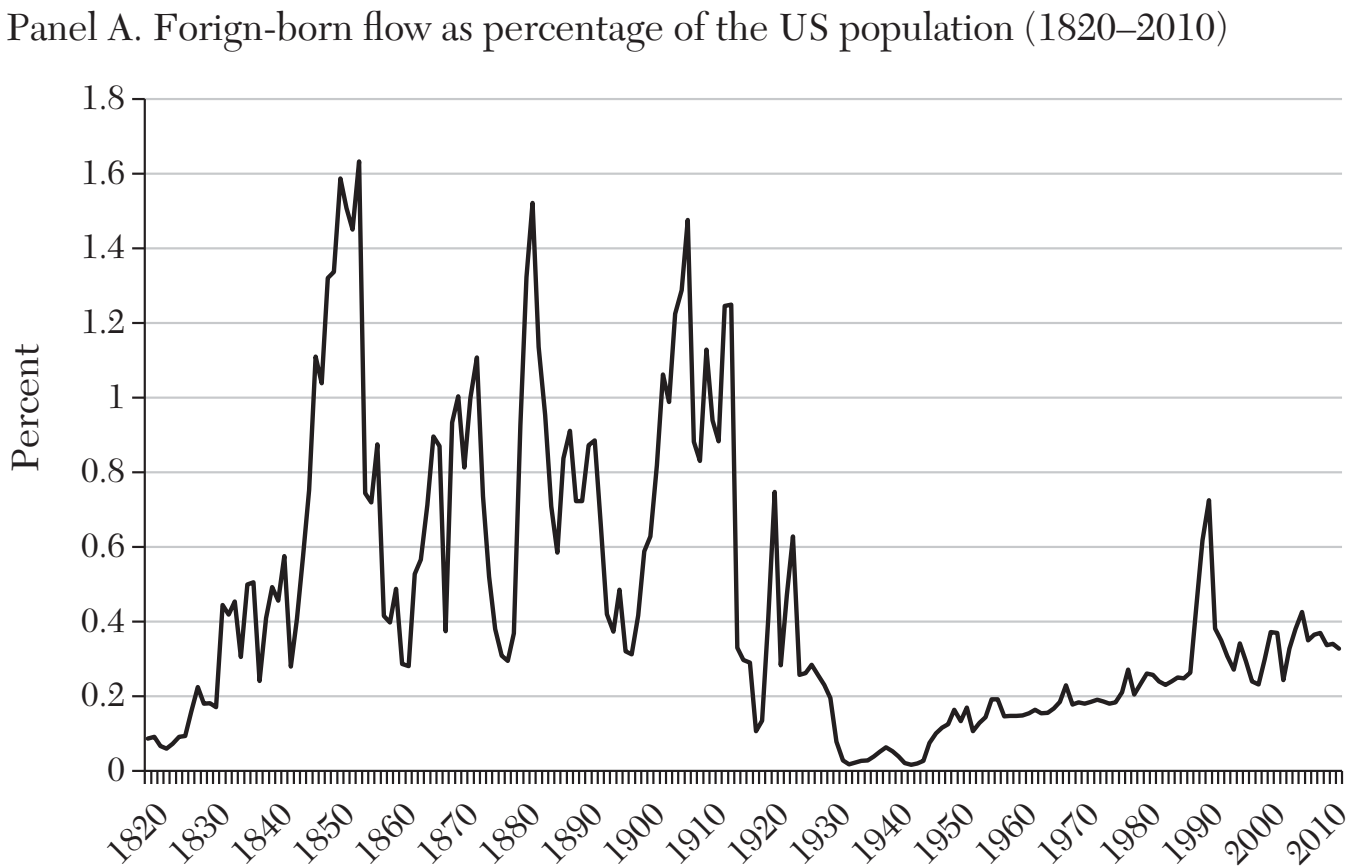
FIGURE 3.12. Marital endogamy among New France descendants, 1950s.

Immigration and Mobility

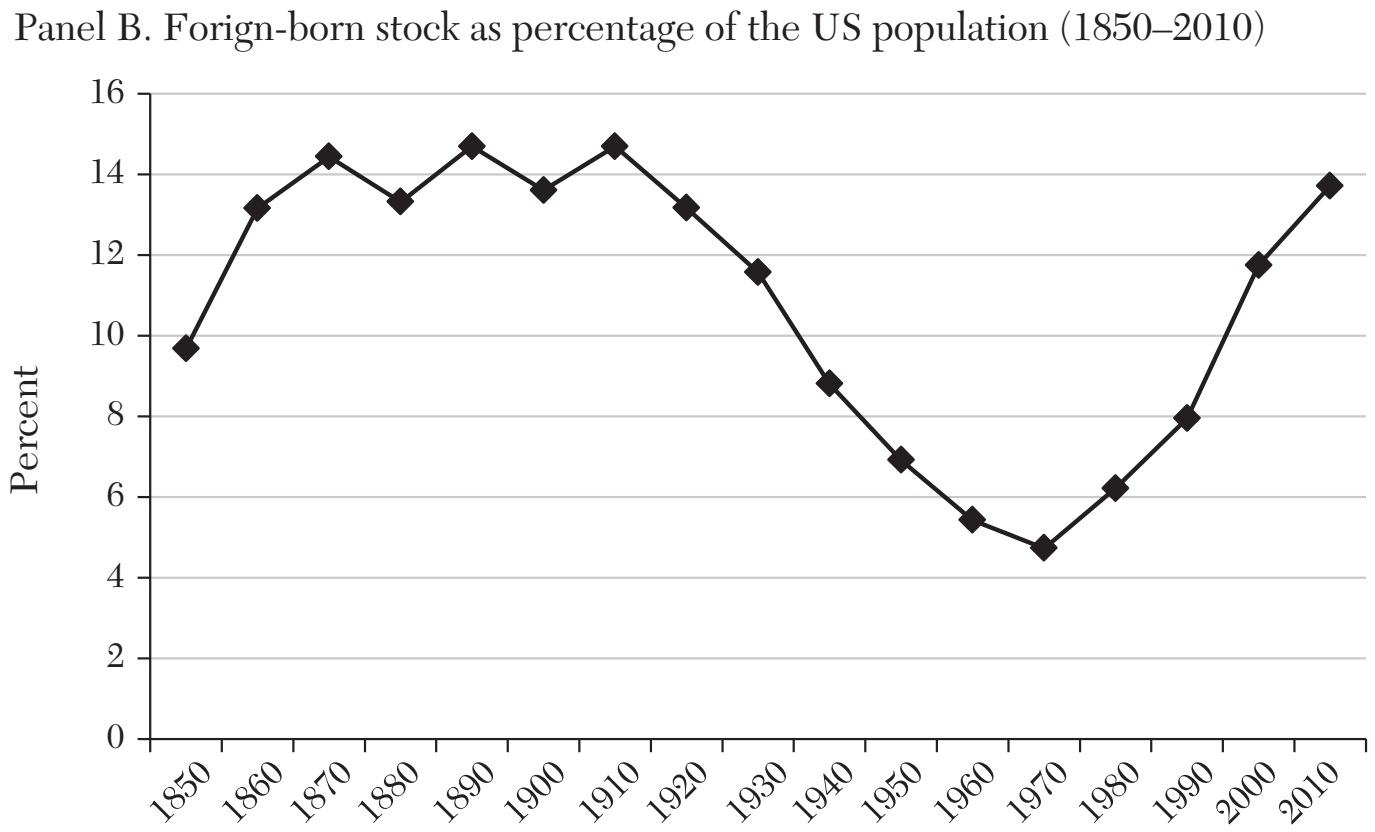


Immigration and Mobility

Immigration Over Time



Immigration Over Time



Immigration Over Time

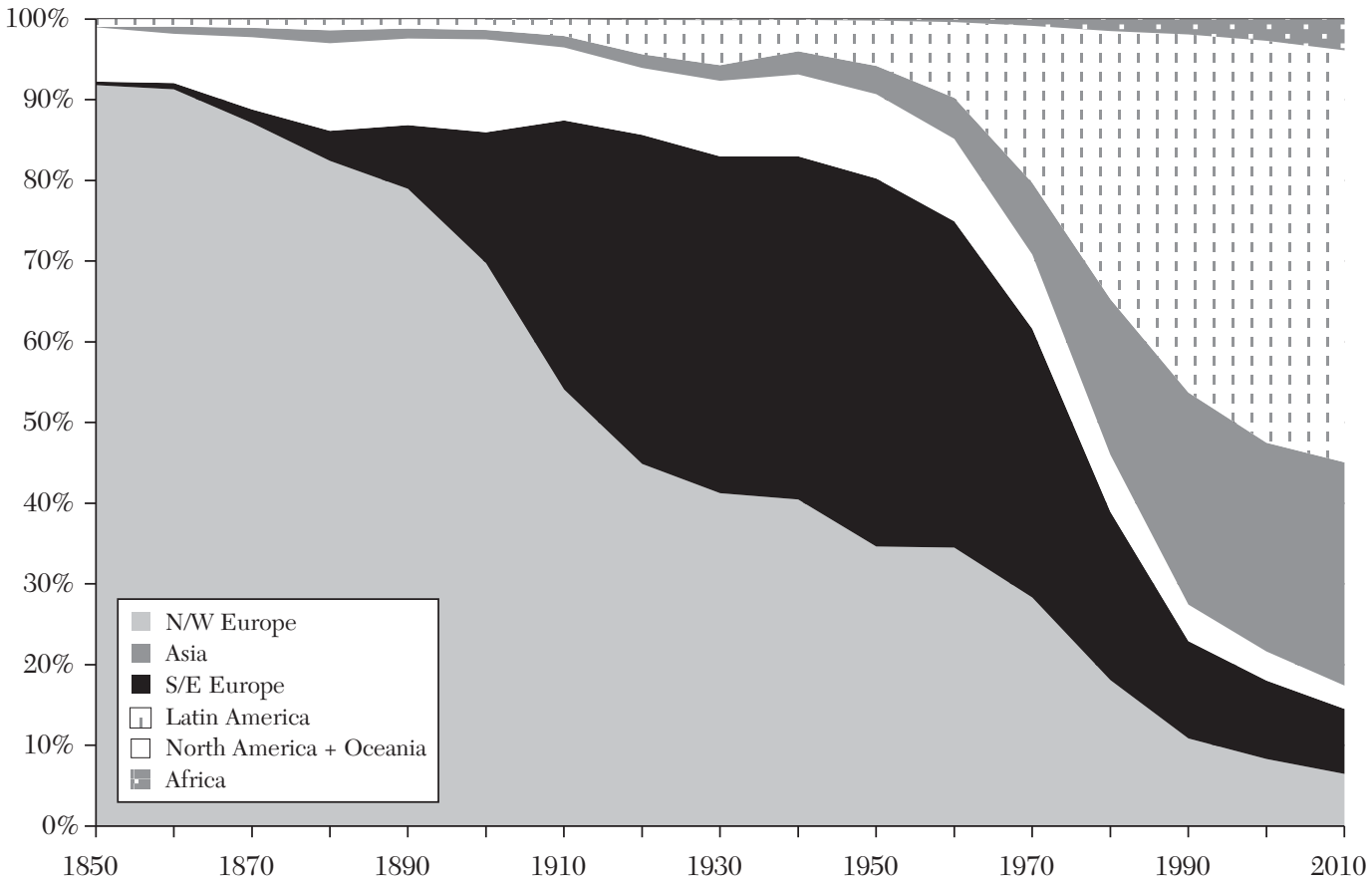


Figure 2. Sending Regions within the Foreign-Born Population, 1850–2010

Immigration Over Time

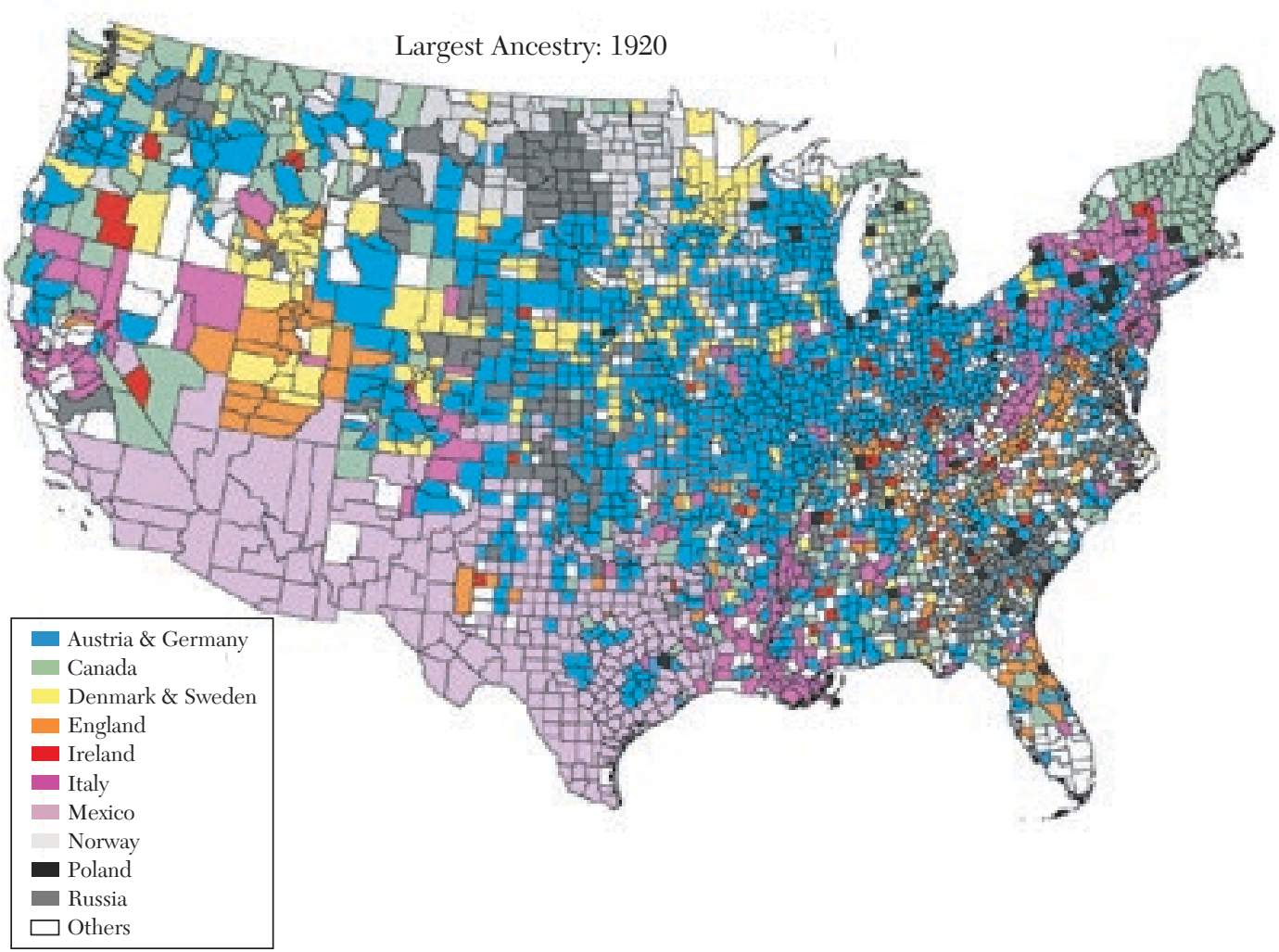
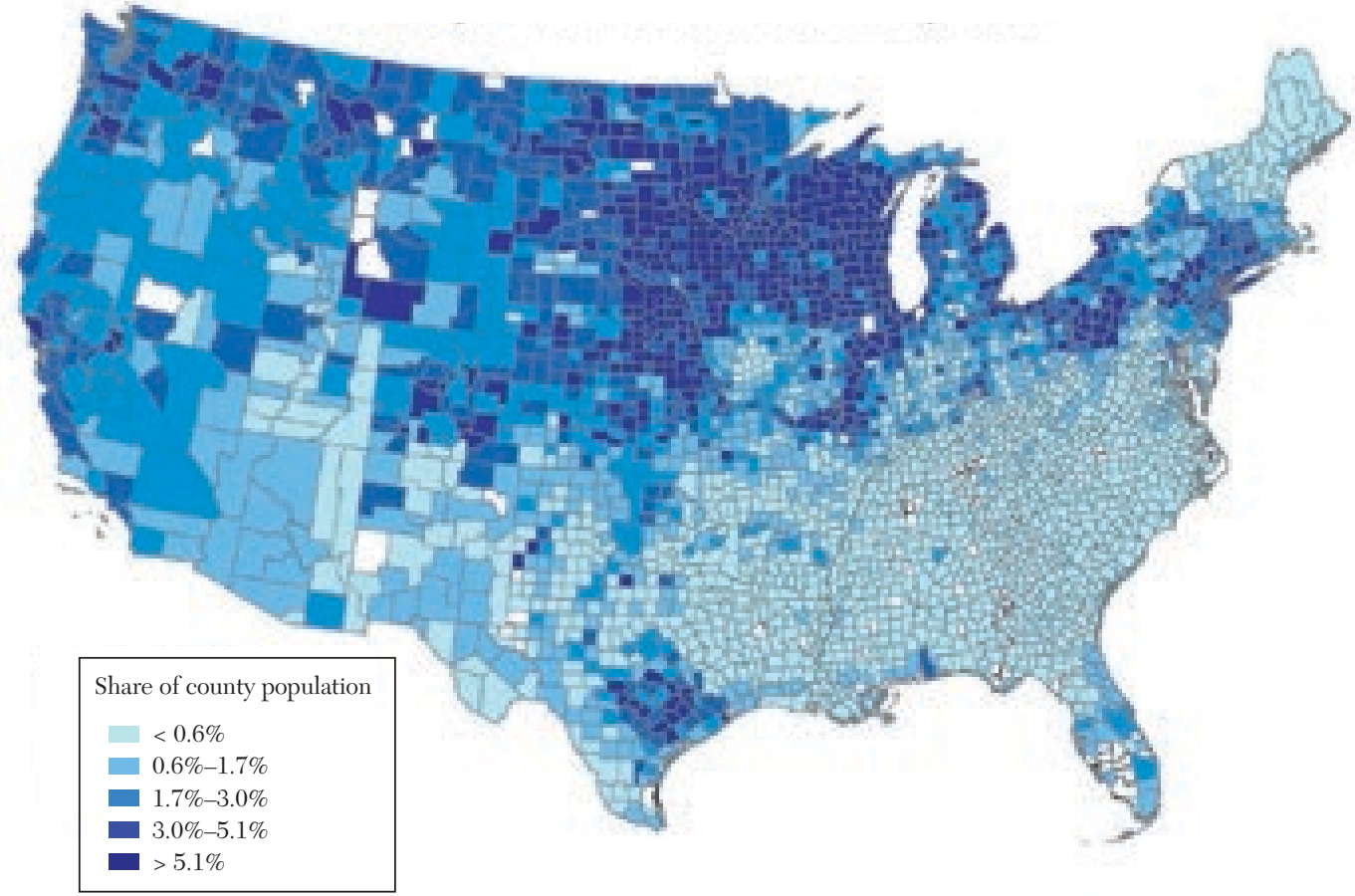


Figure 3. Largest Country-of-Origin Group among Foreign Born by County, 1920

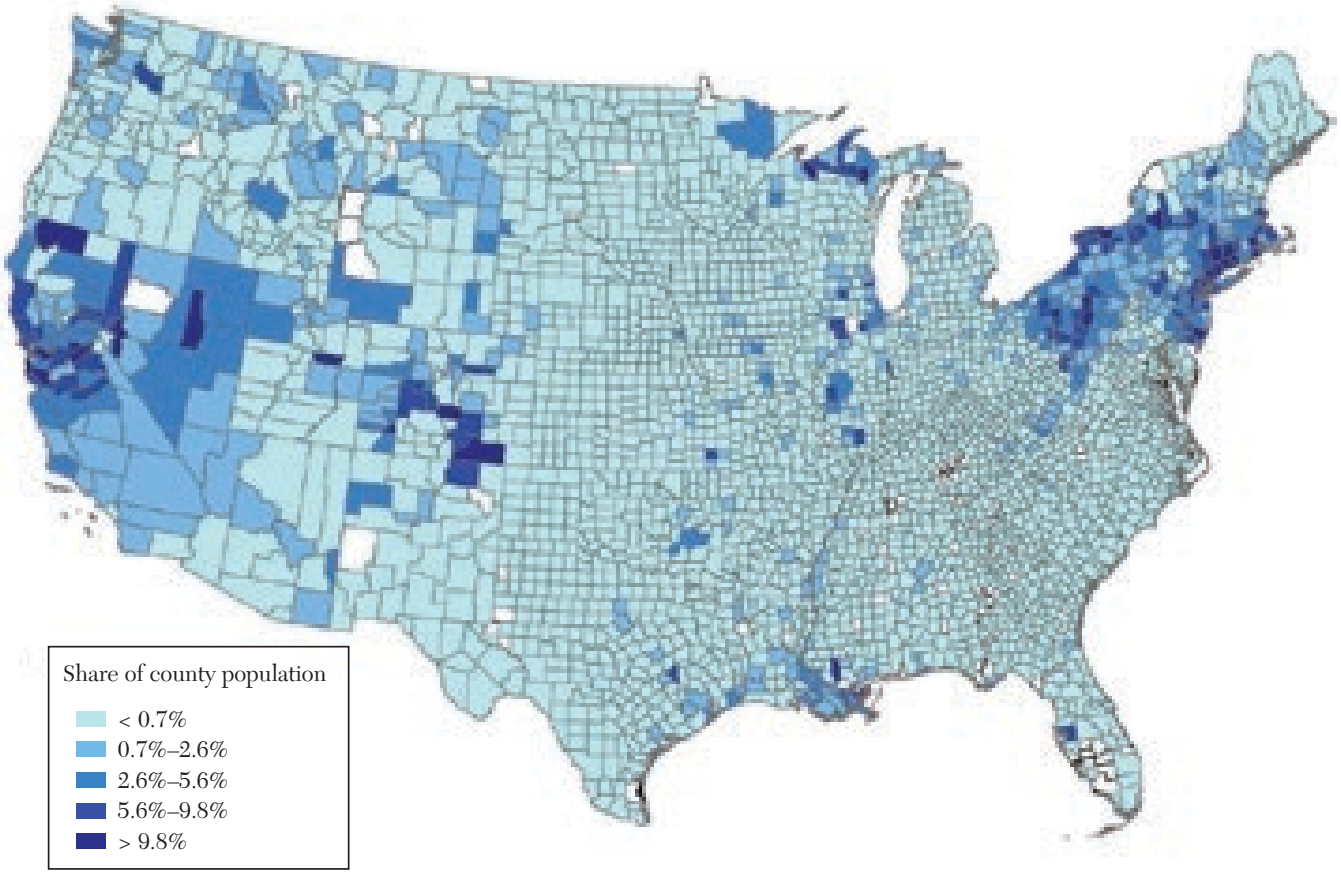
Immigration Over Time

Panel A. Ancestry Share: Austria and Germany, 1920



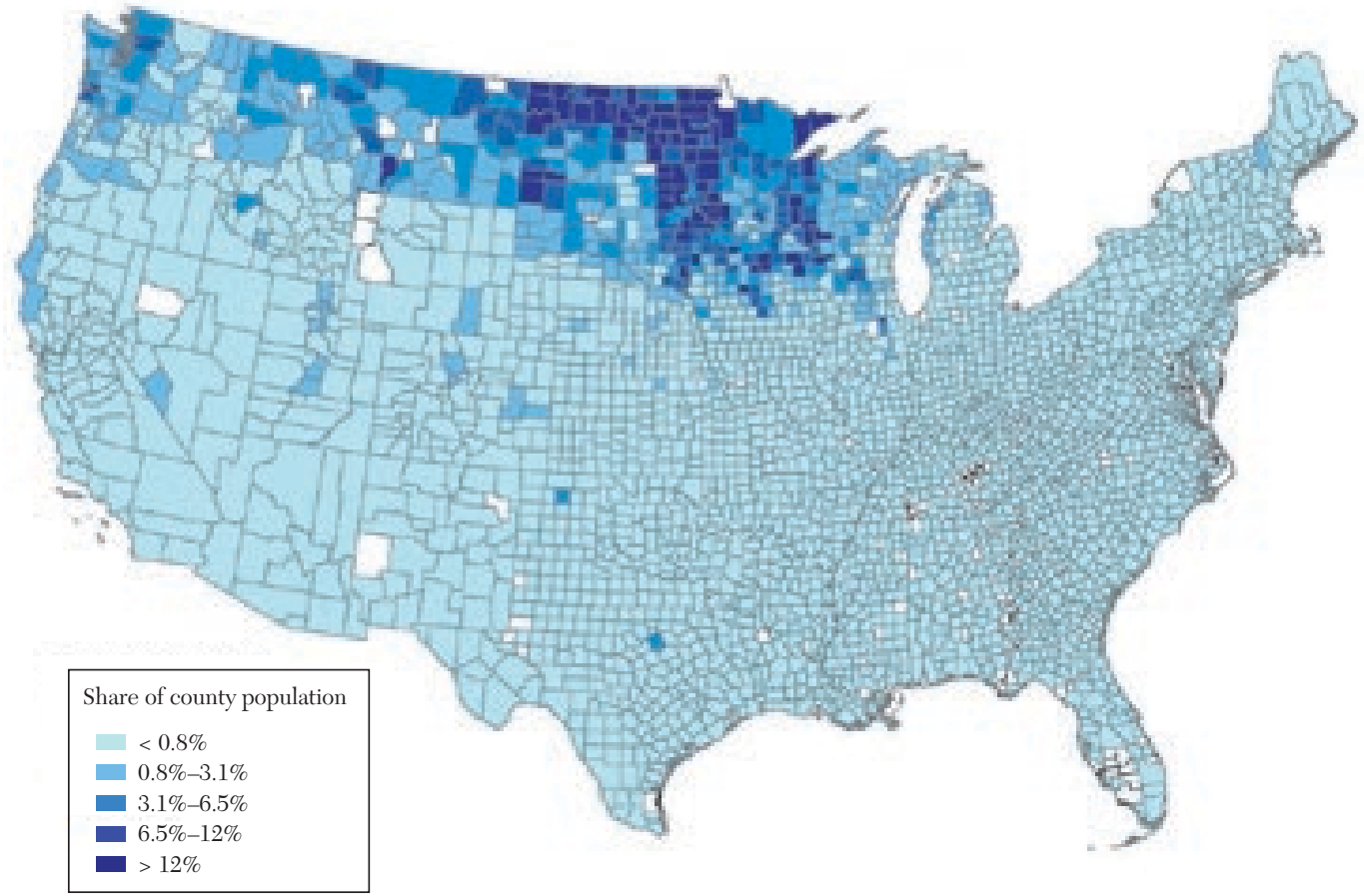
Immigration Over Time

Panel B. Ancestry Share: Italy, 1920

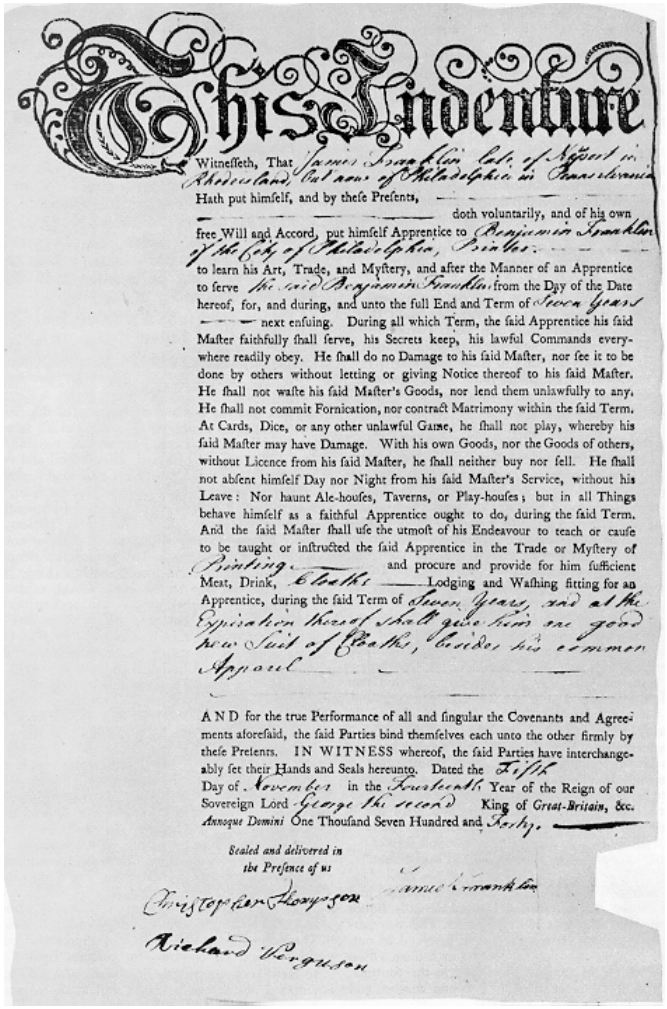


Immigration Over Time

Panel C. Ancestry Share: Norway, 1920



Immigration in the Colonial Period



This Indenture

Witneſſeth, That James Franklin late of Newport in Rhode-Island, but now of Philadelphia in Pennſylvania
Hath put himſelf, and by theſe Preſents, ——— doth voluntarily, and of his own free Will and Accord, put himſelf Apprentice to Benjamin Franklin of the City of Philadelphia, Printer: ——— to learn his Art, Trade, and Myſtery, and after the Manner of an Apprentice to ſerve the ſaid Benjamin Franklin from the Day of the Date hereof, for, and during, and unto the full End and Term of Seven Years ——— next enſuing. During all which Term, the ſaid Apprentice his ſaid Maſter faithfully ſhall ſerve, his Secrets keep, his lawful Commands everywhere readily obey. He ſhall do no Damage to his ſaid Maſter, nor ſee it to be done by others without letting or giving Notice thereof to his ſaid Maſter. He ſhall not waſte his ſaid Maſter's Goods, nor lend them unlawfully to any. He ſhall not commit Fornication, nor contract Matrimony within the ſaid Term. At Cards, Dice, or any other unlawful Game, he ſhall not play, whereby his ſaid Maſter may have Damage. With his own Goods, nor the Goods of others, without Licence from his ſaid Maſter, he ſhall neither buy nor ſell. He ſhall not abſent himſelf Day nor Night from his ſaid Maſter's Service, without his Leave: Nor haunt Ale-houſes, Taverns, or Play-houſes; but in all Things behave himſelf as a faithful Apprentice ought to do, during the ſaid Term. And the ſaid Maſter ſhall uſe the utmoſt of his Endeavour to teach or cauſe to be taught or inſtructed the ſaid Apprentice in the Trade or Myſtery of Printing ——— and procure and provide for him ſufficient Meat, Drink, Cloaths ——— Lodging and Waſhing fitting for an Apprentice, during the ſaid Term of Seven Years, and at the Expiration thereof ſhall give him one good new Suit of Cloaths, conſider his common Apparel ———

AND for the true Performance of all and ſingular the Covenants and Agreements aforeſaid, the ſaid Parties bind themſelves each unto the other firmly by theſe Preſents. IN WITNESS whereof, the ſaid Parties have interchangeably ſet their Hands and Seals hereunto. Dated the 24th Day of November in the Fourteenth Year of the Reign of our Sovereign Lord George the ſecond King of Great-Britain, &c. Annoque Domini One Thouſand Seven Hundred and Forty. ———

Sealed and delivered in the Preſence of us

Chriſtopher Grayſon

Richard Ferguſon

James Franklin

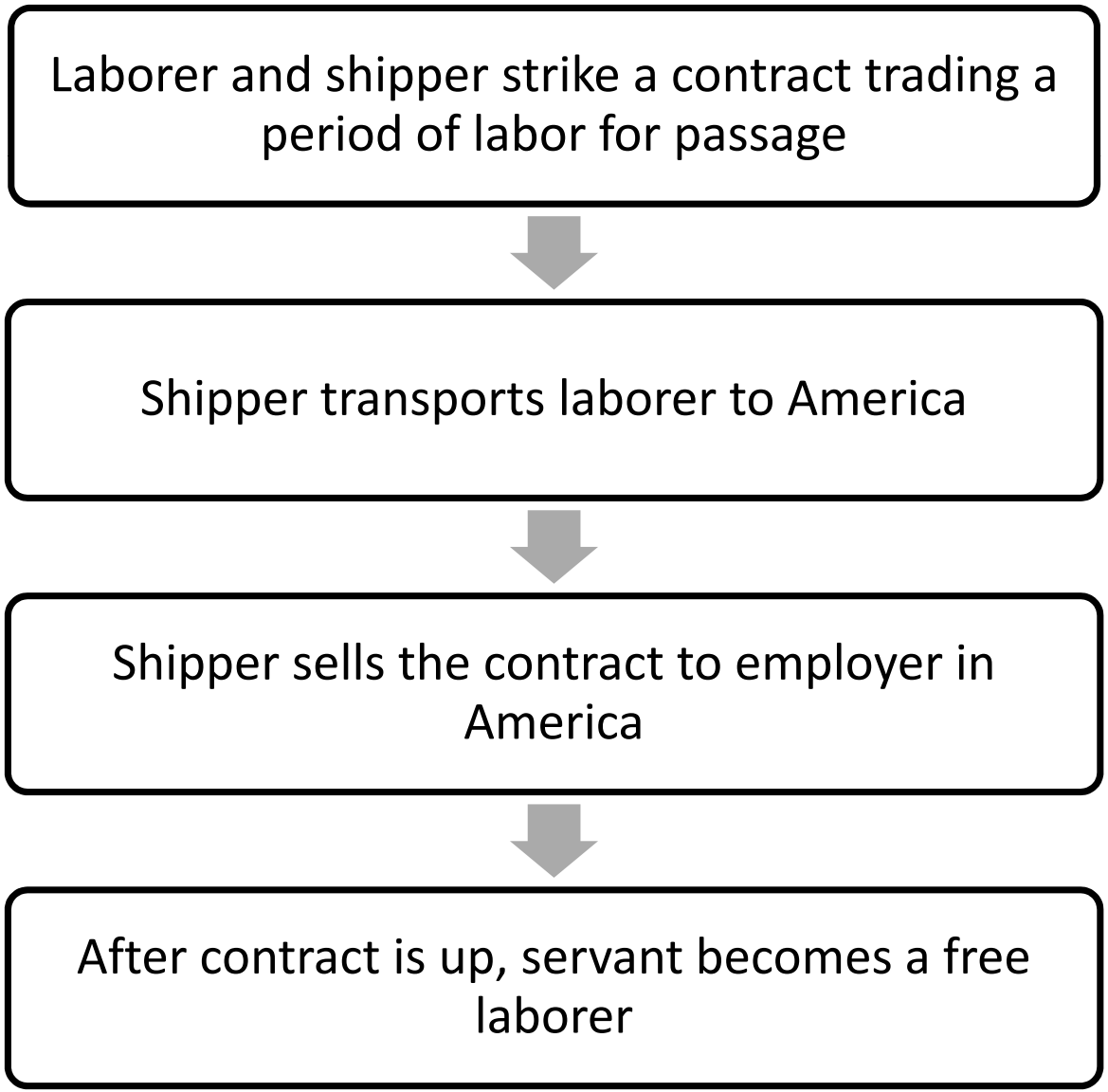
Immigration in the Colonial Period

This indentured...between [Alexander Beard]...of the one part, and [John Dickey]...of the other part, witnesseth, that the said [Alexander Beard] doth hereby covenant, promise and grant, to ...[John Dickey]...and his assigns, from the day of the date hereof until the first and next arrival at [Philadelphia] in America...and during the term of [three] years to serve in such service and employment as the said [John Dickey] or [his] assigns shall there employ [him]...In consideration whereof the said [John Dickey] doth grant...to pay for [his] passage, and to find allow [him] meat, drink, apparel and lodging, with other necessaries, during the said term; and at the end of the said term to pay unto him the usual allowance, according to the custom of the country in the like kind...

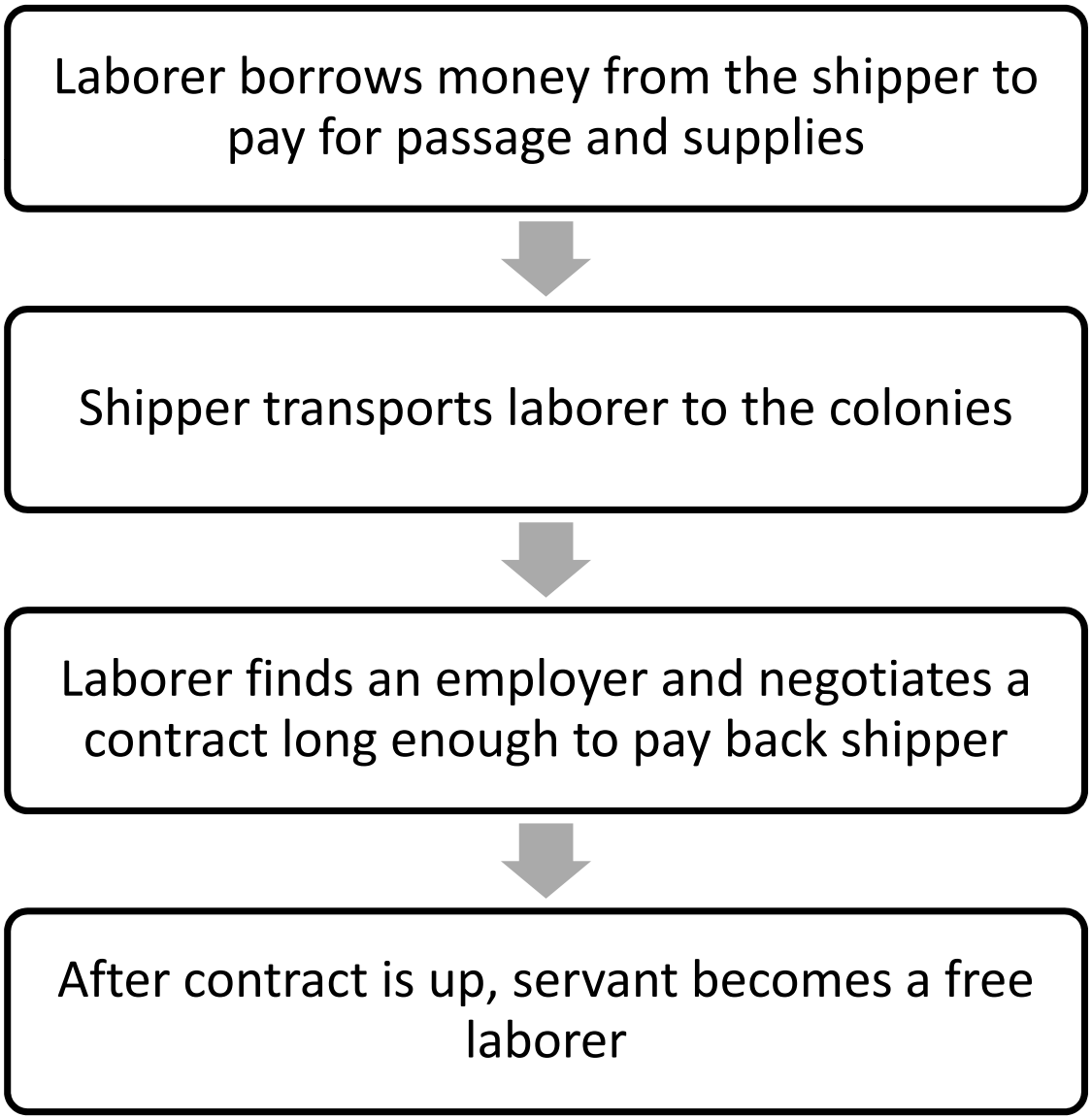
Credit Constraints and Indentured Servitude

- The cost of passage to America was £5 to £10, an amount greater than average annual income at the time
- To put that in perspective, think about college tuition:
  - Average tuition and fees at private four-year colleges is $32,410 (according to the College Board)
  - Median income for a 20 to 24 year old is $26,728 (according to the BLS)
  - If there were no student loans, how would people pay for college?

How Indentured Servitude Works - Standard



How Indentured Servitude Works - Redemptioners



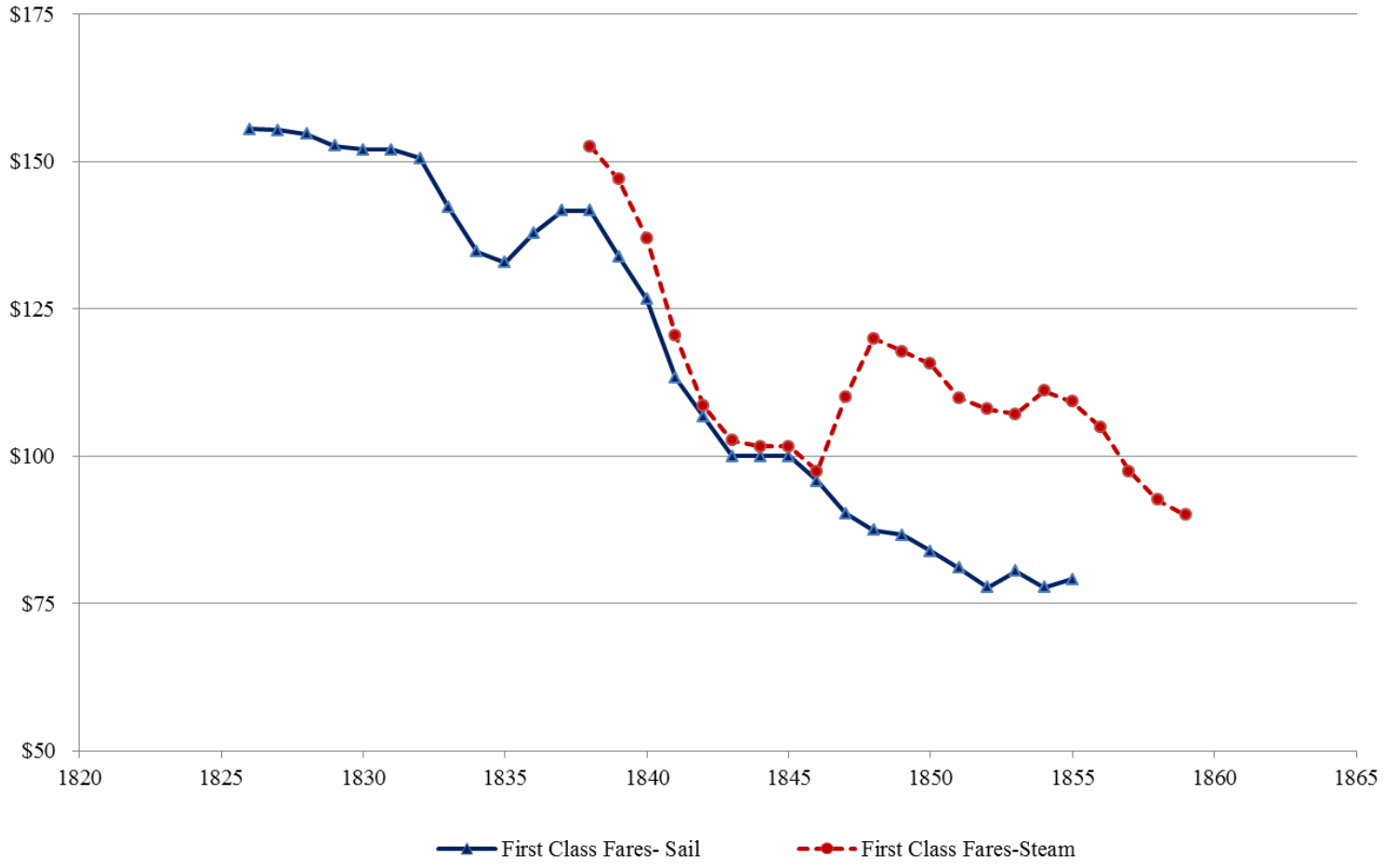
Migration and Human Capital During the Colonial Period

Contract Length and Servant Characteristics

Characteristic	Months More or Less Service
15 years old	26
17 years old	9
19 years old	2
Female	-2
Literate	-1
Farmer	-4
Metalworker	-4
Textile worker	-4

Migration in the Nineteenth Century

Figure 8
Advertised First Class Fares on Sailing and Steam Ships
3 year averages 1826-1859



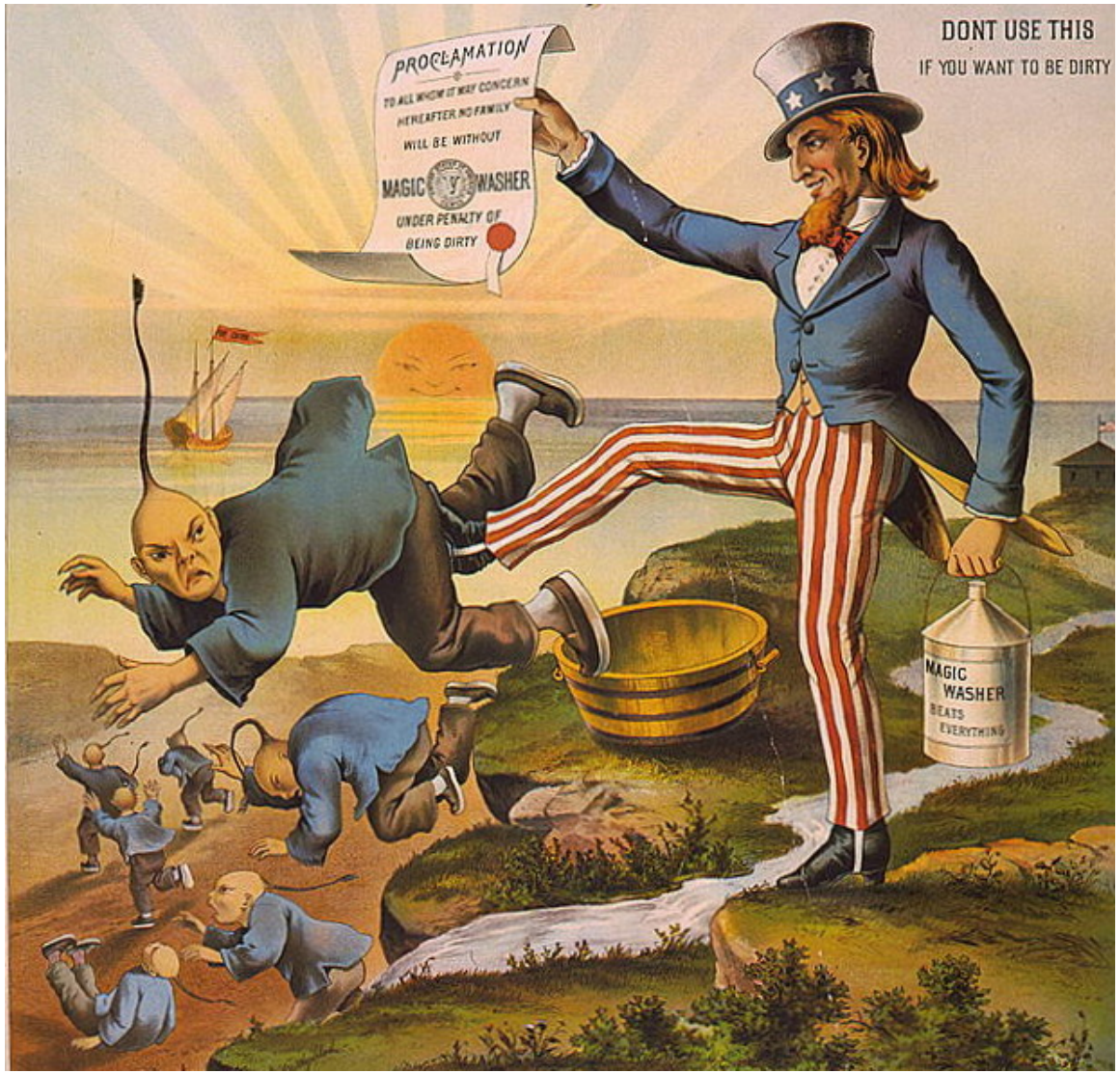
Migration in the Nineteenth Century

Migration in the Nineteenth Century

Migration in the Nineteenth Century

Chinese Exclusion Act - 1882

Chinese Exclusion Act - 1882

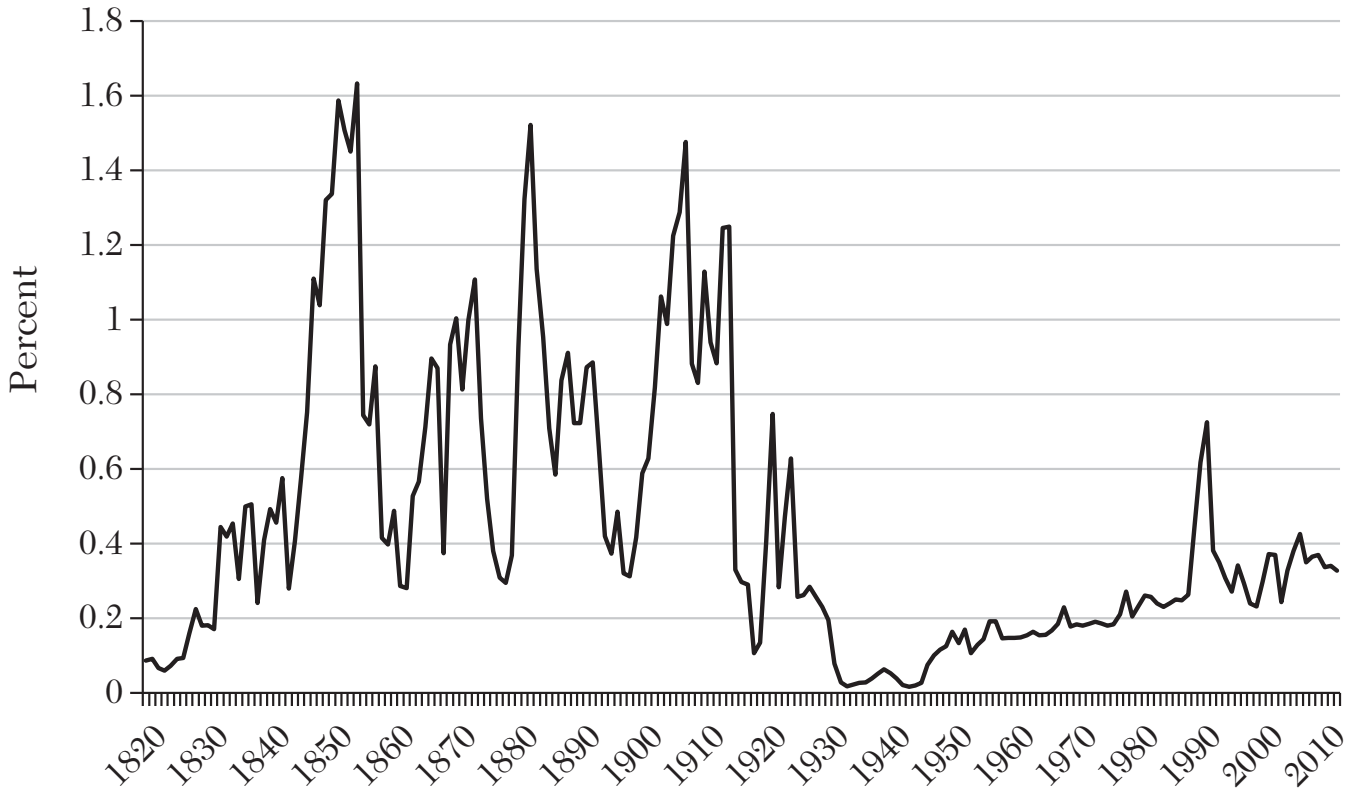


Announcements

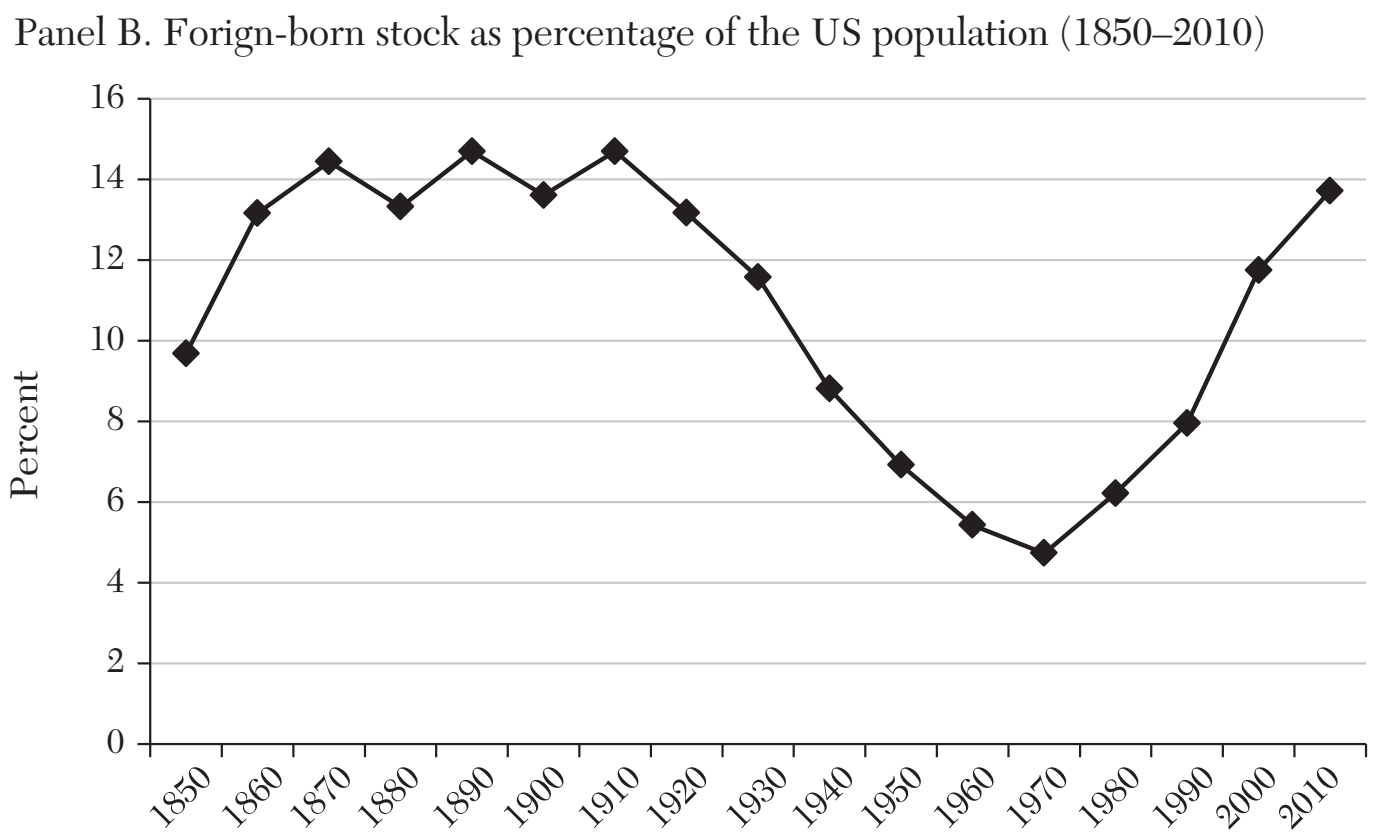
- Grades will be posted tomorrow for your referee reports, I will email each of you comments on your report
- Remember that the second referee report is due March 15th on Miller (2008) "Women's Suffrage, Political Responsiveness, and Child Survival in American History"
- This is a different style of paper, be certain to think about whether they are proving their hypothesis or whether there are alternative explanations
- Also think about making progress on your data projects, feel free to email me if you run into any problems
- No class on Friday, have a great Spring Break

Immigration Over Time

Panel A. Forign-born flow as percentage of the US population (1820–2010)



Immigration Over Time



Migration in the Nineteenth Century

Chinese Exclusion Act - 1882

Chinese Exclusion Act - 1882

Industrial Distribution of Chinese Employment by Region, 1870-1930

	Total	Northeast	Midwest	South	West
		1870			
Restaurants	0.2	--	--	--	0.2
Laundries	11.0	--	--	--	11.0
Food stores	1.3	--	--	--	1.3
All else	87.5	--	--	--	87.5
		1880			
Restaurants	0.4	0.0	0.0	0.0	0.4
Laundries	13.9	100.0	100.0	0.0	11.9
Food stores	1.6	0.0	0.0	0.0	1.6
All else	84.1	0.0	0.0	100.0	86.1
		1900			
Restaurants	0.5	0.0	0.0	0.0	0.6
Laundries	34.1	81.6	100.0	85.7	16.9
Food stores	3.4	0.0	0.0	7.1	4.2
All else	62.0	18.4	0.0	7.3	78.3
		1910			
Restaurants	7.4	9.2	11.1	8.5	5.6
Laundries	20.9	60.2	66.7	31.9	7.4
Food stores	6.2	0.0	0.0	31.9	12.7
All else	65.5	30.6	22.2	27.7	74.3
		1920			
Restaurants	17.3	32.8	47.8	22.2	12.2
Laundries	22.1	55.5	39.1	44.4	11.5
Food stores	7.4	5.5	0.0	7.4	7.0
All else	53.2	6.2	13.1	26.0	69.3
		1930			
Restaurants	27.7	42.0	32.4	34.4	15.6
Laundries	24.7	42.0	50.0	21.9	6.6
Food stores	8.8	0.6	0.0	28.1	13.2
All else	38.8	15.4	17.6	15.6	64.6

Chinese Exclusion Act - 1882

Distribution of the Chinese-American Population, 1870-1960

Year	Percentage of counties with one or more Chinese residents	Median number of Chinese residents in a county with Chinese residents
1870	10.9	19
1880	18.8	1
1890	37.8	4
1900	45.7	4
1910	40.8	5
1920	44.9	4
--	--	--
1960	42	7

Chinese Exclusion Act - 1882

Dissimilarity in Residence by State		
	Chinese vs Non-Chinese	Black vs Non-Black
1870	97.3	66.6
1880	93.6	66.6
1890	85.4	--
1900	70	66.2
1910	65.7	67.7
1920	55.6	49.1
1930	57.4	54.5

The Immigration Act of 1917

Sec. 3. That the following classes of aliens shall be excluded from admission into the United States: All idiots, imbeciles, feeble-minded persons, epileptics, insane persons...persons of constitutional psychopathic inferiority; persons with chronic alcoholism; paupers; professional beggars; vagrants; persons afflicted with tuberculosis...

The Immigration Act of 1917

...persons who have been convicted of or admit having committed a felony or other crime or misdemeanor involving moral turpitude; polygamists; anarchists...[persons] who advocate or teach unlawful destruction of property; ...persons coming to the United States for the purpose of prostitution or for any other immoral purpose...

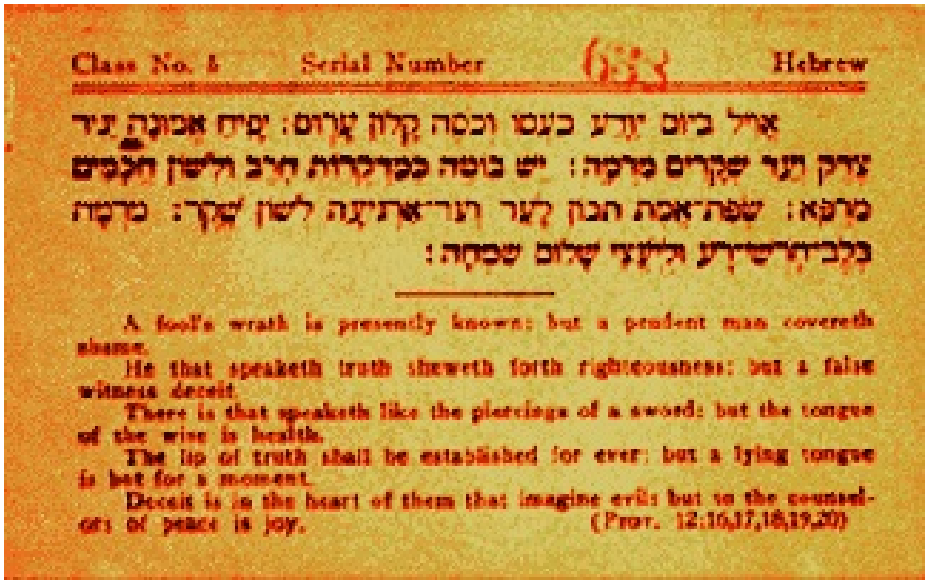
The Immigration Act of 1917

...[The provision] shall not apply to the persons of the following status or occupations: Government officers, ministers or religious teachers, missionaries, lawyers, physicians, chemists, civil engineers, teachers, students, authors, artists, merchants, and travelers for curiosity or pleasure...

The Immigration Act of 1917

All aliens over sixteen years of age, physically capable of reading, who can not read the English language, or some other language or dialect, including Hebrew or Yiddish...That for the purpose of ascertaining whether aliens can read the immigrant inspectors shall be furnished with slips of uniform size...each containing not less than thirty nor more than forty words in ordinary use, printed in plainly legible type of some one of the various languages or dialects of immigrants.

The Immigration Act of 1917


Class No. 4 Serial Number 6[illegible] Hebrew

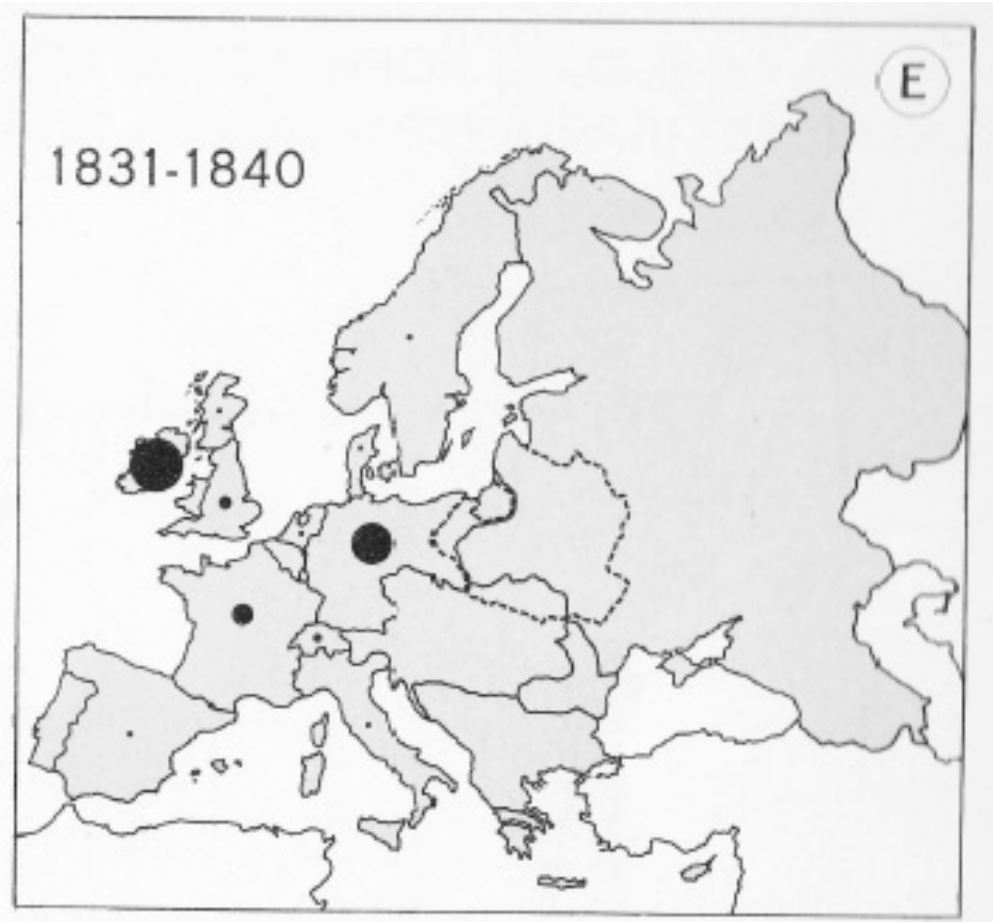
אויל ביום יודע כעסו וכסה קלון ערום: יפיח אמונה יגיד
צדק ועד שקרים מרמה: יש בוטה כמדקרות חרב ולשון חכמים
מרפא: שפת־אמת תכון לעד ועד־אחת־ענה לשון שקר: מרמה
בלב־חרשי־רע ויועצי שלום שמחה:

A fool's wrath is presently known: but a prudent man covereth shame.
He that speaketh truth sheweth forth righteousness: but a false witness deceit.
There is that speaketh like the piercings of a sword: but the tongue of the wise is health.
The lip of truth shall be established for ever: but a lying tongue is but for a moment.
Deceit is in the heart of them that imagine evil: but to the counsellors of peace is joy. (Prov. 12:16,17,18,19,20)

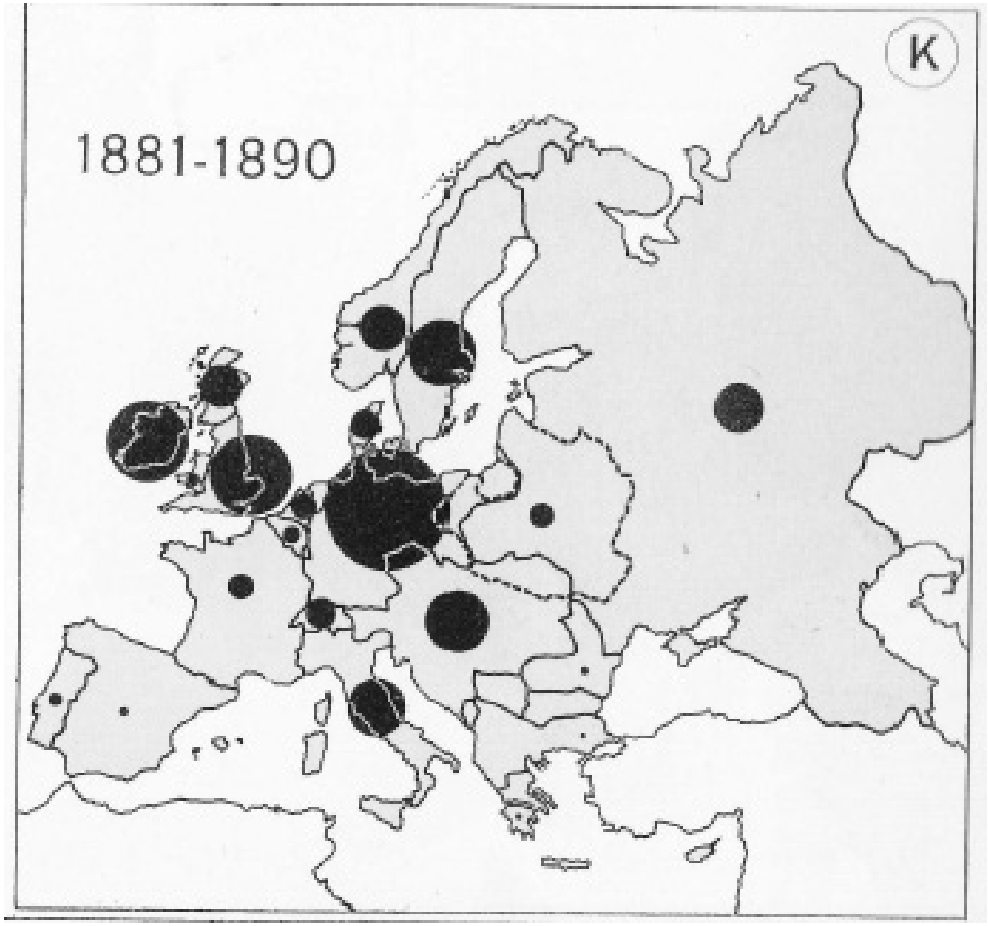
1917 - Literacy Test

Quota Act and National Origins Act - 1920s

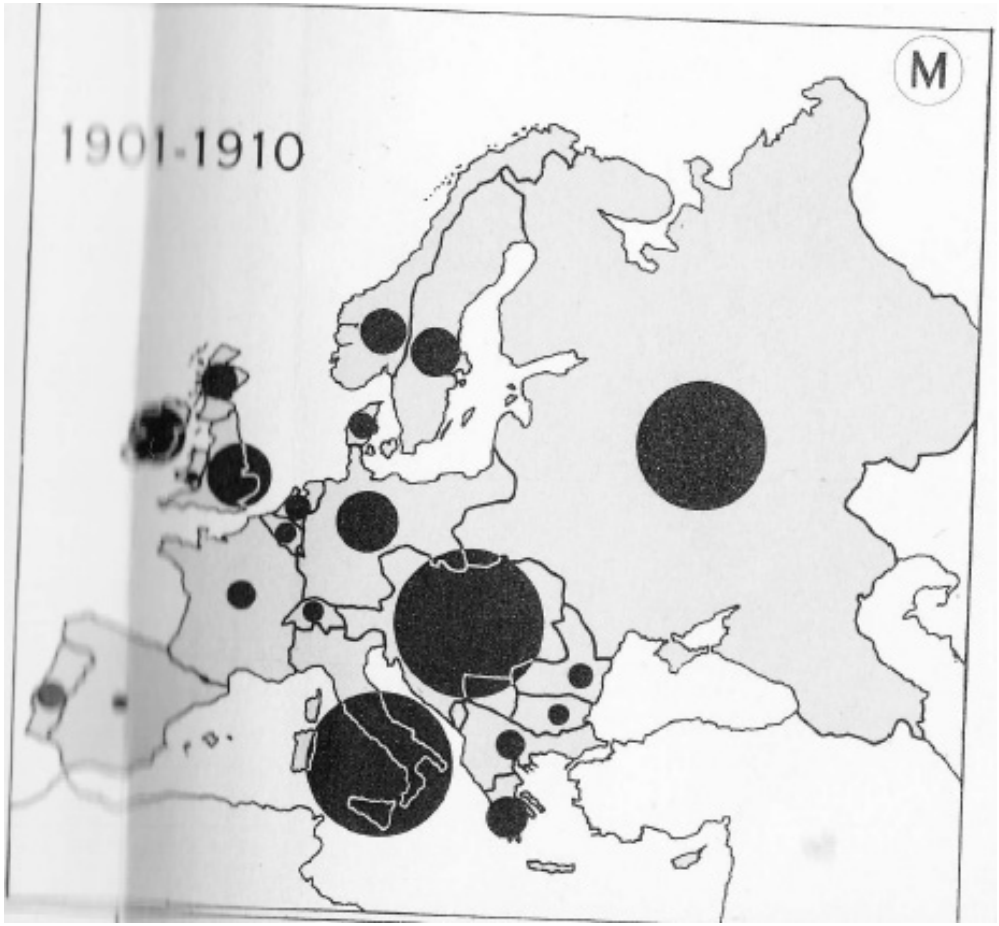
Quota Act and National Origins Act - 1920s



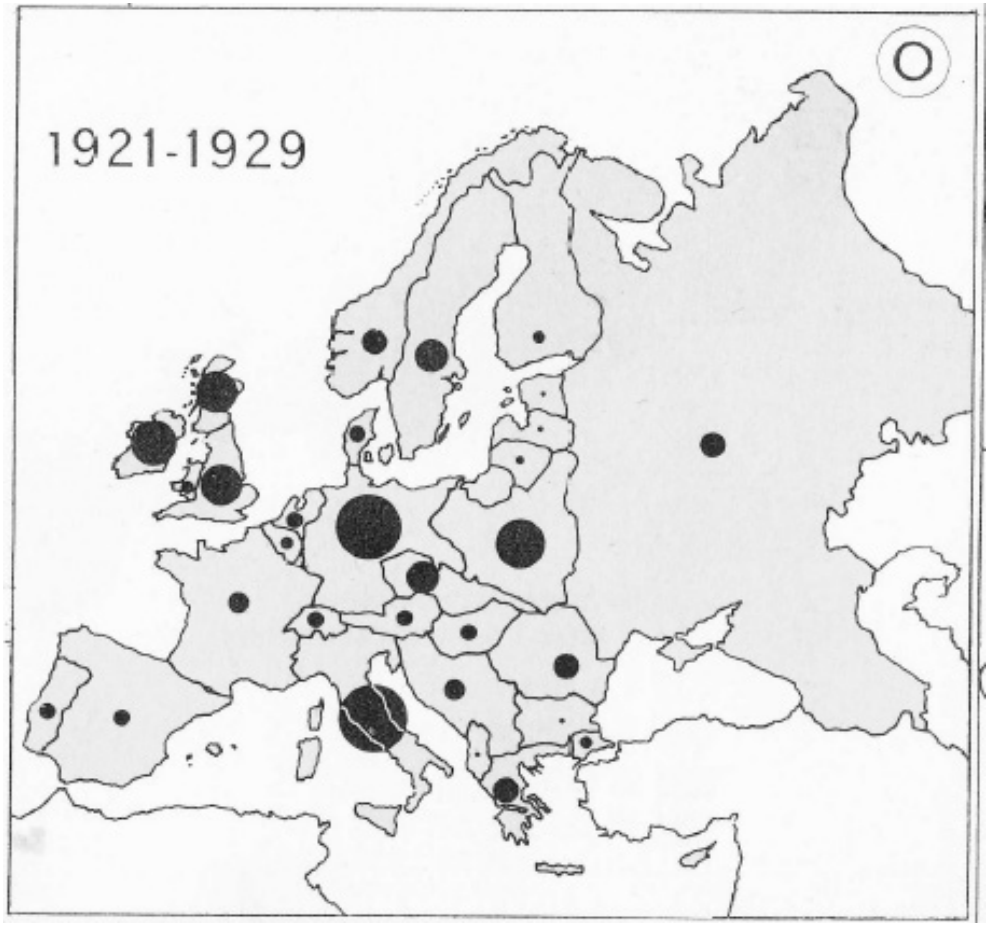
Quota Act and National Origins Act - 1920s



Quota Act and National Origins Act - 1920s



Quota Act and National Origins Act - 1920s

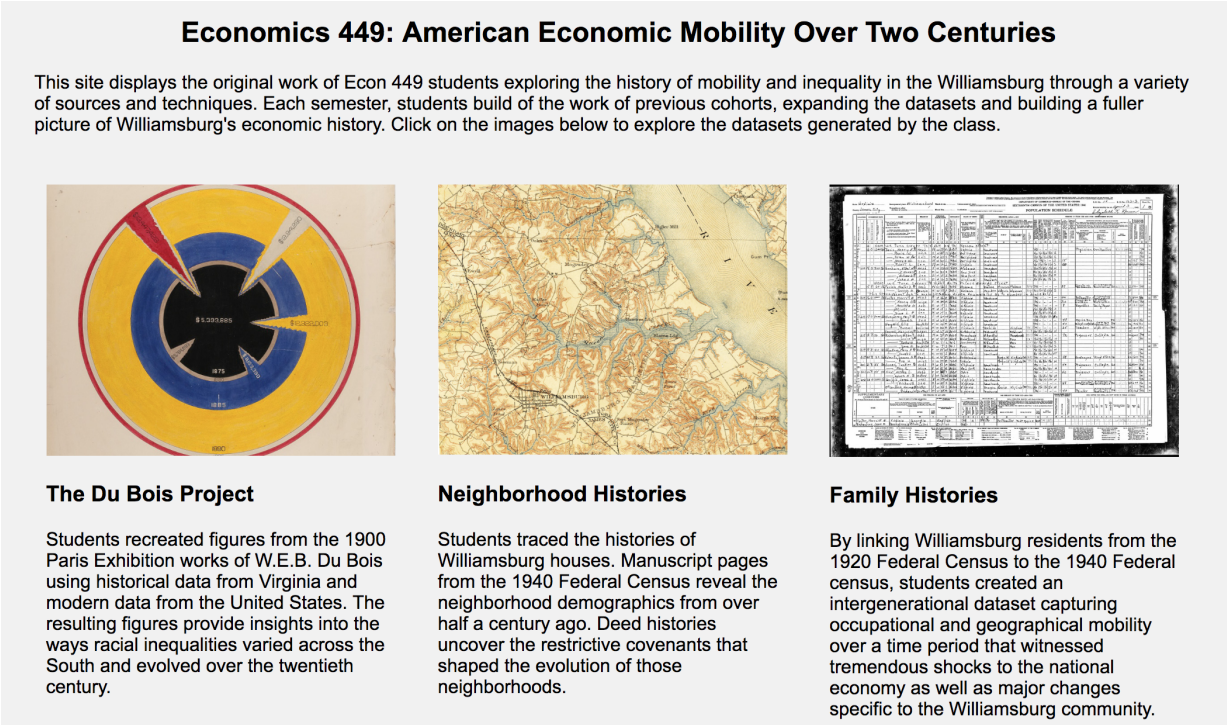


Immigration and Nationality Act - 1965

Announcements

- You should have gotten an email with my feedback on your first referee report, let me know if you didn't
- Remember that the second referee report is due March 15th on Miller (2008) "Women's Suffrage, Political Responsiveness, and Child Survival in American History"
- This is a different style of paper, be certain to think about whether they are proving their hypothesis or whether there are alternative explanations
- It's also time to get moving on your data projects, feel free to email me if you run into any problems
- On Wednesday we'll talk about your final projects

Announcements



Economics 449: American Economic Mobility Over Two Centuries

This site displays the original work of Econ 449 students exploring the history of mobility and inequality in the Williamsburg through a variety of sources and techniques. Each semester, students build of the work of previous cohorts, expanding the datasets and building a fuller picture of Williamsburg's economic history. Click on the images below to explore the datasets generated by the class.

The Du Bois Project

Students recreated figures from the 1900 Paris Exhibition works of W.E.B. Du Bois using historical data from Virginia and modern data from the United States. The resulting figures provide insights into the ways racial inequalities varied across the South and evolved over the twentieth century.

Neighborhood Histories

Students traced the histories of Williamsburg houses. Manuscript pages from the 1940 Federal Census reveal the neighborhood demographics from over half a century ago. Deed histories uncover the restrictive covenants that shaped the evolution of those neighborhoods.

Family Histories

By linking Williamsburg residents from the 1920 Federal Census to the 1940 Federal census, students created an intergenerational dataset capturing occupational and geographical mobility over a time period that witnessed tremendous shocks to the national economy as well as major changes specific to the Williamsburg community.

The beta version of our class website is up (with your data still to come). Suggestions are welcome. Let me know if you want your name included with your figure.

Selection Into Migration

- The outcomes of immigrants will depend crucially on which individuals decide to migrate
- In particular, it will depend on whether there is *positive* or *negative* selection into migration
- Economists have developed a theory of selection into migration that has its roots in Roy's 1951 article "Some Thoughts on the Distribution of Earnings"
- The keywords for the article: hunting, rabbits, fishers, occupations, productivity, trout, logarithms, communities, industrial productivity, relative prices

Selection Into Migration

- While Roy may have started with rabbits and trout, Borjas formalized the Roy model for immigration with his 1987 article "Self-Selection and the Earnings of Immigrants"
- Let's discuss the basics of the model (see Autor's notes for more details)
- There are two countries, 0 (the source country) and 1 (the host country)
- Log earnings in the source country are given by:

$$w_0 = \mu_0 + \varepsilon_0$$

- Log earnings in the host country are given by:

$$w_1 = \mu_1 + \varepsilon_1$$

Selection Into Migration

- The log earnings equations:

$$w_0 = \mu_0 + \varepsilon_0$$

$$w_1 = \mu_1 + \varepsilon_1$$

- μ_i designates the mean log earnings in country i
- ε_i can be thought of as the returns to a worker's skill relative to the mean worker in country i
- Assume that ε_i is distributed normally with mean zero and variance σ_i^2

Selection Into Migration

- It is costly to migrate
- Let's assume for the moment that $\mu_0 \approx \mu_1$ (but do spend some time thinking about alternative cases)
- For migrants, the gains from migration in terms of switching from ε_0 to ε_1 have to outweigh these costs
- So does this mean that migrants will necessarily be very high ability?
- Maybe, maybe not
- The nature of selection will depend crucially on σ_0, σ_1 and ρ (the correlation between σ_0 and σ_1)

Selection Into Migration

Case 1: Immigrants are positively selected from the source country and are above the mean of the host country ($E(\varepsilon_0|migrant) > 0$ and $E(\varepsilon_1|migrant) > 0$)

- This will be the case if $\frac{\sigma_1}{\sigma_0} > 1$ and $\rho > \frac{\sigma_0}{\sigma_1}$
- The first condition says that the return to skill is greater in the host country than the source country
- The second condition essentially says that being high skilled in the source country will translate into being high skilled in the host country

Selection Into Migration

Case 2: Immigrants are negatively selected from the source country and are below the mean of the host country ($E(\varepsilon_0|migrant) < 0$ and $E(\varepsilon_1|migrant) < 0$)

- This will be the case if $\frac{\sigma_1}{\sigma_0} < 1$ and $\rho > \frac{\sigma_1}{\sigma_0}$
- The first condition says that the return to skill is greater in the source country, so low skilled workers would prefer the compressed wage distribution of the host country
- The second condition essentially says that being low skilled in the source country will translate into being low skilled in the host country

Selection Into Migration

Case 3: Immigrants are selected from the lower tail of the source country and end up above the mean of the host country ($E(\varepsilon_0|migrant) < 0$ and $E(\varepsilon_1|migrant) > 0$)

- This can occur if $\rho < min\left(\frac{\sigma_1}{\sigma_0}, \frac{\sigma_0}{\sigma_1}\right)$
- This is effectively saying that there is some reason that a migrant receives relatively low pay in the source country but ends up highly paid in the host country
- When could this happen?

Selection Into Migration

- Let's look at some empirical evidence of selection into migration during the age of mass migration
- It is an interesting time period to study because of the nearly open border
- The central question is whether immigrants were positively or negatively selected
- We'll look at empirical evidence of both cases
- After the question of selection is addressed, there is a second question: what sort of mobility did immigrants experience after arriving in the United States?

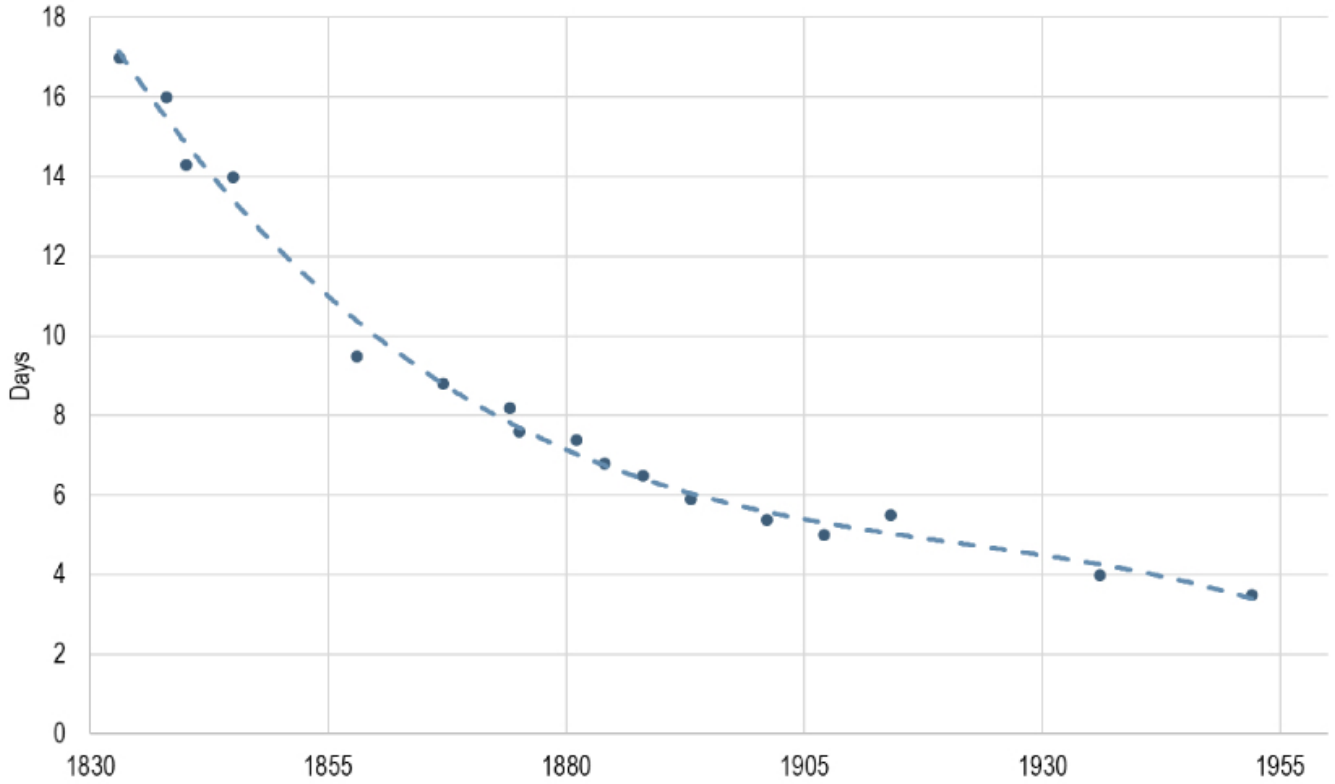
Selection Into Migration

- First, how do we look at who decided to emigrate?
- One classic approach is to turn to passenger lists
- In 1819, the US starts requiring ship captains to file ship records upon arrival in the United States
- Several economic history studies have used these passenger lists to get a sense of the age and occupational distributions of emigrants

Passenger Act of 1819

..[E]very ship or vessel bound on a voyage from the United States to any port on the continent of Europe, at the time of leaving the last port whence such ship or vessel shall sail, shall have on board, well secured under deck, at least sixty gallons of water, one hundred pounds of salted provisions, one gallon of vinegar, and one hundreds pounds of wholesome ship bread, for each and every passenger on board such ship...

Passenger Act of 1819

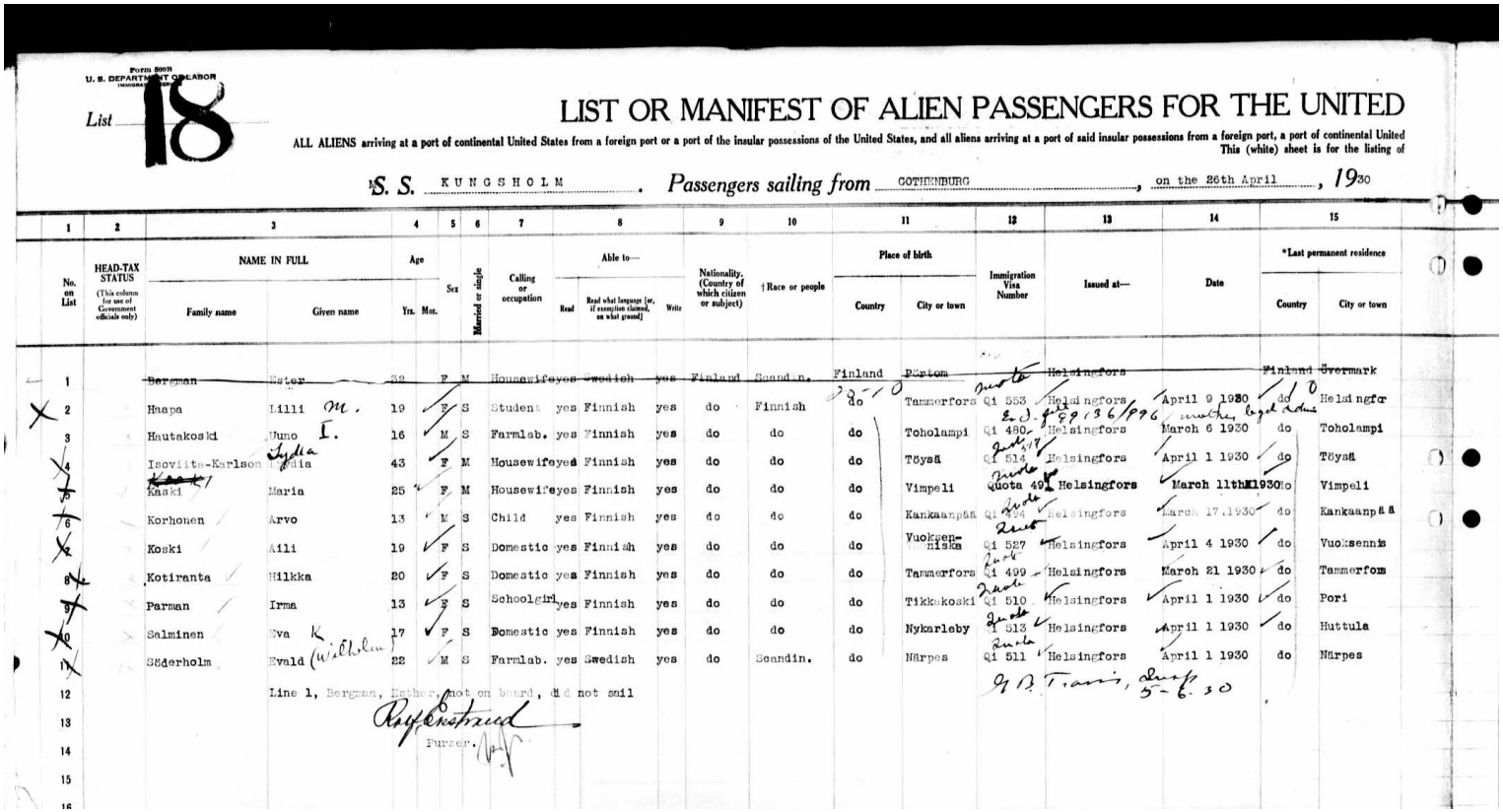


Liner transatlantic crossing times, 1833-1952 (P.J. Hugill (1993))

Passenger Act of 1819

...[T]he captain or master of any ship or vessel arriving in the United States, or any of the territories thereof, from any foreign place whatever...shall also deliver and report, to the collector of the district in which such ship or vessel shall arrive, a list or manifest of all the passengers taken on board of the said ship or vessel at any foreign port or place; in which list or manifest it shall be the duty of the said master to designate, particularly, the age, sex, and occupation, of the said passengers, respectively, the country to which they severally belong, and that of which it is their intention to become inhabitants; and shall further set forth whether any, and what number, have died on the voyage;

Passenger Lists



Form 8001
U. S. DEPARTMENT OF LABOR
Immigration Service

List 18

LIST OR MANIFEST OF ALIEN PASSENGERS FOR THE UNITED

ALL ALIENS arriving at a port of continental United States from a foreign port or a port of the insular possessions of the United States, and all aliens arriving at a port of said insular possessions from a foreign port, a port of continental United States or another insular possession, must be listed. This (white) sheet is for the listing of

S. S. KUNGSHOLM. Passengers sailing from GOTHENBURG, on the 26th April, 1930

1	2	3		4	5	6	7	8			9	10	11		12	13	14	15	
No. on List	HEAD-TAX STATUS (This column for use of Government officials only)	NAME IN FULL		Age				Able to—					Place of birth					*Last permanent residence	
		Family name	Given name	Yrs.	Mos.	Sex	Married or single	Calling or occupation	Read	Read what language [or, if exemption claimed, on what ground]	Write	Nationality. (Country of which citizen or subject)	†Race or people	Country	City or town	Immigration Visa Number	Issued at—	Date	Country	City or town
1		~~Bergman~~	~~Ester~~	~~52~~		~~F~~ M		Housewife	yes	~~Swedish~~	yes	~~Finland~~	Scandin.	Finland	Pietom		~~Helsingfors~~		~~Finland~~	~~Övermark~~
2		Haapa	Lilli M.	19		F	S	Student	yes	Finnish	yes	do	Finnish	do	Tammerfors	Q1 553	Helsingfors	April 9 1930	do	Helsingfor
3		Hautakoski	Juno I. Lydia	16		M	S	Farmlab.	yes	Finnish	yes	do	do	do	Toholampi	Q1 480	Helsingfors	March 6 1930	do	Toholampi
4		Isovirta-Karlson	Lydia	43		F	M	Housewife	yes	Finnish	yes	do	do	do	Töysä	Q1 514	Helsingfors	April 1 1930	do	Töysä
5		~~Kaski~~ Kaski	Maria	25		F	M	Housewife	yes	Finnish	yes	do	do	do	Vimpeli	quota 491	Helsingfors	March 11th 1930	do	Vimpeli
6		Korhonen	Arvo	13		M	S	Child	yes	Finnish	yes	do	do	do	Kankaanpää	Q1 [illegible]	Helsingfors	March 17,1930	do	Kankaanpää
7		Koski	Aili	19		F	S	Domestic	yes	Finnish	yes	do	do	do	Vuokyen-niska	Q1 527	Helsingfors	April 4 1930	do	Vuoksenniska
8		Kotiranta	Hilkka	20		F	S	Domestic	yes	Finnish	yes	do	do	do	Tammerfors	Q1 499	Helsingfors	March 21 1930	do	Tammerfors
9		Parman	Irma	13		F	S	Schoolgirl	yes	Finnish	yes	do	do	do	Tikkakoski	Q1 510	Helsingfors	April 1 1930	do	Pori
10		Salminen	Eva K	17		F	S	Domestic	yes	Finnish	yes	do	do	do	Nykarleby	Q1 513	Helsingfors	April 1 1930	do	Huttula
11		Söderholm	Evald (Wilhelm)	22		M	S	Farmlab.	yes	Swedish	yes	do	Scandin.	do	Närpes	Q1 511	Helsingfors	April 1 1930	do	Närpes
12			Line 1, Bergman, Esther, not on board, did not sail																	
13																				
14																				
15																				
16																				

Purser.

Passenger Lists

- So what can we do with passenger lists?
- The most basic thing is to look at the gender and age distributions of individuals arriving in the US
- Some passenger lists include occupation, although there is a (very interesting) question of how that will correspond to occupation in the US
- One thing we would really like if we're focusing on selection is a measure of human capital

Passenger Lists and Human Capital

- The manifests do not contain a direct measurement of education level beyond (sometimes) literacy
- If we don't know education level or literacy, what else can we use?
- It turns out that a decent source of information on human capital is actually age
- We can infer something about the basic human capital of immigrants by looking at the distribution of ages, specifically the distribution of the last digit of age
- Let's head over to Stata to see how this works

Passenger Lists and Human Capital

- Mokyr and O'Grada use this concept of age heaping to look at Irish immigrants
- They use a slightly more complex measure than what we just saw
- In particular, they are measuring

$$\gamma = \sum_{i=15}^{34} \left(\frac{n_i}{\sum n_i} - \frac{\hat{n}_i}{\sum \hat{n}_i} \right)^2$$

- This is basically saying how much deviation do we see from a smooth age distribution

Passenger Lists and Human Capital

Values of γ based on census data

Population	Males	Females	Total
Ireland, 1841	54.18	72.36	66.09
Ulster, 1841	47.46	61.54	54.49
Leinster, 1841	43.98	66.40	54.80
Munster, 1841	57.72	88.21	72.32
Connaught, 1841	80.14	117.70	98.98
Ireland, 1851	63.23	89.24	88.42
Ulster, 1851	53.52	57.89	62.33
Leinster, 1851	50.00	61.80	74.88
Munster, 1851	72.13	115.50	92.27
Connaught, 1851	97.00	117.53	140.44
U.S. 1880			10.23
U.S. 1970			1.16
Mexico, 1960			22.72

Passenger Lists and Human Capital

Values of γ based on passenger lists

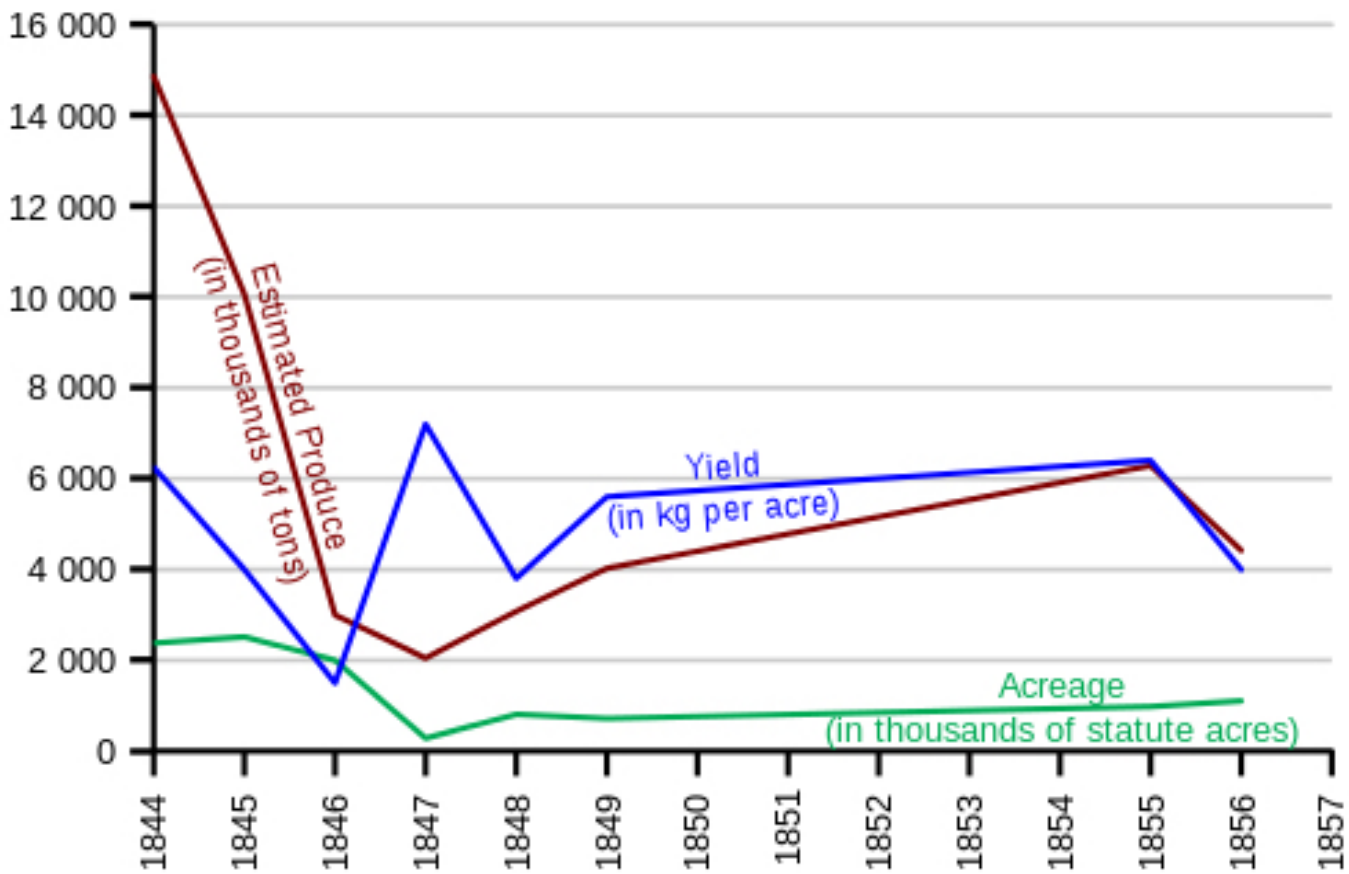
Sample	Males	Females	Total
1803-06, Ireland	89.8	130.8	96.3
1819-20, Ireland	146	108.5	106.5
1830-39, Derry	262.9	319	268.5
1820-48, New York	113.2	158.7	122.9
1822-39, Boston	125.9	117	111.9
1820-50, Sweden	22.1	27.9	20.5

Passenger Lists and Human Capital

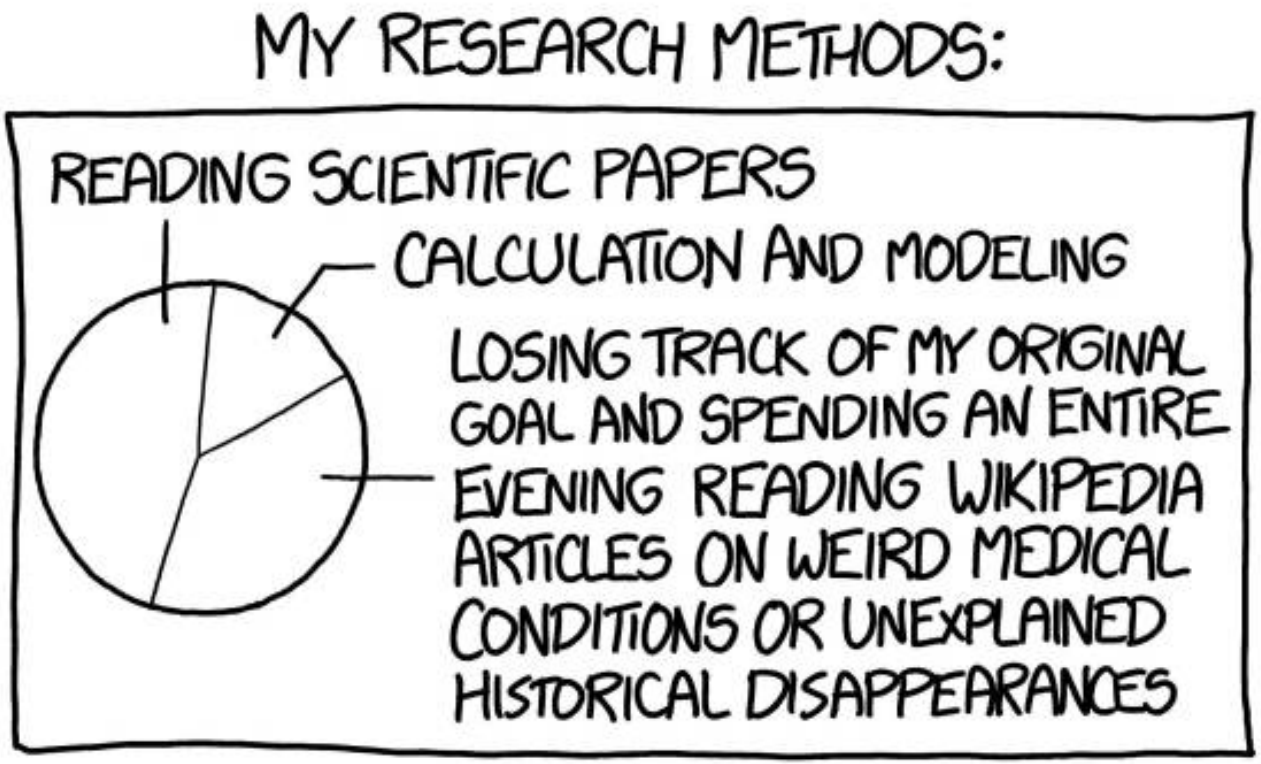
Values of γ based on New York passenger lists only

Period	Males	Females	Total
1820-24	103.6	209.8	107.4
1825-29	113.1	170.2	114.2
1830-34	96	109.5	96.7
1835-39	100.5	143.9	104.9
1840-44	92.2	141.5	103.8
1845-46	158.7	175.1	150
1847-48	217	296.3	239.9
Total	113.2	158.7	112.9

Passenger Lists and Human Capital



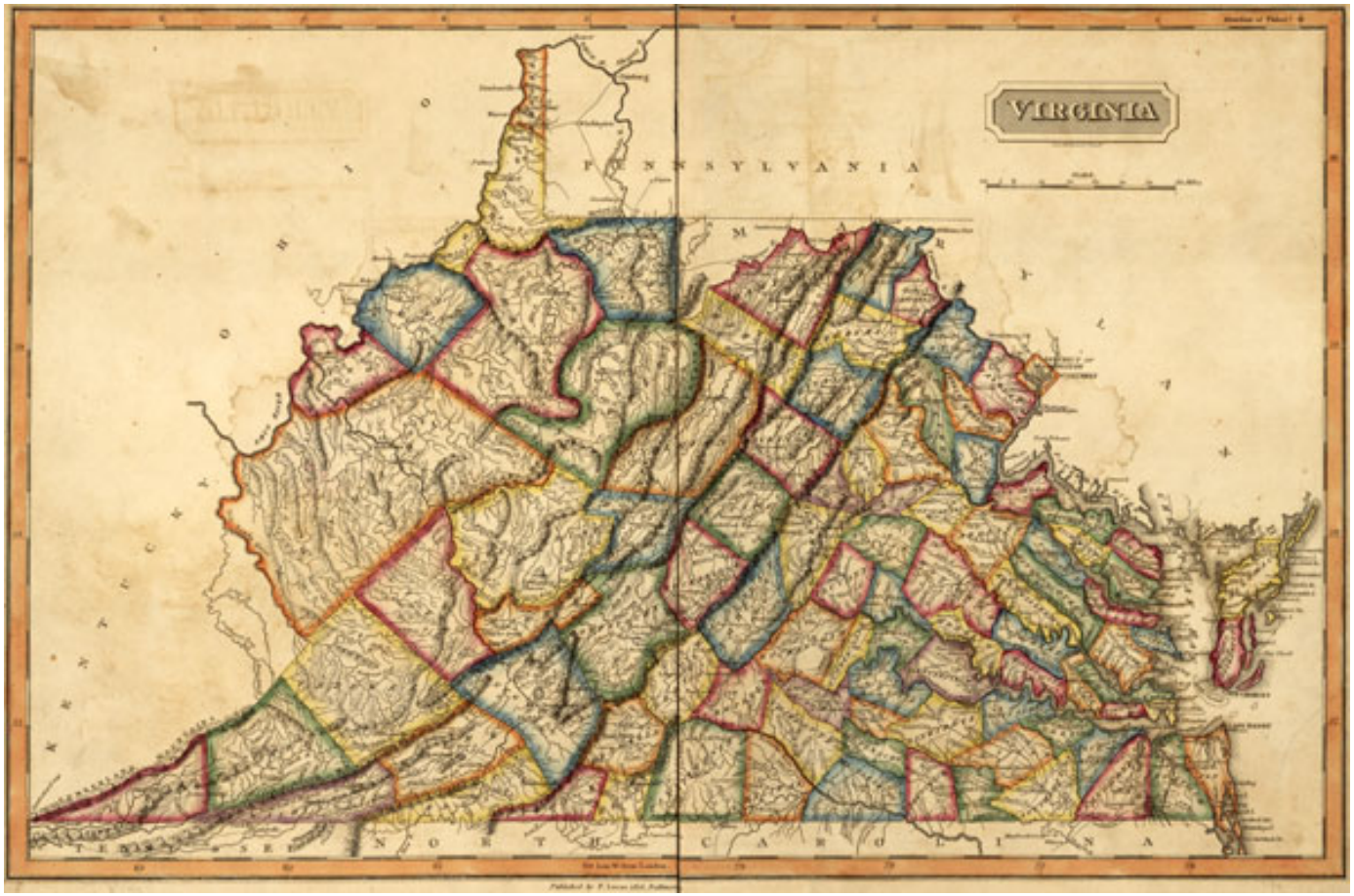
Final Project



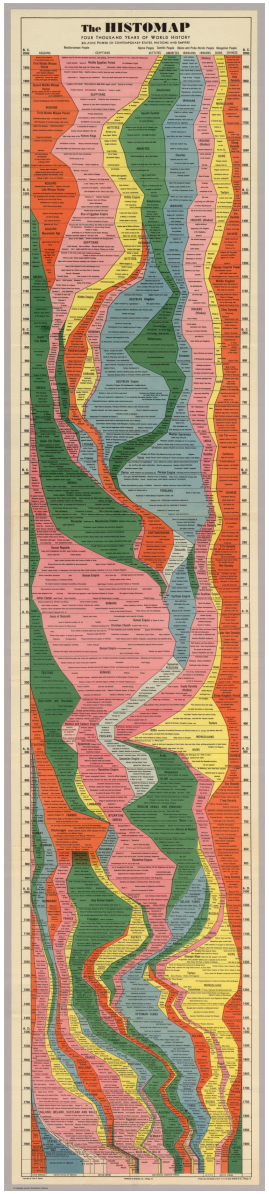
The Basics

- You need to do an original empirical study related to local mobility or inequality in a historical context
- You may want to focus on changes in mobility or inequality between groups over time
- You may also want to consider the impact of a historical shock on mobility or inequality
- Either way, you need to do original research that includes:
  - Developing a research question with a testable hypothesis
  - Collecting micro-level data (you can use the class datasets, you can also decide to use other data sources)
  - Presenting summary statistics and stylized facts using those data
  - Presenting novel regression analysis to test your hypothesis
- You will produce both a research paper and a policy memo

What is local?



What is historical?



The Research Paper

- The research paper should be approximately 15 pages double-spaced inclusive of figures and references
- Write it in the style of an economics journal article
- Typical components:
  - Introduction
  - Literature review
  - Description of data and methods
  - Presentation of empirical analysis
  - Conclusion
- Key thing is that this requires original data analysis

The Policy Memo

- Two pages double-spaced with at least one figure or table (this is a strict limit)
- It should contain:
  - A succinct statement of the policy issue
  - A concise summary of your methodology and findings
  - Discussion of the importance of your findings to policy
  - Recommendation for policy change
- Must be written for non-economists (and non-locals)

More Details

- Both the paper and policy memos are due by 5pm on April 29th
- Submit them by email to me (jmparman@wm.edu)
- They should be submitted as separate pdf's (both can be sent in the same email)
- Pay attention to formatting, it does matter for the grading
- I'm happy to look over drafts and give feedback

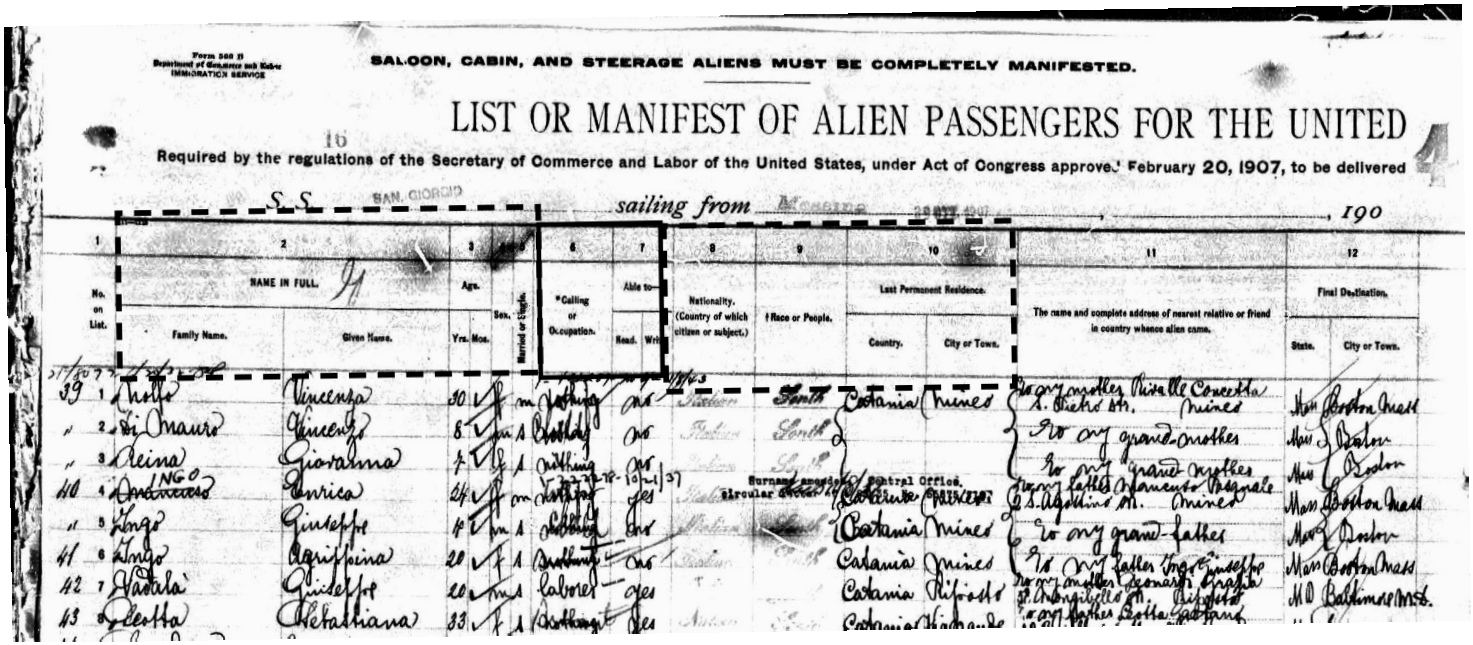
Failure is an option, sort of

Every failure is a step to success. Every detection of what is false directs us towards what is true: every trial exhausts some tempting form of error. Not only so; but scarcely any attempt is entirely a failure; scarcely any theory, the result of steady thought, is altogether false; no tempting form of Error is without some latent charm derived from Truth. – W. Whewall, Lectures on the History of Moral Philosophy in England, 1852

Some Quick Tips

- Get started early, even if you plan to use the class data (you know what variables will be in there)
- If using Stata, take advantage of the tutorials on our course website and the online help files and forums
- Use my office hours, especially if you are getting stuck on something technical
- Also consider meeting with the research librarians
- Check out the resources in the final project guidelines on Blackboard (and that we covered when talking about the Du Bois project)
- Google Scholar and Google Books are your friends

Passenger Lists and Human Capital

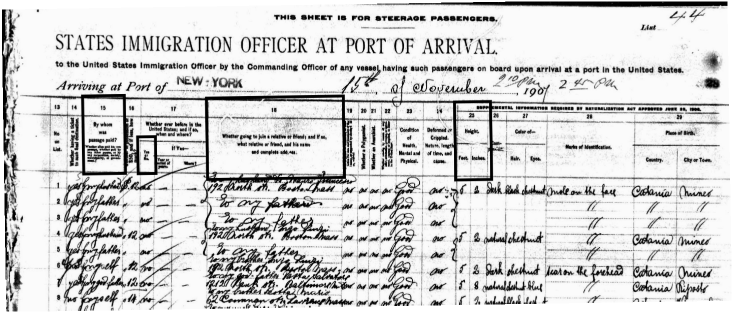


SALOON, CABIN, AND STEERAGE ALIENS MUST BE COMPLETELY MANIFESTED.

LIST OR MANIFEST OF ALIEN PASSENGERS FOR THE UNITED

Required by the regulations of the Secretary of Commerce and Labor of the United States, under Act of Congress approved February 20, 1907, to be delivered

S.S. SAN. GIORGIO sailing from ... 190


THIS SHEET IS FOR STEERAGE PASSENGERS.
List 44
STATES IMMIGRATION OFFICER AT PORT OF ARRIVAL.
to the United States Immigration Officer by the Commanding Officer of any vessel having such passengers on board upon arrival at a port in the United States.
Arriving at Port of NEW YORK
By whom was passage paid?
Whether ever before in the United States; and if so, when and where?
Whether going to join a relative or friend; and if so, what relative or friend, and his name and complete address.
Whether a Polygamist.
Whether an Anarchist.
Condition of Health, Mental and Physical.
Deformed or Crippled. Nature, length of time, and cause.
Height.
Feet.
Inches.
Complexion.
Color of—
Hair.
Eyes.
Marks of Identification.
Place of Birth.
Country.
City or Town.

Passenger Lists and Human Capital

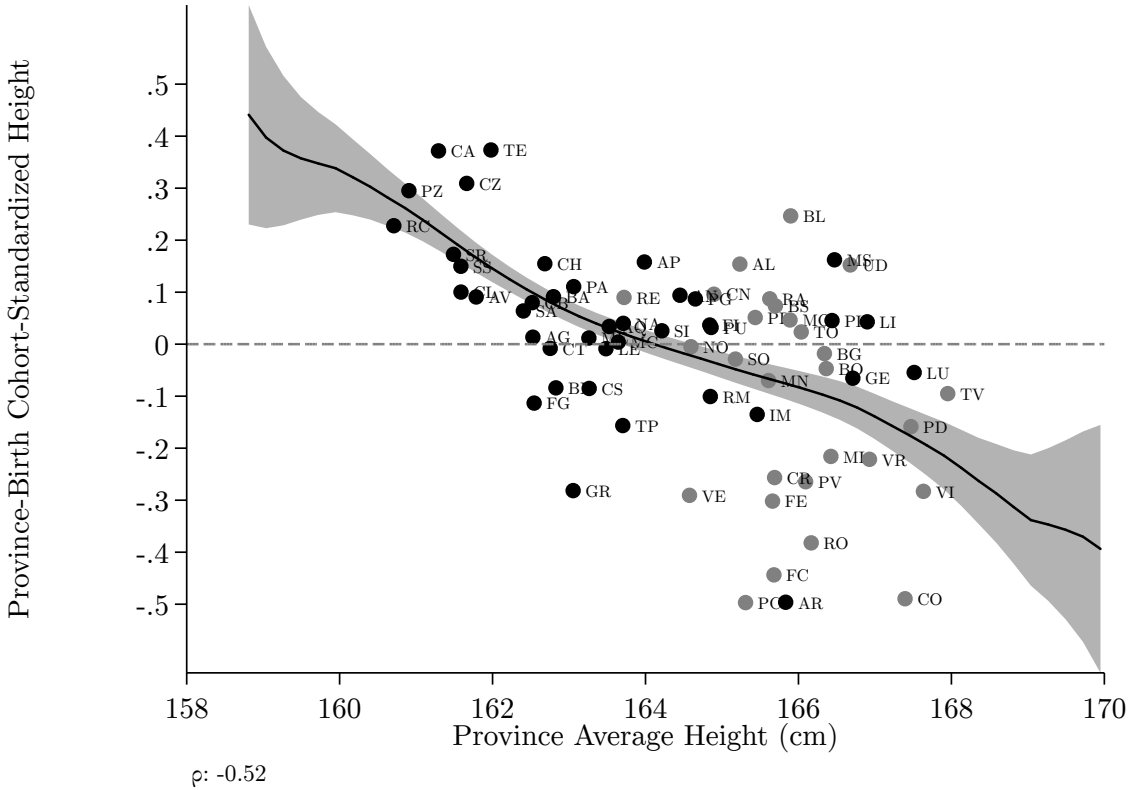


Figure 3: Province-cohort z-score and average province height.

Passenger Lists and Human Capital

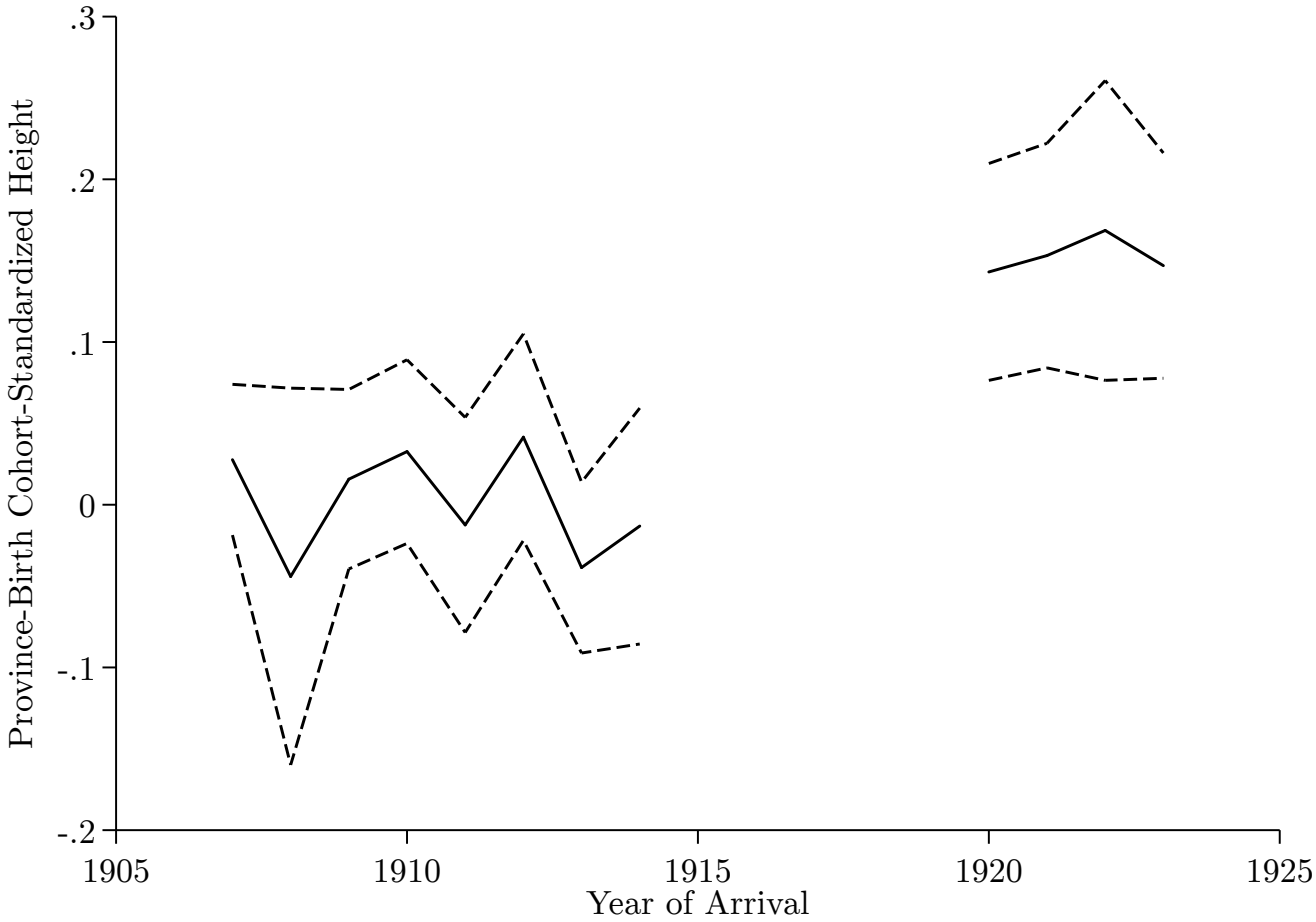


Figure 4: Local selection by year of arrival.

Passenger Lists and Human Capital

- So height gives us another way to get at human capital
- Height does suffer from some limitations (genetic differences across different population groups, inability to capture variation after full height is achieved, ...)
- One thing it does show us in the context of Spitzer and Zimran's article is that the nature of selection can vary across localities within a sending country

Selection Into Migration

- Let's switch to something familiar, Long and Ferrie's work on the US versus Britain
- We've already looked at how Long and Ferrie could link individuals within the US and within Britain
- This strategy also works for linking individuals between the US and Britain
- That means you can see the characteristics of who moved and who stayed and trace the mobility patterns of who moved and those who stayed

Selection of British Migrants

- So Long and Ferrie have linked records for US sons and fathers, British fathers and with sons who stayed in Britain, and British fathers with sons who moved to the US
- They are going to think about migration decisions being the outcome of three equations:

$$y_{1i} = \beta_1 X_{1i} + \varepsilon_{1i} \text{ if } M_i = 1$$

$$y_{0i} = \beta_0 X_{0i} + \varepsilon_{0i} \text{ if } M_i = 0$$

$$M_i = \begin{cases} 1 & \text{if } \gamma_1 Z_i + \gamma_2 (y_{1i} - y_{0i}) + u_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Selection of British Migrants

$$y_{1i} = \beta_1 X_{1i} + \varepsilon_{1i} \text{ if } M_i = 1$$

$$y_{0i} = \beta_0 X_{0i} + \varepsilon_{0i} \text{ if } M_i = 0$$

$$M_i = \begin{cases} 1 & \text{if } \gamma_1 Z_i + \gamma_2 (y_{1i} - y_{0i}) + u_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Here y_{ji} is the individual i's outcome in country j and X_{ji} is the individual's observable characteristics
- y_{ij} is an ordered set of occupational outcomes, from best to worst it is: white collar, farmer, skilled and semiskilled, and unskilled
- M_i equals one if the individual migrates and zero if the individual doesn't

Selection of British Migrants

$$y_{1i} = \beta_1 X_{1i} + \varepsilon_{1i} \text{ if } M_i = 1$$

$$y_{0i} = \beta_0 X_{0i} + \varepsilon_{0i} \text{ if } M_i = 0$$

$$M_i = \begin{cases} 1 & \text{if } \gamma_1 Z_i + \gamma_2 (y_{1i} - y_{0i}) + u_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- You can estimate these equations using the observed migrants and stayers
- Once you estimate the equations, you can use the coefficients to predict what outcomes would have been if migrants stayed or stayers migrated

Selection of British Migrants

	Ordered Probit Switching Regression							
1881	Movers		Stayers		Structural Probit (Move)			
Characteristic	β	*t*-stat.	β	*t*-stat.	β	S.E.	[90% C.I.]	
Father's Class: 1. WC	0.56	4.39***	0.97	13.35***				
Father's Class: 2. F	0.48	2.30**	0.88	10.02***				
Father's Class: 3. SS	0.25	2.40**	0.37	6.35***				
Age	0.13	0.93	0.10	1.47	0.16	0.15	[-0.10	0.41]
Age^2	0.00	0.85	0.00	1.44	-0.01	0.01	[-0.01	0.00]
Father's Age	0.00	0.23	0.01	3.60***	0.00	0.01	[-0.01	0.01]
Father in Agric.	-0.06	0.33	-0.43	6.19***				
One Servant in HH	0.34	2.40**	0.34	4.07***	0.11	0.17	[-0.17	0.40]
2+ Servants in HH	0.42	2.48**	0.55	4.96***	0.02	0.23	[-0.35	0.40]
Age Discrepancy	-0.02	0.30	-0.06	2.80***	0.16	0.06	[0.06	0.26]
Eldest Child	-0.06	0.74	-0.04	0.91	-0.12	0.09	[-0.28	0.04]
Oldest Brother in HH					-0.07	0.06	[-0.17	0.02]
Children in HH					0.04	0.01	[0.02	0.06]
Mother Employed					-0.19	0.07	[-0.31	-0.07]
Parish ≠ Birth Parish					-0.04	0.04	[-0.10	0.03]
$\hat{y}_M - \hat{y}_S$					-0.90	0.37	[-1.51	-0.29]
Constant	-0.22	0.21	-0.37	0.74	-2.18	1.08	[-3.95	-0.39]

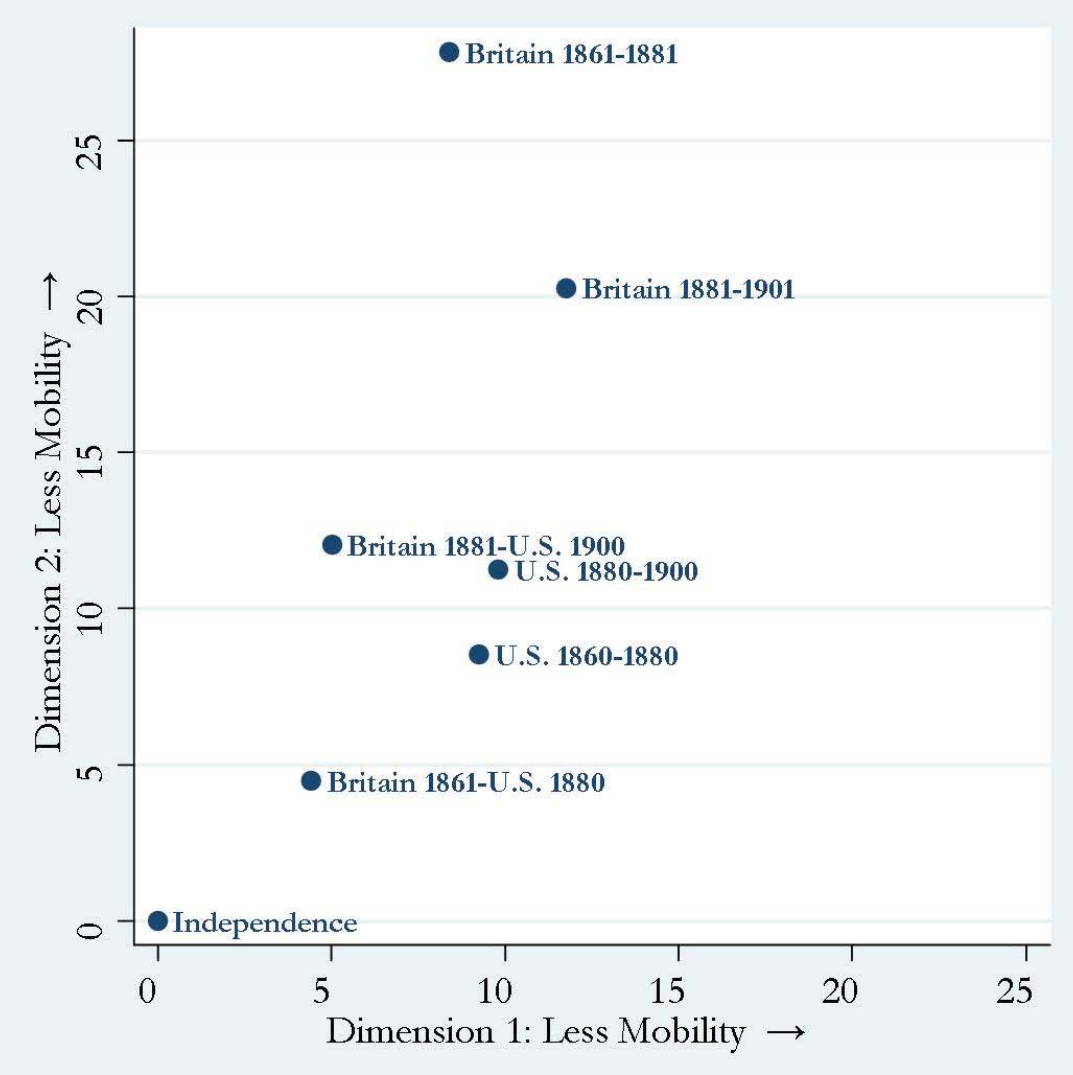
Note: Observations: 5,025. Omitted categories are "Father's Class: 4. U," "No Servants in HH," "<2 Servants in HH," "Not Eldest Child," "Not Oldest Brother in HH," "Mother Not Employed," and "Parish=Birth Parish." Structural Probit SEs and CIs calculated by bootstrapping via data resampling with 500 repetitions.
* significant at 10%; ** significant at 5%; *** significant at 1%

Selection of British Migrants

Parameter	Estimate	S.E.	[90% C.I.]	
(1) $\hat{y}_M$, Movers	1.093	0.127	[0.883	1.302]
(2) $\hat{y}_M$, Stayers	1.028	0.138	[0.801	1.254]
(3) $\hat{y}_S$, Movers	1.128	0.153	[0.876	1.380]
(4) $\hat{y}_S$, Stayers	1.002	0.130	[0.787	1.216]
(5) s_M, Selection of migrants=(1)-(2)	0.065	0.027	[0.021	0.109]
(6) s_S, Selection of stayers=(4)-(3)	-0.127	0.029	[-0.174	-0.080]
(7) τ_M, Treatment Effect: Treated=(1)-(3)	-0.036	0.196	[-0.359	0.288]
(8) τ_S, Treatment Effect: Not Treated=(2)-(4)	0.026	0.186	[-0.281	0.333]
Average Treatment Effect	0.013	0.188	[-0.297	0.322]

Note: SEs and CIs are calculated by bootstrapping via data resampling with 500 repetitions.

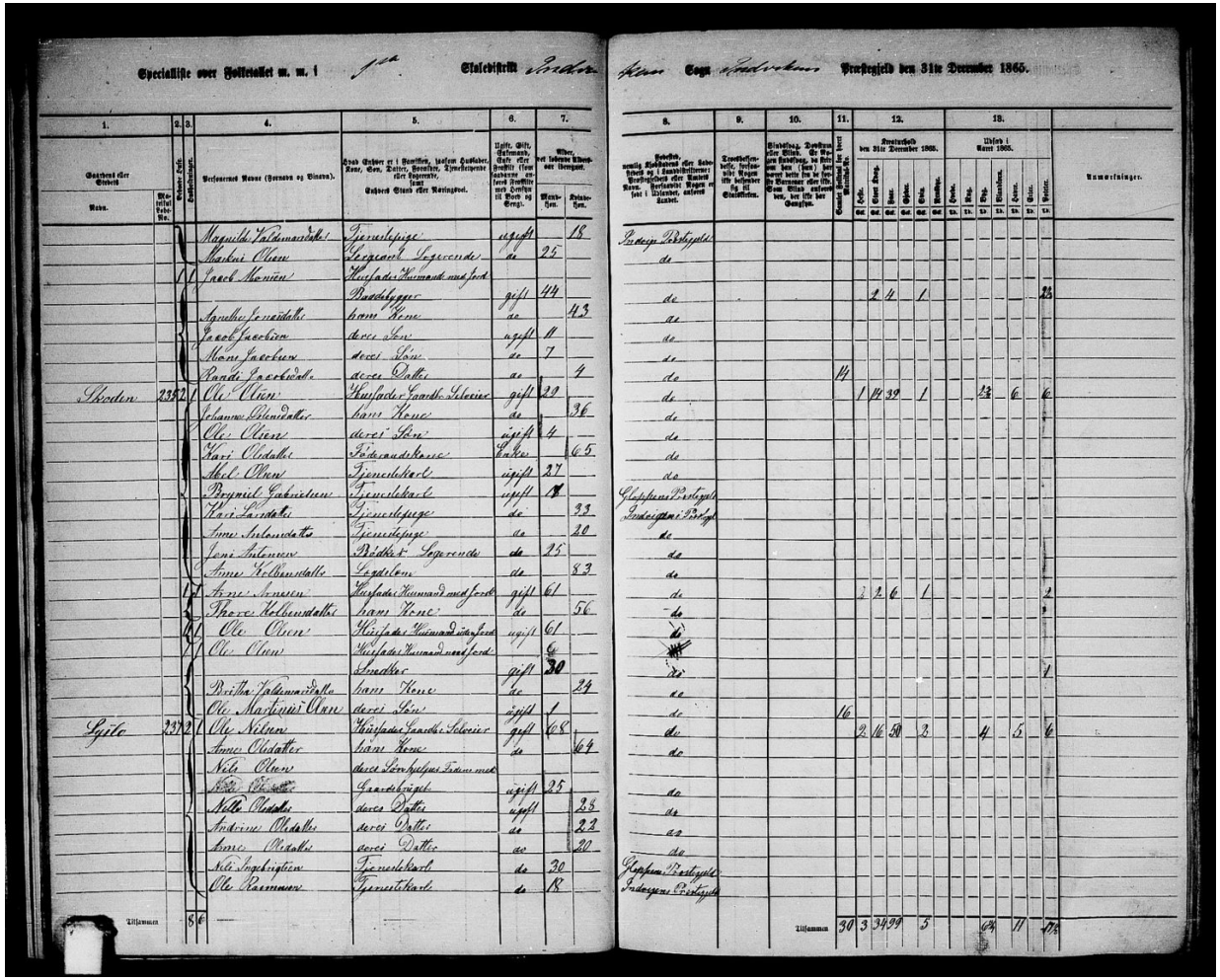
Selection of British Migrants



Selection of British Migrants

- So it looks from Long and Ferrie like British migrants may have been positively selected
- It also looks like they experienced substantial occupational mobility when arriving in the US, even more than US sons were experiencing
- But we have a bit of a problem
- We only truly observe the impacts of migration for those who decide to migrate
- We can't really say what would have happened if a non-migrant was randomly chosen to migrate
- Time to get more ambitious with data linking and switch our focus to Norway

The Norwegian Census



Matching Migrants

- Abramitzky et al. are going to match individuals between the 1865 Norwegian census and the 1900 Norwegian and United States censuses
- They use three different techniques:
  - Match 1: Use a standard iterative matching technique to match the population of Norwegian-born men in 1900 to their childhood households in 1865 using name, age, and country of birth
  - Match 2: Add province of birth for men who remain in Norway
  - Match 3: Restrict to men who are unique by name within a five-year age band in both censuses
- These different approaches represent tradeoffs between sample size and false matches

What We Can and Can't Learn From Individuals

- By just observing migrants and non-migrants, we can learn a few things
- At most basic level, we can see occupational distributions
- Let's take a look at overall occupational distributions and then at the occupational distributions for migrants and non-migrants separately
- This will give us some (limited) insight into selection and returns to migration

Norwegians and the Roy Model

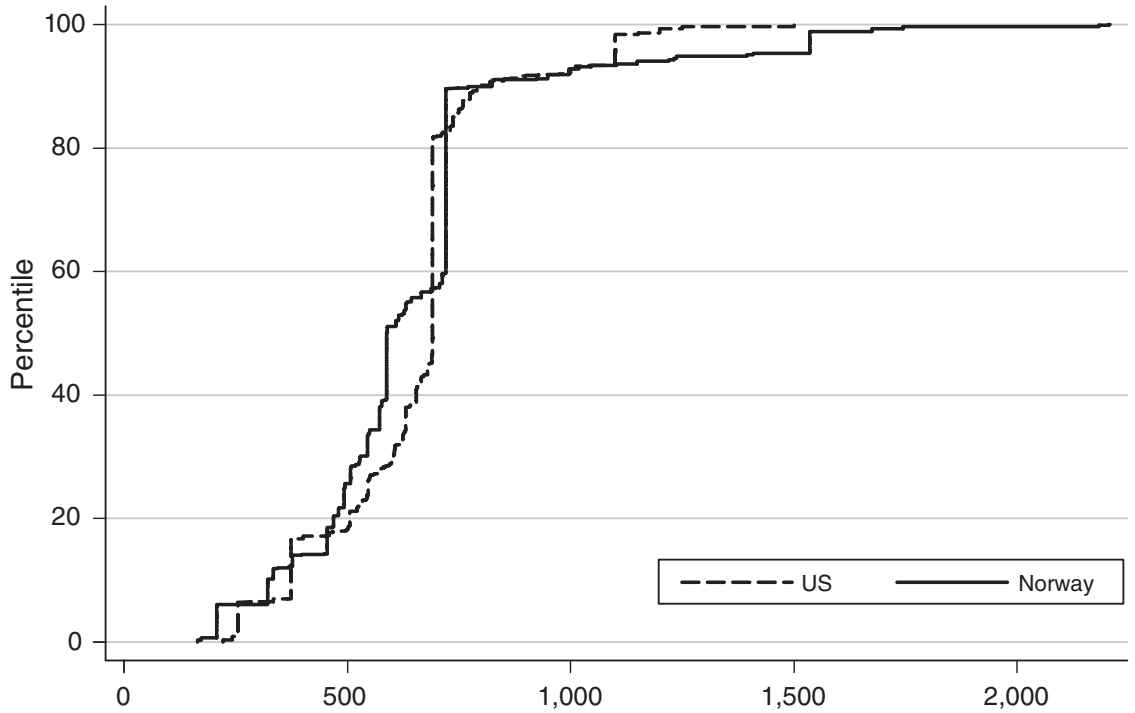


FIGURE 1. CUMULATIVE INCOME DISTRIBUTION FUNCTIONS IN THE UNITED STATES AND NORWAY IN 1900

Notes: US and Norwegian distributions contain all men aged 38 to 50 in the respective censuses of 1900. The x-axis is scaled in 1900 US dollars. Individuals are assigned the mean earnings for their occupation and are arrayed from lowest- to highest-paid occupations. The Norwegian distribution is rescaled to have the same mean as the US distribution (the actual Norwegian and US means are US$(1900)350 and US$(1900)643, respectively).

Norwegians and the Roy Model

Panel B. Top ten occupations in matched sample, Norwegian-born men living in Norway in 1900

Rank	Occupation	Frequency	Percentage	Earnings
1	General farmers	4,189	22.26	393
2	Farmer and fisherman	1,522	8.09	321
3	Merchants and dealers	722	3.84	837
4	Fisherman	709	3.77	248
5	Husbandmen or cottars	658	3.50	114
6	Farm workers	597	3.17	175
7	Carpenters	505	2.68	312
8	Shipmasters and captains	459	2.44	298
9	Cottar and fisherman	412	2.19	321
10	Seamen	351	1.87	182
Total		10,124	53.79	

Notes: $N = 18{,}820$. Historical International Standard Classification of Occupations (HISCO) occupation categories. Annual earnings by occupation data from Statistik Centralbureau (1905) and Grytten (2007). Values reported in year 1900 dollars. Average incomes of owner-occupier farmers and fishermen are estimated using data from the Norwegian census of agriculture.

Norwegians and the Roy Model

Panel A. Top ten occupations in matched sample, Norwegian-born men living in the United States in 1900

Rank	Occupation	Frequency	Percentage	Earnings
1	Farmers and planters	1,012	35.81	691
2	Laborers (general)	256	9.05	373
3	Carpenters and joiners	174	6.15	630
4	Farm laborers	101	3.57	255
5	Painters, glaziers, and varnishers	66	2.33	624
6	Sailors	60	2.12	467
7	Saw and planing mill workers	42	1.49	572
8	Machinists	39	1.38	736
9	Railroad laborers	36	1.27	460
10	Salesmen	32	1.13	680
Total		1,818	64.33	

Notes: $N = 2{,}826$. Occupation data collected by hand from census manuscripts on Ancestry.com. Annual earnings by occupation data from the 1901 Cost of Living Survey reported in Preston and Haines (1991) in year 1900 dollars. Average income of owner-occupier farmers is estimated using data from the US census of agriculture.

Announcements

- Let me know if you didn't get an email confirmation for your second referee report
- Time to really focus on your data projects and thinking about research project ideas
- Today and Wednesday we are going to wrap up immigration
- Up next: education, returns to skill, and inequality
  - Gray (2013) "Taking Technology to Task: The Skill Content of Technological Change in Early 20th Century United States"
  - Parman (2011) "American Mobility and the Expansion of Public Education"

Norwegians and the Roy Model

- So in the context of Norway, it looks like we would expect negative selection of migrants
- Why? The US earnings distribution was more compressed at the time than that of Norway
- One thing to think about: could low earning Norwegians really afford the trip?
- Seems that way, the total cost of migration was about 18 percent of the annual earnings of a Norwegian farm laborer (compare this to our discussion of indentured servitude)
- How will this negative selection impact estimates of the return to migration?

The Returns to Migration

- Let's start with a very naive estimate of the returns to migration:

$$ln(\text{Earnings}_i) = \alpha + \beta_1(\text{Migrant}_i) + \beta_2(\text{Age}_i) + \beta_3(\text{Age}_i)^2 + \varepsilon_i$$

- In this equation, we can think of β_1 as representing the return to migration
- Let's see what Abramitzky et al. get when running this regression

The Returns to Migration - Historical

TABLE 2—OLS REGRESSIONS OF THE RETURN TO MIGRATION FROM NORWAY TO THE UNITED STATES

	Dependent variable = ln(earnings)						
	Population	Match 1	Match 2	Match 3	Weighted	Match 1 Iowa data	Add penalty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
In US	0.609	0.606	0.644	0.572	0.641	0.554	0.466
	(0.017)	(0.009)	(0.009)	(0.015)	(0.024)	(0.010)	(0.009)
N	122,620	17,501	33,641	7,596	14,647	17,352	17,501

Notes: Standard errors are reported in parentheses. All regressions control a quadratic in age. The first column contains a representative sample of the population of Norwegian-born men between the ages of 38–50 in 1900 from the 100 percent 1900 Norwegian census and 1 percent 1900 US census sample (IPUMS). Column 2 reports estimates from the first matched sample, which is based on an iterative matching strategy that searches first for an exact match and then for matches in a one- or two-year age band. Column 3 uses the second matched sample, which allows men to match in Norway by name, age, and province of birth. Column 4 reports estimates from the third matched sample, which instead requires that matched observations be unique within a five-year age band. Columns 5 through 7 return to the first matched sample. In column 5, US migrants are assigned earnings from the 1915 Iowa census (appropriately adjusted for inflation). We lose 157 observations whose occupations do not match categories in the Iowa census. In column 6, we reduce the Cost of Living earnings by 13 log points in each occupation based on the earnings penalty for Scandinavian migrants reported in Hatton and Williamson (1994). Column 7 weights the matched sample to reflect the urban status, asset holdings and occupational distribution of fathers in the full population. We lose 2,905 observations because of missing information (primarily missing data on fathers' occupations).

The Returns to Migration and Selection

$$ln(\text{Earnings}_i) = \alpha + \beta_1(\text{Migrant}_i) + \beta_2(\text{Age}_i) + \beta_3(\text{Age}_i)^2 + \varepsilon_i$$

- But there is a big problem interpreting β_1 as the return to migration
- If migrants are negatively (or positively) selected, chances are ε_i will be correlated with Migrant_i
- This is going to give us biased estimates
- So what can we do about this?

The Returns to Migration and Selection

- We can think about taking advantage of brothers in the sample
- Suppose that ε_{ij} can actually be thought of as having two components:

$$\varepsilon_{ij} = \alpha_j + \nu_{ij}$$

- Here we've added j to indicate individuals i's household
- α_j is a household-specific component of the error term and ν_{ij} is the individual-specific component
- What happens if we run a regression with household fixed effects?

The Returns to Migration and Selection

$$ln(\text{Earnings}_{ij}) = \alpha_j + \beta_1'(\text{Migrant}_{ij}) + \beta_2'(\text{Age}_{ij}) + \beta_3'(\text{Age}_{ij})^2 + \nu_{ij}$$

- Including household fixed effects through α_j eliminates common household characteristics (both observed and unobserved) from the error term
- If migrants are negatively selected, we would expect β_1' to be bigger than β_1
- If migrants are positively selected, we would expect β_1' to be less than β_1
- Let's see what we get

The Returns to Migration and Selection

TABLE 3—OLS AND WITHIN-HOUSEHOLD ESTIMATES OF THE RETURN TO MIGRATION. HOUSEHOLDS WITH TWO OR MORE MEMBERS IN THE MATCHED SAMPLE

Dependent variable = ln(earnings); Coefficient on = 1 if migrant			
	Full sample, 1865	Rural, 1865	Urban, 1865
Panel A. Unweighted			
OLS	0.545	0.607	0.384
	(0.027)	(0.034)	(0.044)
Within household	0.511	0.508	0.508
	(0.035)	(0.045)	(0.057)
Chi-squared	1.49	7.47	8.31
p-value	0.2218	0.0063	0.0039
N	2,655	1,823	832
Number of migrant-stayer pairs	326	167	159
Panel B. Weighted			
OLS	0.586	0.609	0.443
	(0.029)	(0.033)	(0.067)
Within household	0.542	0.529	0.561
	(0.039)	(0.042)	(0.049)
Chi-squared	2.13	4.60	5.65
p-value	0.1441	0.0320	0.0175
N	2,241	1,666	306
Number of migrant-stayer pairs	269	140	129

Notes: Each cell contains coefficient estimates and standard errors from regressions of ln(earnings) on a dummy variable equal to one for individuals living in the United States in 1900. Regressions also include controls for age and age squared. In each panel, the first row conducts an OLS regression for the restricted sample of households that have at least two matched members in the dataset and the second row adds household fixed effects. Panel B contains results from regressions weighted to reflect the urban status (full sample only), asset holdings, and occupational distribution of fathers in the full population. We conduct chi-squared tests of the null hypothesis that the OLS and within-household coefficients are equal.

The Returns to Migration and Selection

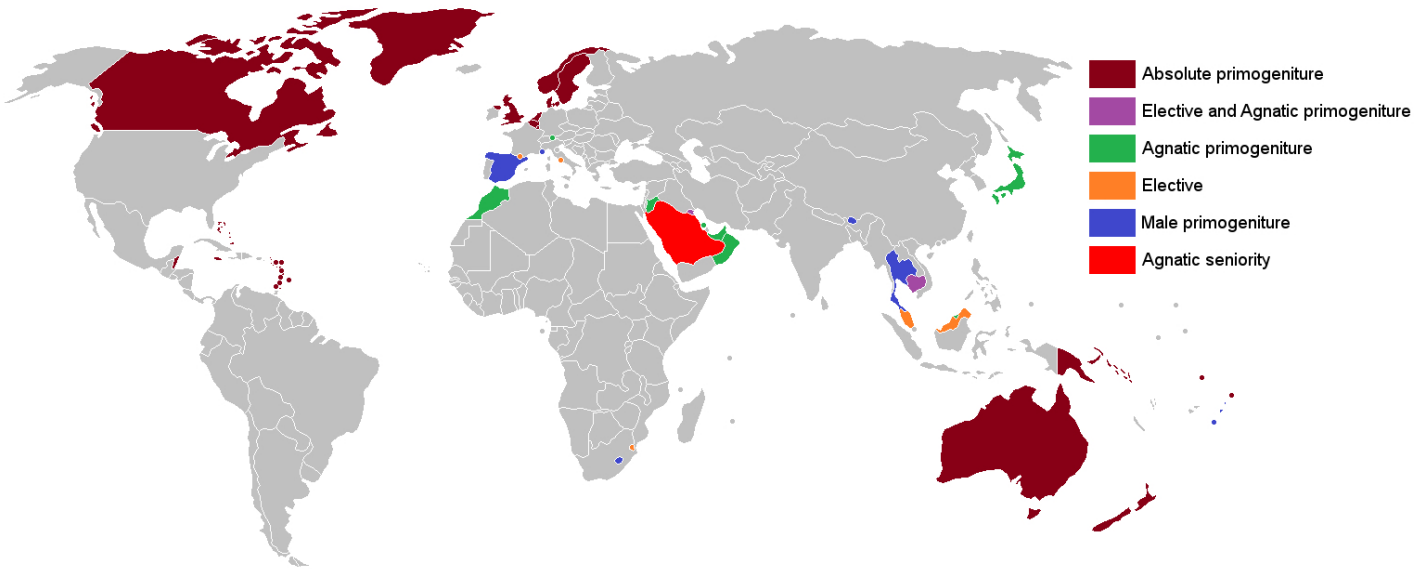
- So have we solved our selection issue?
- Only if the problematic component was α_j, the characteristics common to both brothers
- These are certainly important (think about the role of parents) but may not capture everything that matters
- What can we do if the problem is really ν_{ij}?
- Household fixed effects effects won't help us here

The Returns to Migration and Selection

The Returns to Migration and Selection

- So what can we do about ν_{ij}?
- We can't rely on fixed effects to get rid of ν_{ij}
- What we can do is try to find some variation in the decision to migrate that is uncorrelated with ν_{ij}
- More specifically, we would like an instrument that is correlated with the decision to migrate but uncorrelated with all of the other unobserved characteristics that influence earnings
- What can we use?

The Returns to Migration and Selection



The Returns to Migration and Selection

Appendix Table A1—Birth Order and Number of Brothers as Instruments for Migration to the United States

	(1)	(2)	(3)
Panel A. First stage	Dependent variable = In US in 1900		
Number of brothers	0.016		0.011
	(0.006)		(0.006)
2nd brother		−0.000	—
		(0.012)	
3rd brother		0.047	0.037
		(0.019)	(0.019)
4th or higher brother		0.076	0.058
		(0.035)	(0.036)
Panel B. OLS	Dependent variable = ln(earnings in 1900)		
In US in 1900		0.642	
		(0.019)	
Panel C. IV	Dependent variable = ln(earnings in 1900)		
In US in 1900	0.669	0.696	0.668
	(0.436)	(0.381)	(0.338)
Over-ID test (p-value)			0.869
N	4031	4031	4031

Notes: Standard errors are reported in parentheses. The sample includes men in Match 1 who lived in a rural household that had some assets in 1865 and whose mother is 42 years old or younger in 1865. The regressions also include a quadratic in age and dummy variables for total number of siblings in the household (see equation (3) in the text). In column 3, we report the p-value from a Sargan (chi-squared) test of overidentification.

The Returns to Migration and Selection

- There is one big question to think about with these results
- The returns to migration are being estimated from mean earnings by occupation
- What if migrants earn more or less within a given occupation?
- This seems like a very plausible scenario
- Let's work through this as a class in Stata

The Returns to Migration and Selection

- We'll use the dataset 1940-census-income-data.dta
- These data come from the IPUMS 1% sample of the 1940 federal census
- Variables are exactly as downloaded from IPUMS, check definitions here
- As a class, we are going to figure out how to best identify whether immigrant and native earnings differ within occupational categories and much this changes our interpretation of the returns to migration
- Let's head over to Stata

Announcements

- A cluster of class cancellations: March 29th, April 3rd, April 5th
- Internet may be spotty for me April 3rd through April 7th
- Consider getting a jump on data and the final research project so you can get questions answered before April 3
- Up next: education, returns to skill, and inequality
  - Gray (2013) "Taking Technology to Task: The Skill Content of Technological Change in Early 20th Century United States"
  - Parman (2011) "American Mobility and the Expansion of Public Education"

Assimilation after Migration

- So we have (sort of) established the returns to migration
- The next question is what inequality and mobility look like for migrants after arrival
- Do they converge with the native population in terms of earnings, education, social norms, etc.?
- Let's start by thinking about earnings convergence for the migrants themselves (their kids are a whole other story)

Estimating Earnings Convergence

- Given a single cross section of data, we might think about estimating earnings convergence with the following equation:

$$ln(y_i) = \alpha + \beta_1 FB_i + \beta_2 FB_i \cdot T_i + \beta_3 FB_i \cdot T_i^2 + \gamma X_i + \varepsilon_i$$

- Here $\ln(y_i)$ is log annual earnings, FB_i is an indicator for being foreign born, T_i is the number of years since migration, and X_i are observable characteristics that may influence earnings
- β_1 then tells us the earnings gap between native-born individuals and migrants when the migrant first arrives
- β_2 and β_3 then let us describe the speed at which earnings converge (or don't converge) to those of natives

Estimating Earnings Convergence

TABLE 2

REGRESSION ANALYSIS OF EARNINGS FOR NATIVE- AND FOREIGN-BORN ADULT WHITE MEN, 1970

	NATIVE BORN	NATIVE AND FOREIGN BORN			FOREIGN BORN
	(1)	(2)	(3)	(4)	(5)
EDUC	.07154	.07058	.07004	.07164	.05740
	(53.78)	(55.68)	(55.18)	(54.11)	(12.93)
T	.03167	.03050	.03071	.03097	.02028
	(22.99)	(22.86)	(22.99)	(23.10)	(3.47)
$T2$	−.00052	−.00049	−.00050	−.00051	−.00031
	(−20.77)	(−20.45)	(−20.78)	(−20.93)	(−3.18)
LN WW	1.10335	1.10326	1.10169	1.10111	1.07151
	(81.75)	(84.78)	(84.70)	(84.67)	(21.97)
RURALEQ1	−.17222	−.16970	−.17080	−.16915	−.05821
	(−20.28)	(−20.25)	(−20.39)	(−20.18)	(−1.13)
SOUTHEQ1	−.12090	−.12620	−.12530	−.12389	−.21587
	(−14.17)	(−15.01)	(−14.91)	(−14.74)	(−4.38)
NOTMSP	−.30647	−.31078	−.30947	−.30874	−.34498
	(−27.76)	(−28.97)	(−28.86)	(−28.79)	(−7.66)
FOR	*	.02951	−.16359	.00990	*
		(1.75)	(−4.32)	(0.18)	
(FOR) (YSM)	*	*	.01461	.01555	.01500
			(3.98)	(4.23)	(3.87)
(FOR) (YSM2)	*	*	−.00016	−.00018	−.00019
			(−2.47)	(−2.79)	(−2.82)
(FOR) (EDUC)	*	*	*	−.01619	*
				(−4.23)	
CONSTANT	−1.03646	−1.01537	−1.00016	−1.02156	−.78891
Observations (N)	34,321	36,245	36,245	36,245	1,924
R	.55423	.55455	.55533	.55564	.58194
R^2	.30717	.30753	.30839	.30873	.33866
Standard error	.70900	.71008	.70966	.70949	.71676

SOURCE.—U.S. Bureau of the Census 1972.
NOTE.—t-ratios in parentheses; dependent variable: natural logarithm of earnings in hundreds of dollars.
* Variable not entered.

Estimates from Chiswick (1978).

Estimating Earnings Convergence

TABLE 3

REGRESSION ANALYSIS OF EARNINGS FOR ADULT FOREIGN-BORN WHITE MEN WITHIN COUNTRY CATEGORIES, 1970

	BORN IN MEXICO OR NATIVE BORN OF SPANISH SURNAME*		BORN IN ENGLISH-SPEAKING DEVELOPED COUNTRIES†	FOREIGN BORN OTHER THAN ENGLISH-SPEAKING DEVELOPED COUNTRIES†	
	(1)	(2)	(3)	(4)	(5)
EDUC	.03573	.04324	.09217	.05211	.05086
	(4.01)	(4.28)	(5.70)	(10.27)	(7.06)
T	.01211	.01373	.06139	.01147	.01070
	(1.15)	(1.30)	(5.11)	(1.67)	(1.42)
$T2$	−.00028	−.00030	−.00095	−.00018	−.00017
	(−1.62)	(−1.74)	(−4.49)	(−1.59)	(−1.33)
LN WW	1.16436	1.16567	1.06921	1.05887	1.05879
	(12.47)	(12.50)	(11.39)	(18.72)	(18.71)
RURALEQ1	−.14442	−.14008	−.10296	−.05025	−.05122
	(−1.72)	(−1.67)	(−1.30)	(−.77)	(−.79)
SOUTHEQ1	−.24159	−.22760	−.12351	−.24956	−.24896
	(−3.81)	(−3.56)	(−1.31)	(−4.36)	(−4.34)
NOTMSP	−.45087	−.45043	−.41734	−.32680	−.32709
	(−5.91)	(−5.91)	(−5.09)	(−6.16)	(−6.17)
FOR	−.33633	−.18680	‡	‡	‡
	(−2.55)	(−1.15)			
(FOR) (EDUC)	‡	−.02402	‡	‡	‡
		(−1.57)			
(FOR) (YSM)	.02715	.03027	.01456	.01877	.01799
	(2.05)	(2.26)	(1.43)	(4.15)	(3.25)
(FOR) (YSM2)	−.00033	−.00038	−.00004	−.00024	−.00024
	(−1.38)	(−1.59)	(−.33)	(−3.09)	(−3.08)
(EDUC) (YSM)	‡	‡	−.00103	‡	.00007
			(−2.06)		(.24)
CONSTANT	−.73694	−.84163	−1.48900	−.62107	−.59879
Observations (N)	804	804	439	1,485	1,485
R	.55229	.55424	.63190	.56761	.56764
R^2	.30503	.30718	.39930	.32218	.32221
Standard error	.80627	.80533	.61350	.74032	.74056

SOURCE.—U.S. Bureau of the Census 1972.
NOTE.—t-ratios in parentheses: dependent variable: natural logarithm of earnings in hundreds of dollars.
* For the five Southwestern states Arizona, California, Colorado, New Mexico, and Texas.
† The English-speaking developed countries are Great Britain, Ireland, Canada, Australia, and New Zealand.
‡ Variable not entered.

Estimating Earnings Convergence

$$ln(y_i) = \alpha + \beta_1 FB_i + \beta_2 FB_i \cdot T_i + \beta_3 FB_i \cdot T_i^2 + \gamma X_i + \varepsilon_i$$

- These estimates often suggest rapid earnings convergence ($\sim$ 15 years) but there is a big problem with this interpretation
- If we use a single cross section, our T_i variable is picking up two very different things
  - Earnings convergence due to additional time in the US
  - Trends in cohort quality over time
- How can we get around this problem?
- One thing we can try is *repeated* cross sections

Estimating Earnings Convergence

- Consider the earnings difference in the year 2010 between the cohort migrating in 1980 and the cohort migrating in 1990:

$$ln(y_{10,80}) - ln(y_{10,90})$$

- The naive interpretation is that this difference is due to the extra ten years of time in the US for the 1980 cohort
- However, consider this decomposition of the earnings difference:

$$ln(y_{10,80}) - ln(y_{10,90}) = (ln(y_{10,80}) - ln(y_{00,80})) +$$

$$(ln(y_{00,80}) - ln(y_{10,90}))$$

Estimating Earnings Convergence

$$ln(y_{10,80}) - ln(y_{10,90}) = (ln(y_{10,80}) - ln(y_{00,80})) + (ln(y_{00,80}) - ln(y_{10,90}))$$

- The term on the left is capturing earnings growth for the 1980 cohort from 2000 to 2010
- The term on the right is capturing difference in earnings between the 1980 and 1990 cohorts when each had been in the country for 20 years

Estimating Earnings Convergence

Table 3
Decomposition of Cross-Section Growth in Immigrant Earnings

Group and Year of Immigration	Cross-Section Growth	Within-Cohort Growth	Across-Cohort Growth
White			
1965–69	.0665	.0029	.0636
	(1.46)	(.20)	(1.61)
1960–64	.1690	−.0111	.1801
	(3.33)	(−.17)	(4.14)
1950–59	.0558	.0089	.0469
	(1.80)	(.59)	(1.54)
Black			
1965–69	.2662	.0041	.2621
	(5.73)	(.10)	(3.35)
1960–64	.0066	−.1540	.1606
	(.10)	(−1.49)	(1.65)
1950–59	−.0831	−.2303	.1472
	(−1.33)	(−2.18)	(1.58)
Asian:			
1965–69	.2829	.1972	.0857
	(20.26)	(4.62)	(2.27)
1960–64	.1754	.1105	.0649
	(8.48)	(1.80)	(1.30)
1950–59	.0204	.0327	−.0123
	(.96)	(.39)	(−.01)

Estimates from Borjas (1985).

Estimating Earnings Convergence

- So do repeated cross sections solve all of our problems?
- Not quite, there is still a (potentially big) issue of return migration
- If return migration is not random, which it most certainly isn't, our remaining migrants in later census waves may be a pretty select group
- What to do? Switch to *longitudinal* data
- Why not do this in the first place? Same reason we've had to get creative with mobility measures
- Longitudinal data can be very hard to get your hands on and often leaves you with very small sample sizes

Estimating Earnings Convergence

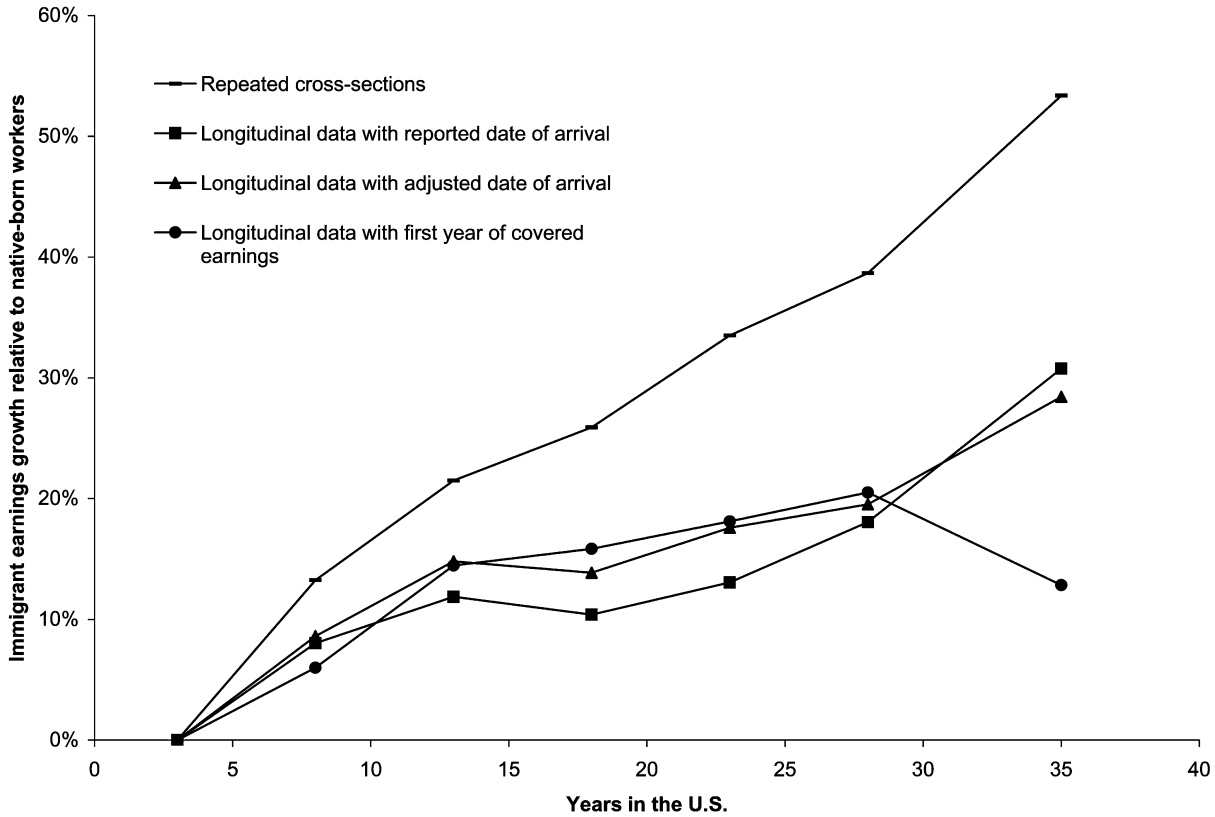
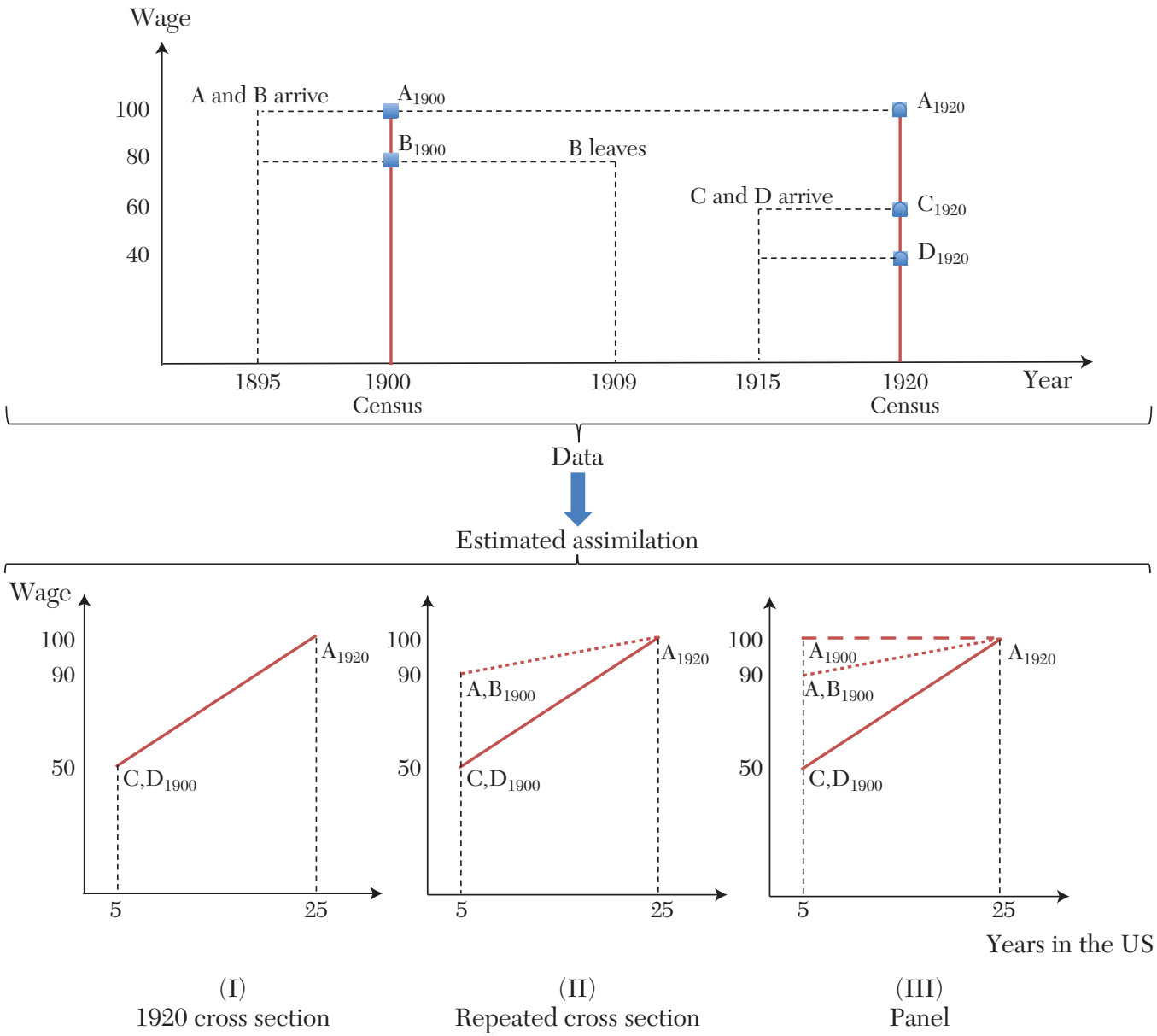


FIG. 2.—Immigrant earnings growth in repeated cross-sectional and longitudinal data. The figure plots the effect of immigrants' time in the United States on the immigrant-native earnings gap. Data are taken from estimates in table 5.

Estimating Earnings Convergence



Announcements

- A cluster of class cancellations: March 29th, April 3rd, April 5th
- Internet may be spotty for me April 3rd through April 7th
- Consider getting a jump on data and the final research project so you can get questions answered before April 3
- Up next: education, returns to skill, and inequality
  - Gray (2013) "Taking Technology to Task: The Skill Content of Technological Change in Early 20th Century United States"
  - Parman (2011) "American Mobility and the Expansion of Public Education"

Earnings Convergence in the Age of Mass Migration

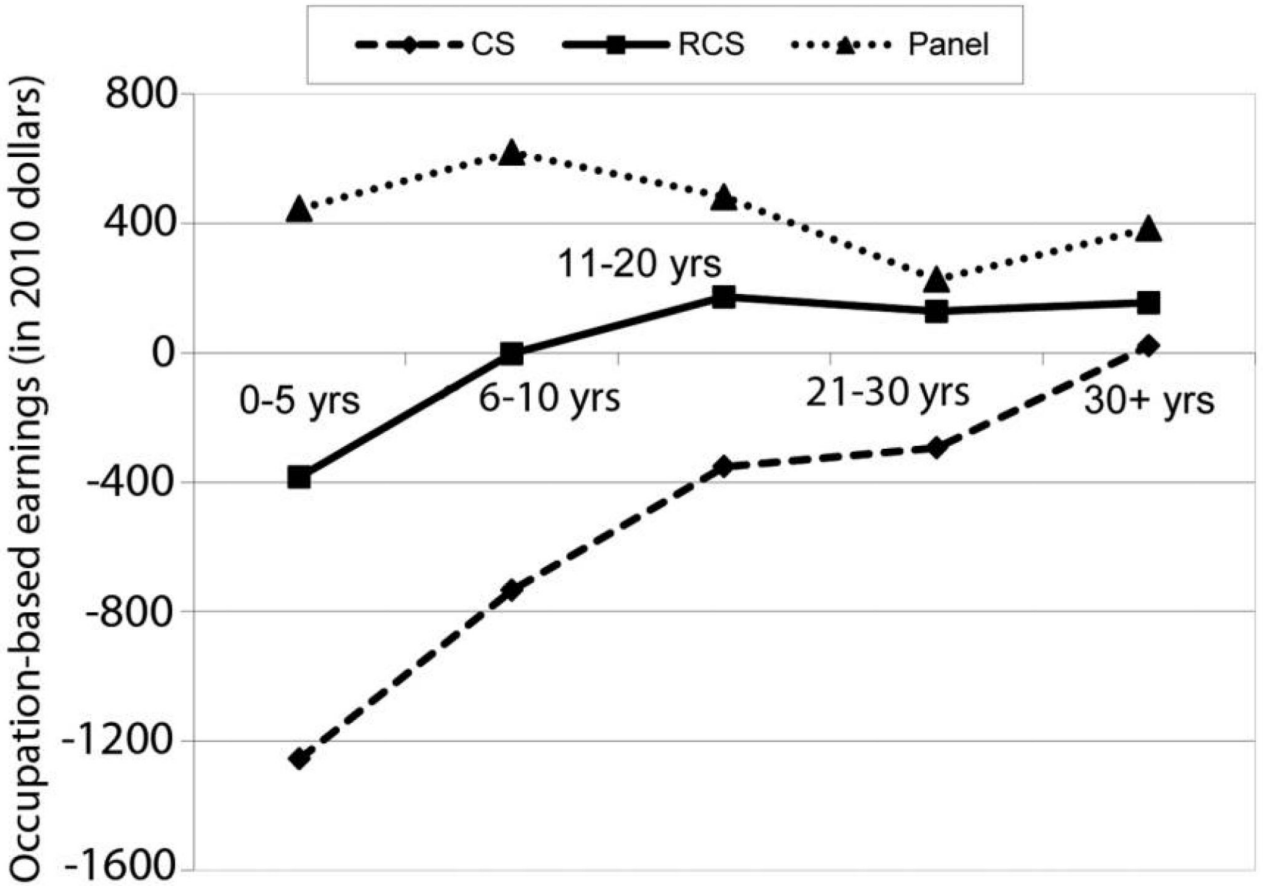


Fig. 2.
—Convergence in occupation score between immigrants and native-born workers by time spent in the United States, cross-sectional and panel data, 1900–1920. The graph plots coefficients for years spent in the United States indicators in equation (1). Note that for the panel line, we subtract the native-born dummy from the years in the United States indicators (because the omitted category in that regression is natives in the panel sample). See table 4 for coefficients and standard errors.

Earnings Convergence in the Age of Mass Migration

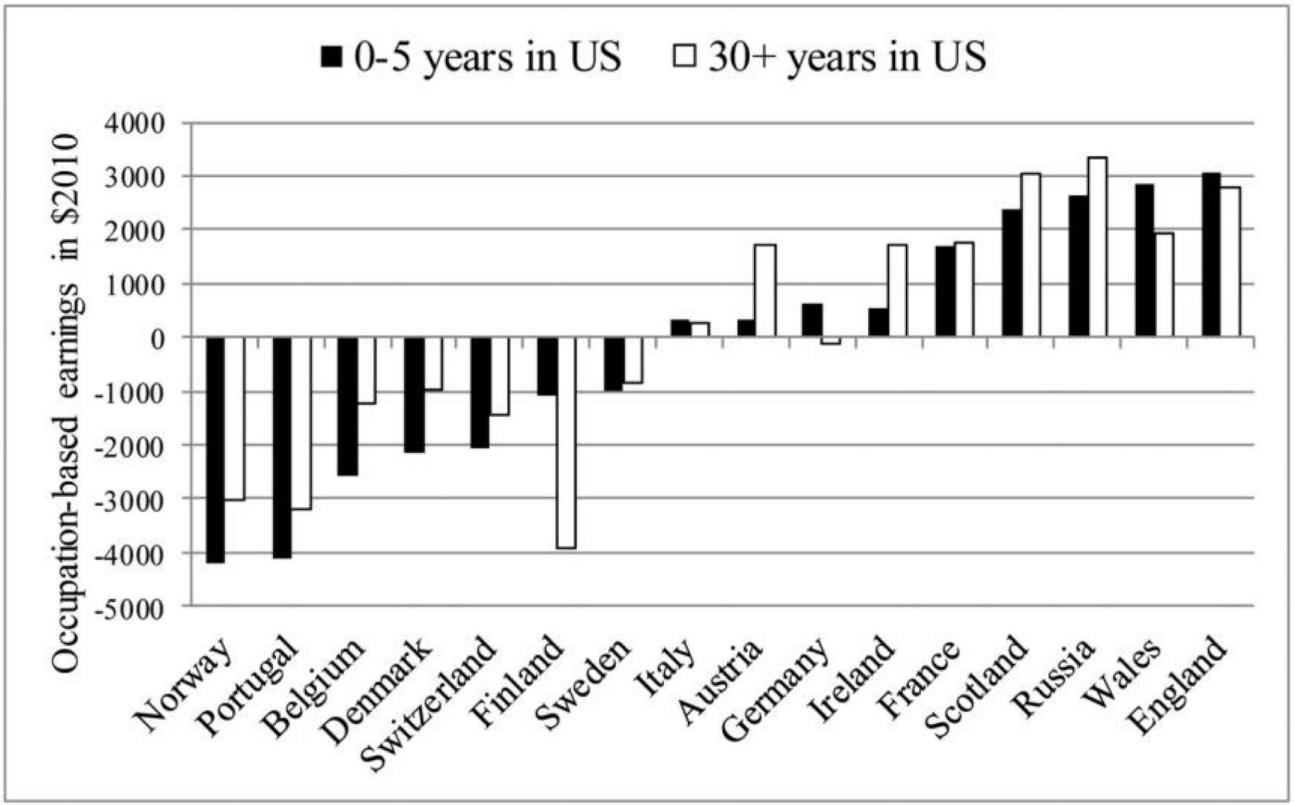
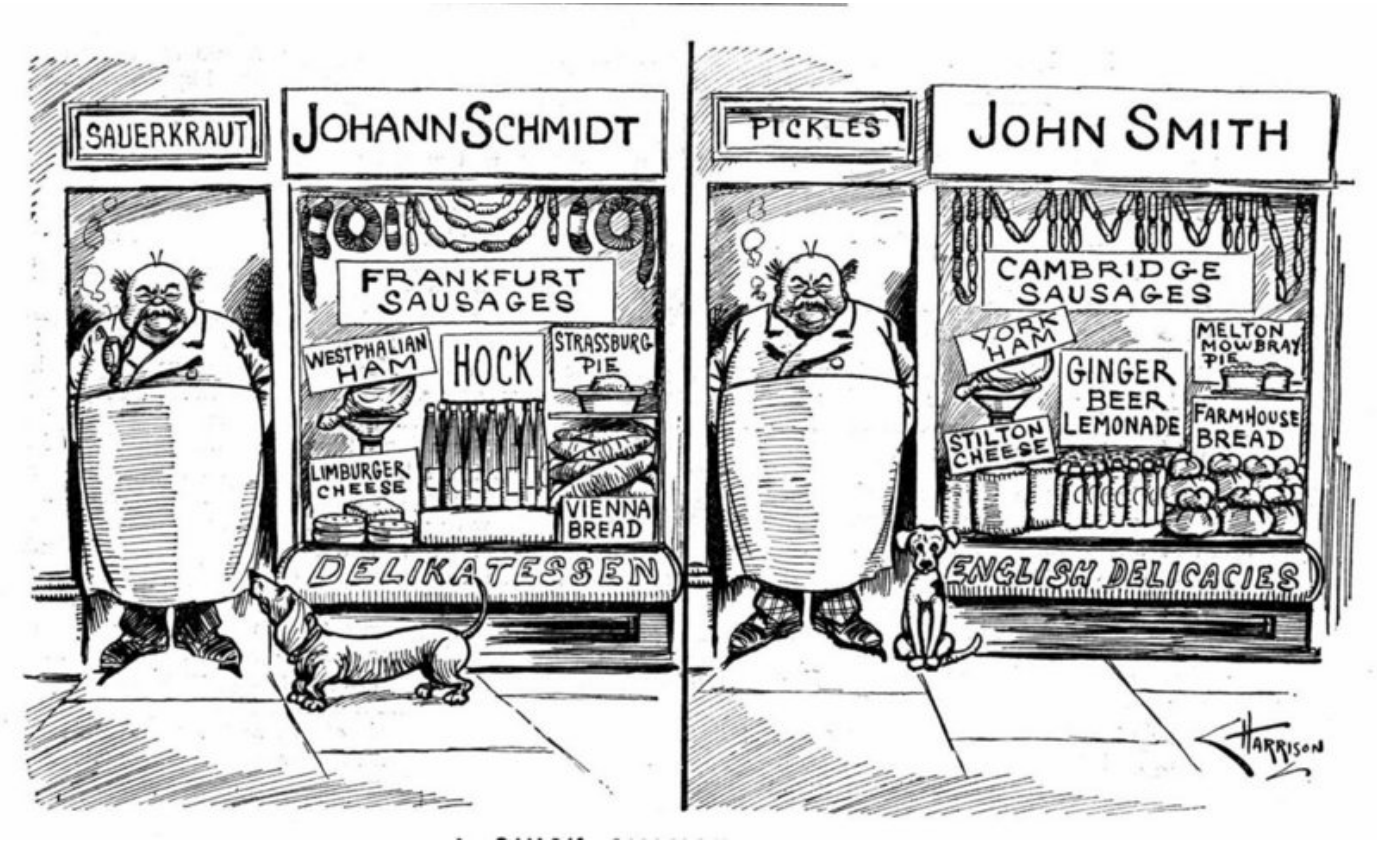


Fig. 3.
—Earnings gap between the native- and foreign-born in the panel sample: natives versus immigrants upon first arrival (0–5 years in the United States) and after time in the United States (30+ years in the United States), by country of origin. The graph reports co-efficients on the interaction between country-of-origin fixed effects and dummy variables for being in the United States for 0–5 years or for 30+ years from regression of equation (1) in the panel sample. All coefficients for the 0–5 year interaction are significant except those for Austria, Germany, Ireland, Italy, and Sweden. None of the differences between the 0–5 year and 30+ year coefficients are significant except for those of Finland and Ireland.

What Aids Assimilation?



What Aids Assimilation?

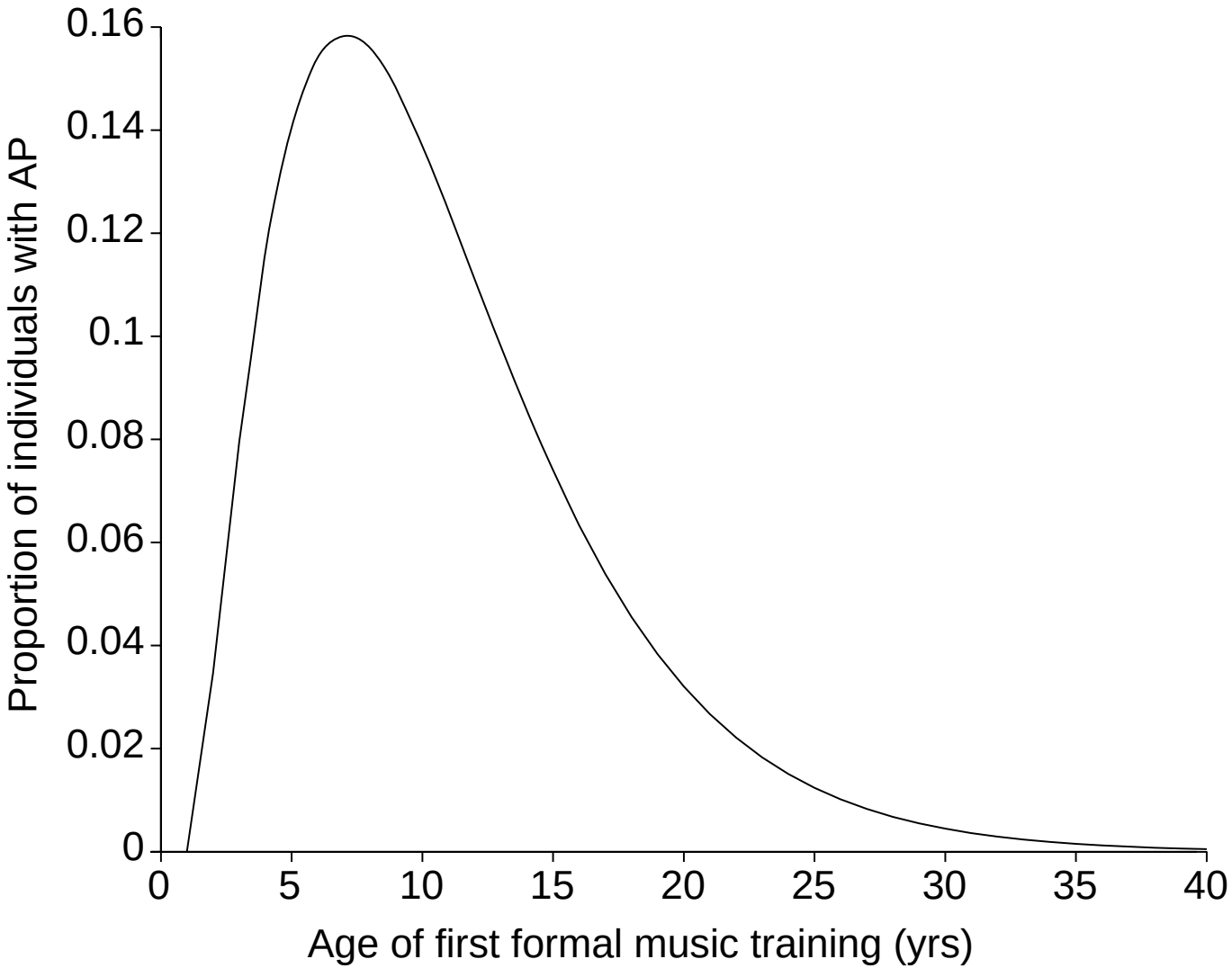
- Many things might help or hinder earnings assimilation
- We'll consider a few different things:
  - Language
  - Ethnic enclaves
  - Discrimination
  - Cultural assimilation

What Aids Assimilation?

- One critical thing may be language
- Certainly fluency in English likely helps in the US labor market
- However, we have a pretty big identification problem
- English-speaking migrants are coming from a very different set of countries than non-English-speaking migrants
- How do we tell what's due to English and what's due to other factors that differ by these countries?
- Let's take a look at the approach of Bleakley and Chin (2004)

English Fluency and Assimilation

English Fluency and Assimilation

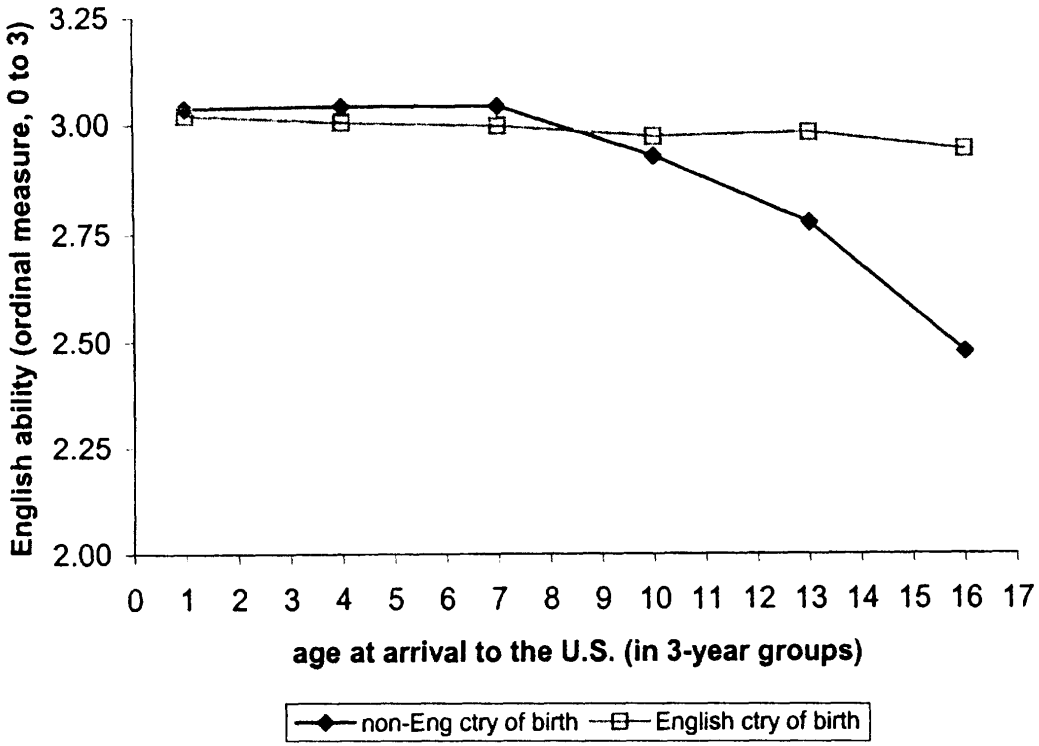


TRENDS in Cognitive Sciences

English Fluency and Assimilation

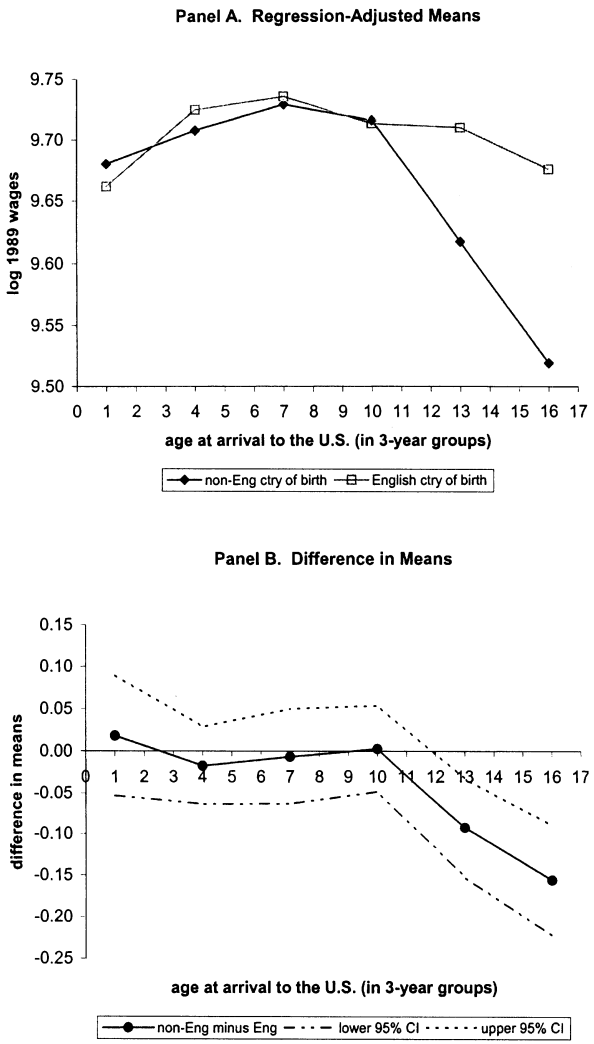
FIGURE 1.—ENGLISH-SPEAKING ABILITY BY AGE AT ARRIVAL

Panel A. Regression-Adjusted Means



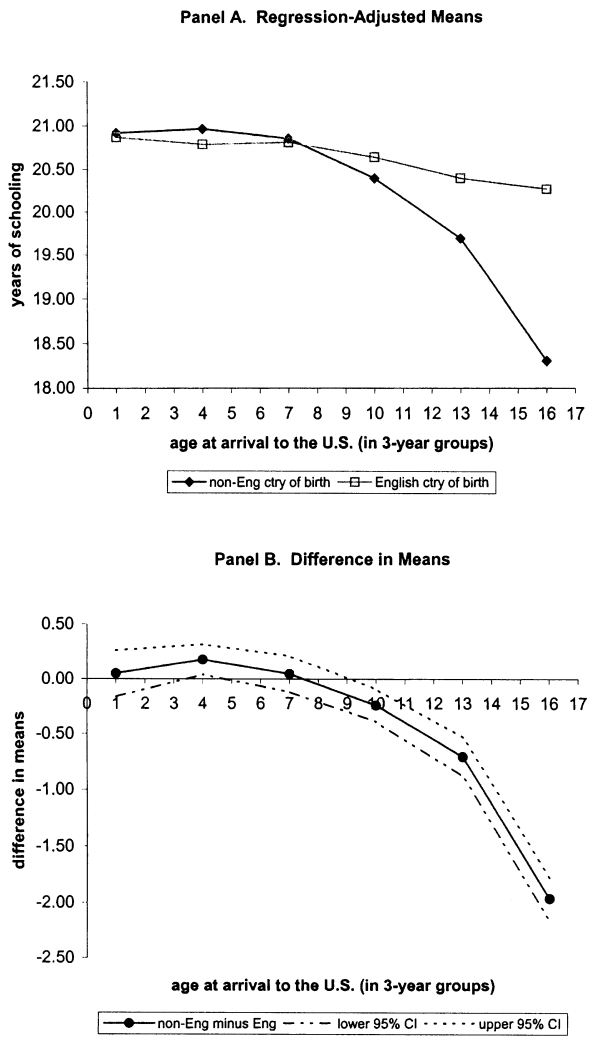
English Fluency and Assimilation

FIGURE 2.—LOG ANNUAL WAGES BY AGE AT ARRIVAL



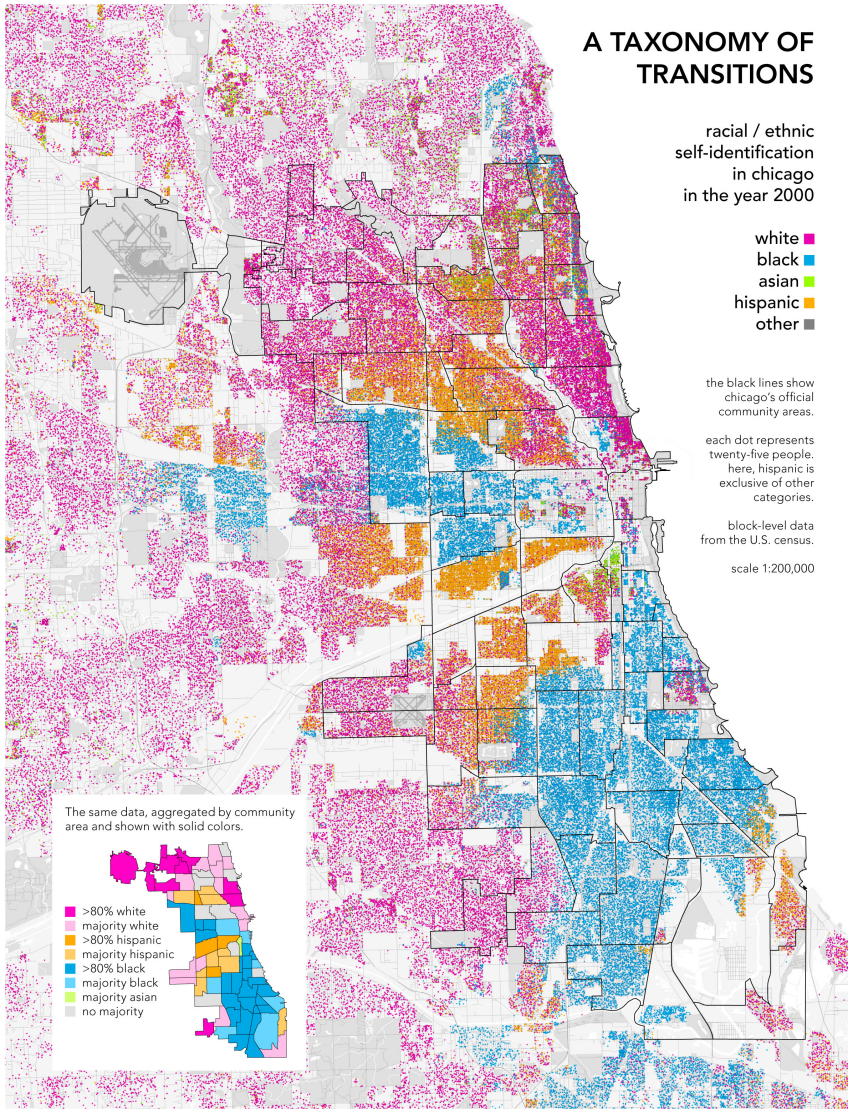
Notes: Data from 1990 IPUMS weighted by IPUMS weights. Sample size is 47,422 (composed of individuals who arrived in the United States by age 17 between 1960 and 1974 and are currently aged 25 to 38). Means have been regression-adjusted for age, race, Hispanic, and female dummies.

FIGURE 3.—YEARS OF SCHOOLING BY AGE AT ARRIVAL



Notes: Data from 1990 IPUMS weighted by IPUMS weights. Sample size is 65,214 (composed of individuals who arrived in the United States by age 17 between 1960 and 1974 and are currently aged 25 to 38). Means have been regression-adjusted for age, race, Hispanic, and female dummies.

Immigrant Enclaves and Assimilation



Immigrant Enclaves and Assimilation

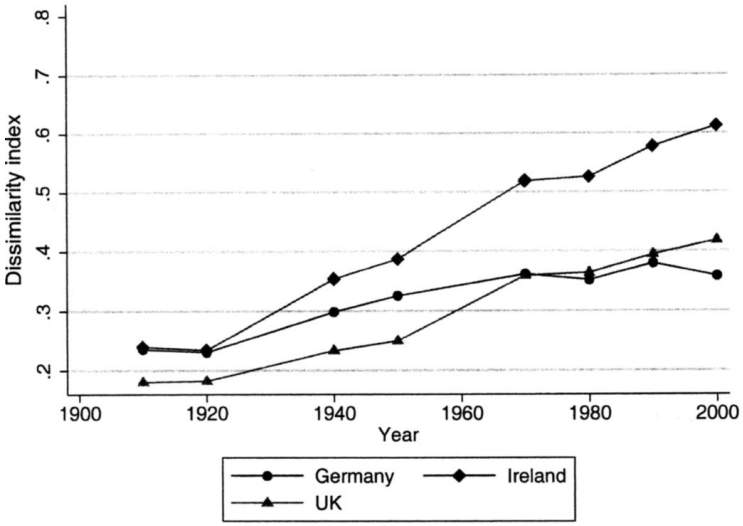
- A quick note on measuring segregation
- There are a variety of measures capturing different dimensions of segregation
- The dissimilarity index is a measure of *evenness*
- Let N_i be the number of native-born individuals in neighborhood i and F_i be the number of foreign-born individuals
- The dissimilarity index is then:

$$D = \frac{1}{2} \sum_{i=1}^{n} \left| \frac{F_i}{F_{total}} - \frac{N_i}{N_{total}} \right|$$

- A larger value corresponds to greater segregation

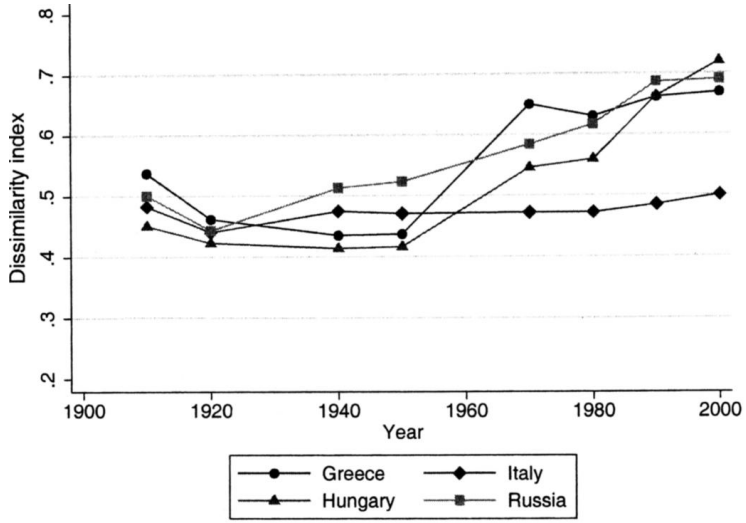
Immigrant Enclaves and Assimilation

FIGURE 2.—DISSIMILARITY FOR OLDER IMMIGRANT GROUPS, 1910–2000



Observations are weighted averages of statistics for immigrant communities, with weights equal to the number of immigrants in the community.

FIGURE 4.—DISSIMILARITY FOR NEW IMMIGRANT GROUPS, 1910–2000



Observations are weighted averages of statistics for immigrant communities, with weights equal to the number of immigrants in the community.

Immigrant Enclaves and Assimilation

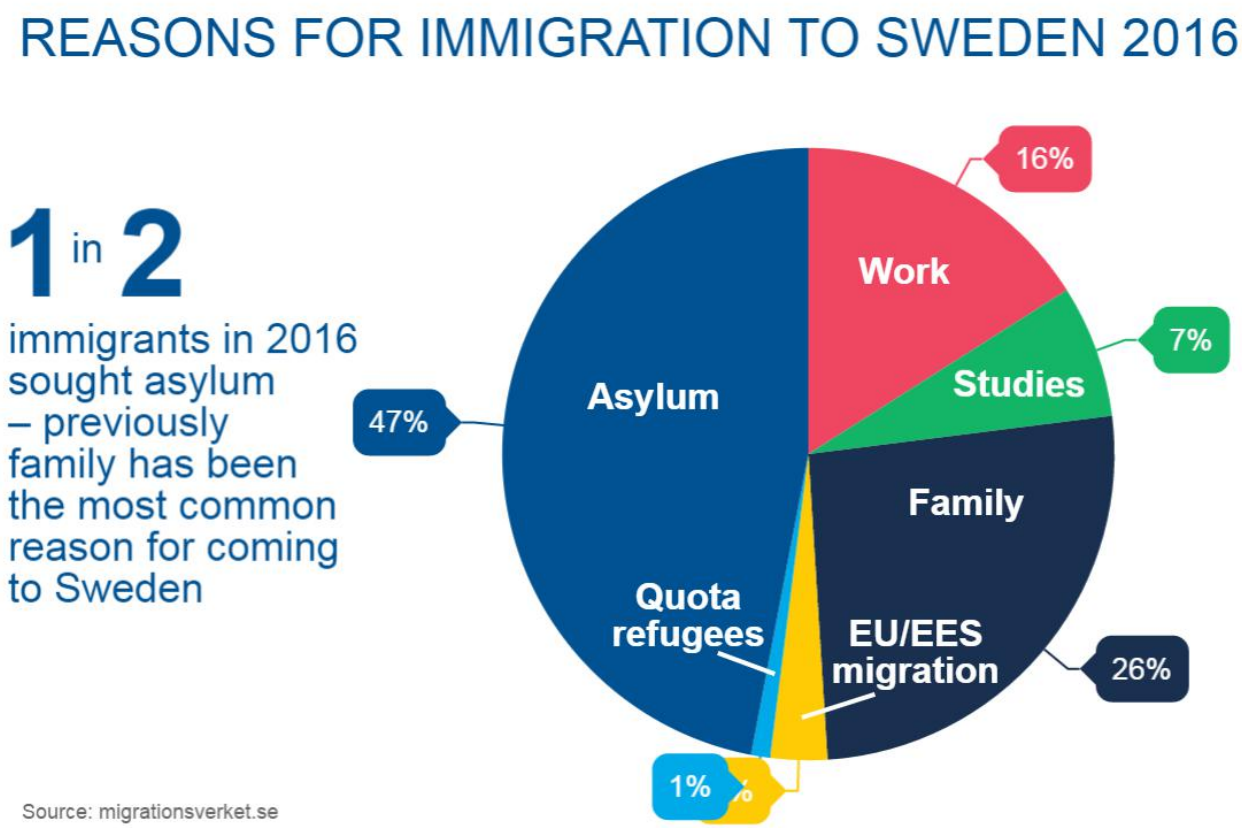
Why might immigrant enclaves matter for earnings (and other types of) assimilation?

- Slower rate of acquisition of host country skills - hinders moves to better jobs and resulting earnings growth
- Network effects - networks present employment opportunities and a means of disseminating information about the labor market
- Spatial mismatch - immigrants may be forced to segregate in an enclave far from employment opportunities
- Human capital externalities - if an enclave has a high stock of human capital, this may benefit new arrivals

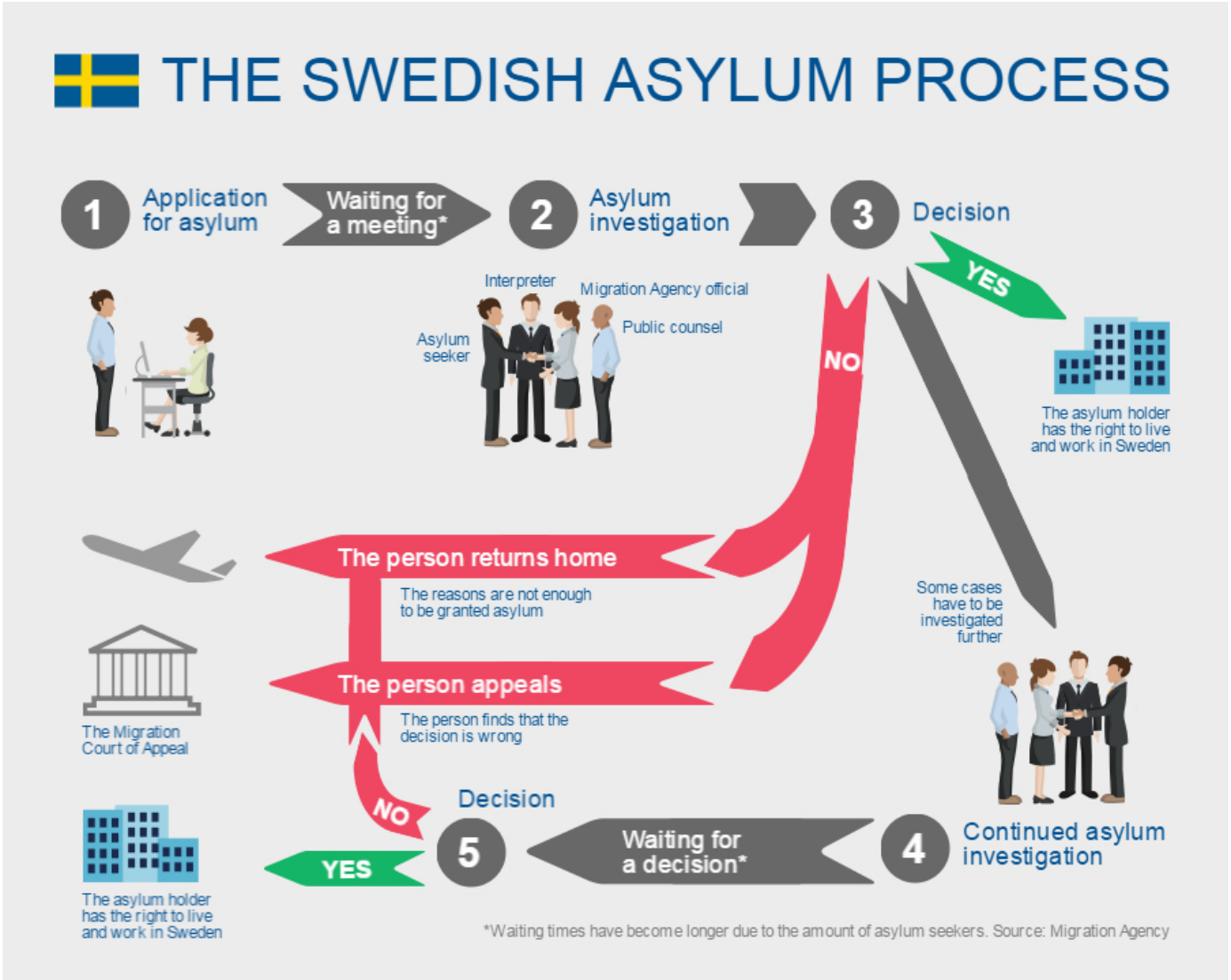
Immigrant Enclaves and Assimilation

- So immigrant enclaves might matter for all kinds of reasons
- But if a large share of immigrants live in enclaves, how do we disentangle the effects of individual characteristics versus enclave influence?
- Can we compare immigrants living in enclaves to those not living in enclaves?
- Probably not a good idea if there is self-selection into enclaves
- Let's take a look at the approach of Edin, Fredricksson and Aslund (2003)

Immigrant Enclaves and Assimilation



Immigrant Enclaves and Assimilation



Immigrant Enclaves and Assimilation

TABLE I

INDIVIDUAL CHARACTERISTICS BY INITIAL PLACEMENT

	Initial placement	
	Enclave	No enclave
Female	.44	.45
Age	37.3	37.6
	(7.7)	(7.4)
Years of schooling	11.3	11.7
	(3.0)	(2.9)
Married	.63	.62
Kid $\leq$15 years of age	.55	.57
No. of individuals	3094	3324

Standard deviations are in parentheses. An enclave is defined as described in the main text. Years of schooling are imputed from highest degree attained. Individuals with missing information on education were given the same number of years of schooling as those with less than nine years of schooling. All characteristics are measured eight years after immigration. The sample is restricted to those with positive earnings at that point in time.

Immigrant Enclaves and Assimilation

TABLE III
BASELINE ESTIMATES—DEPENDENT VARIABLE: ln(EARNINGS)

	Full sample		Low education (10 years or less)		High education (more than 10 years)	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
ln(ethnic group)	−.056	.012	−.053	.174	−.050	−.057
	(.022)	(.050)	(.024)	(.088)	(.030)	(.080)
Female	−.071	−.069	−.087	−.050	−.004	−.004
	(.081)	(.082)	(.128)	(.132)	(.098)	(.098)
Age	.066	.068	.079	.099	.054	.054
	(.023)	(.022)	(.038)	(.040)	(.030)	(.031)
Age squared (*10^{-2})	−.074	−.075	−.090	−.112	−.062	−.062
	(.028)	(.027)	(.049)	(.052)	(.036)	(.037)
Married	.210	.210	.289	.278	.168	.167
	(.084)	(.084)	(.162)	(.166)	(.072)	(.073)
Kid	−.027	−.004	−.115	−.050	.083	.081
	(.075)	(.082)	(.132)	(.138)	(.086)	(.102)
Married*female	−.049	−.032	−.226	−.207	.012	.011
	(.100)	(.100)	(.153)	(.162)	(.106)	(.106)
Kid*female	−.262	−.278	−.144	−.223	−.391	−.389
	(.125)	(.124)	(.222)	(.229)	(.137)	(.139)
Education missing and <9 years	Ref.	Ref.	Ref.	Ref.		
9–10 years	.078	.077	.097	.084		
	(.060)	(.059)	(.069)	(.070)		
High school ≤2 years	.204	.209			Ref.	Ref.
	(.088)	(.087)				
High school >2 years	.196	.204			−.013	−.013
	(.070)	(.069)			(.081)	(.081)
University <3 years	.181	.180			.006	.007
	(.071)	(.070)			(.072)	(.072)
University ≥3 years	.525	.526			.341	.341
	(.081)	(.082)			(.076)	(.076)

Immigrant Enclaves and Assimilation

TABLE IV
THE "QUALITY" OF ENCLAVES—DEPENDENT VARIABLE: ln(EARNINGS)

	Full sample		Low education (10 years or less)		High education (more than 10 years)	
	(1) IV	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV
ln(ethnic group)	−.221	−.138	−.031	.027	−.315	−.267
	(.109)	(.071)	(.220)	(.114)	(.161)	(.113)
ln(ethnic group)*ethnic inc	.044		.039		.047	
	(.015)		(.040)		(.019)	
ln(ethnic group)*ethnic self-employment rate		3.212		2.964		4.592
		(.926)		(1.589)		(1.309)
Female	−.068	−.071	−.049	−.039	−.011	−.018
	(.081)	(.080)	(.130)	(.133)	(.099)	(.098)
Age	.065	.069	.095	.096	.053	.056
	(.022)	(.023)	(.040)	(.041)	(.031)	(.031)
Age squared (*10^{-2})	−.072	−.077	−.108	−.110	−.061	−.065
	(.027)	(.028)	(.052)	(.053)	(.037)	(.038)
Married	.204	.203	.287	.290	.147	.143
	(.088)	(.084)	(.168)	(.162)	(.076)	(.073)
Kid	−.022	−.025	−.068	−.080	.057	.059
	(.083)	(.082)	(.140)	(.134)	(.106)	(.102)
Married*female	−.044	−.046	−.216	−.236	.002	.008
	(.101)	(.102)	(.163)	(.152)	(.108)	(.111)
Kid*female	−.264	−.264	−.213	−.215	−.367	−.366
	(.126)	(.127)	(.228)	(.228)	(.141)	(.141)

Immigrant Enclaves and Assimilation

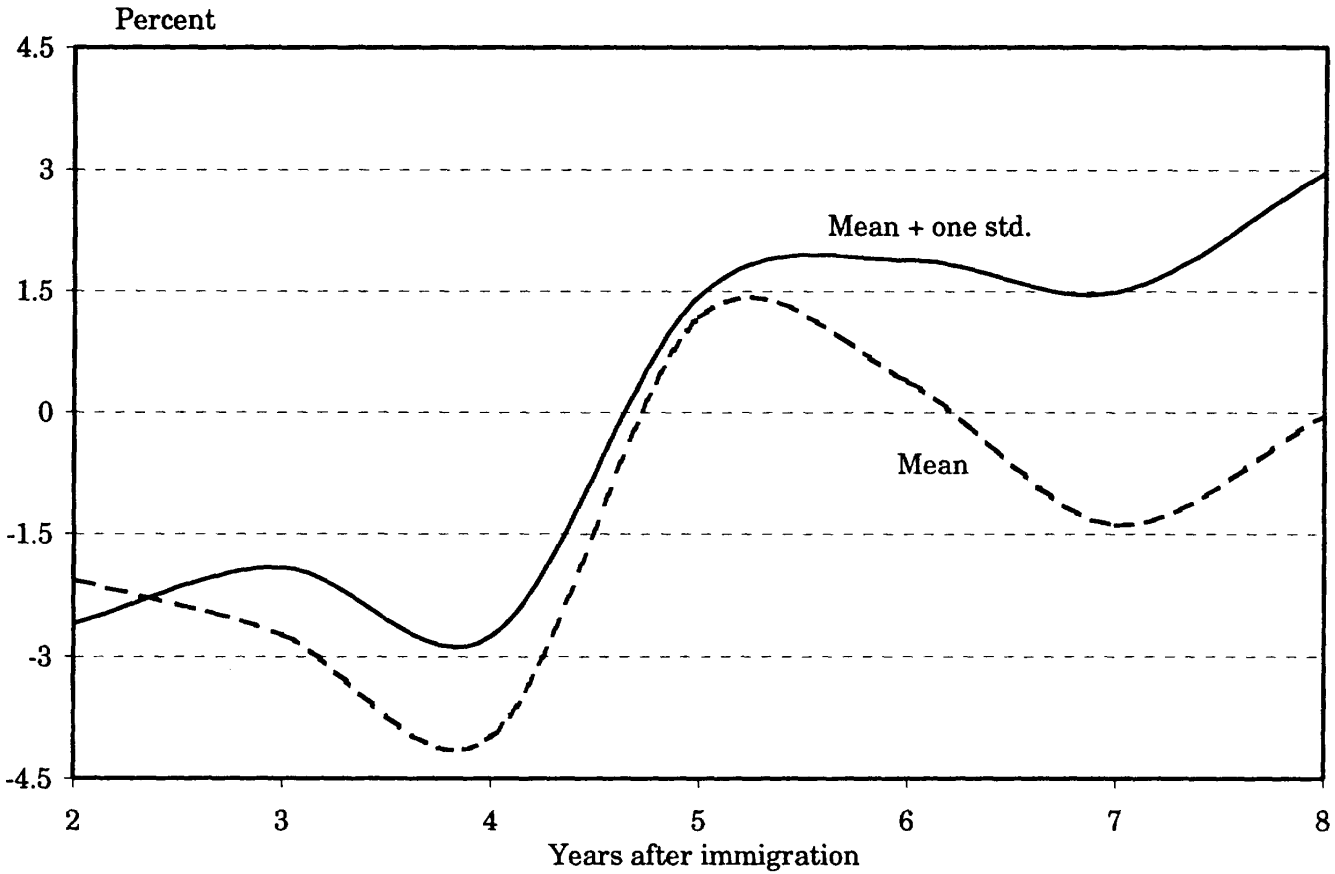


FIGURE II

The Reduced-Form Earnings Effect of Ethnic Concentration over Time

The figure is based on Table V. The dashed line shows the effect evaluated at mean ethnic income; the solid line shows the effect evaluated at one standard deviation above the mean.

Immigrant Outcomes and Discrimination

- The immigrant enclaves raise the possibility of discrimination in the housing market influencing immigrant outcomes
- What about discrimination against immigrants in labor markets?
- If discrimination exists, it could hinder earnings convergence (or, more generally, earnings growth) after migration
- One problem is that it will be hard to find variation in discrimination within a particular ethnic group
- One solution to this problem is presented by Moser (2012)

The German American Experience

TABLE I

CONSECUTIVE ORDER OF TEN ETHNIC VARIETIES IN THE POPULATION OF THE UNITED STATES ACCORDING TO THEIR MEAN RATING IN TEN SELECTED TRAITS

	Order of Native White Americans	Order of Germans	Order of English	Order of Polish and Russian Hebrews	Order of Scandinavians	Order of Irish	Order of French Canadians	Order of Austrian Slavs	Order of South Italians	Order of Negroes
Physical vigor	3	2	5	8	1	4	6	7	9	10
Intellectual ability	1	2	4	3	5	6	8	7	9	10
Self-control	3	1	2	5	4	7	6	8	9	10
Moral integrity	4	1	2	5	3	8	7	6	9	10
Sympathy	6	4	10	5	9	1	8	7	3	2
Co-operation	1	2	3	5	6	4	8	7	9	10
Leadership	1	4	2	6	5	3	7	8	9	10
Perseverance	4	1	3	2	5	8	7	6	9	10
Efficiency	1	2	3	4	5	6	7	8	9	10
Aspiration	2	4	3	1	6	5	7	8	9	10
All qualities	1	2	3	4	5	6	7	8	9	10

Survey results from Woolston, *American Journal of Sociology* (1916)

The German American Experience

The German American Experience

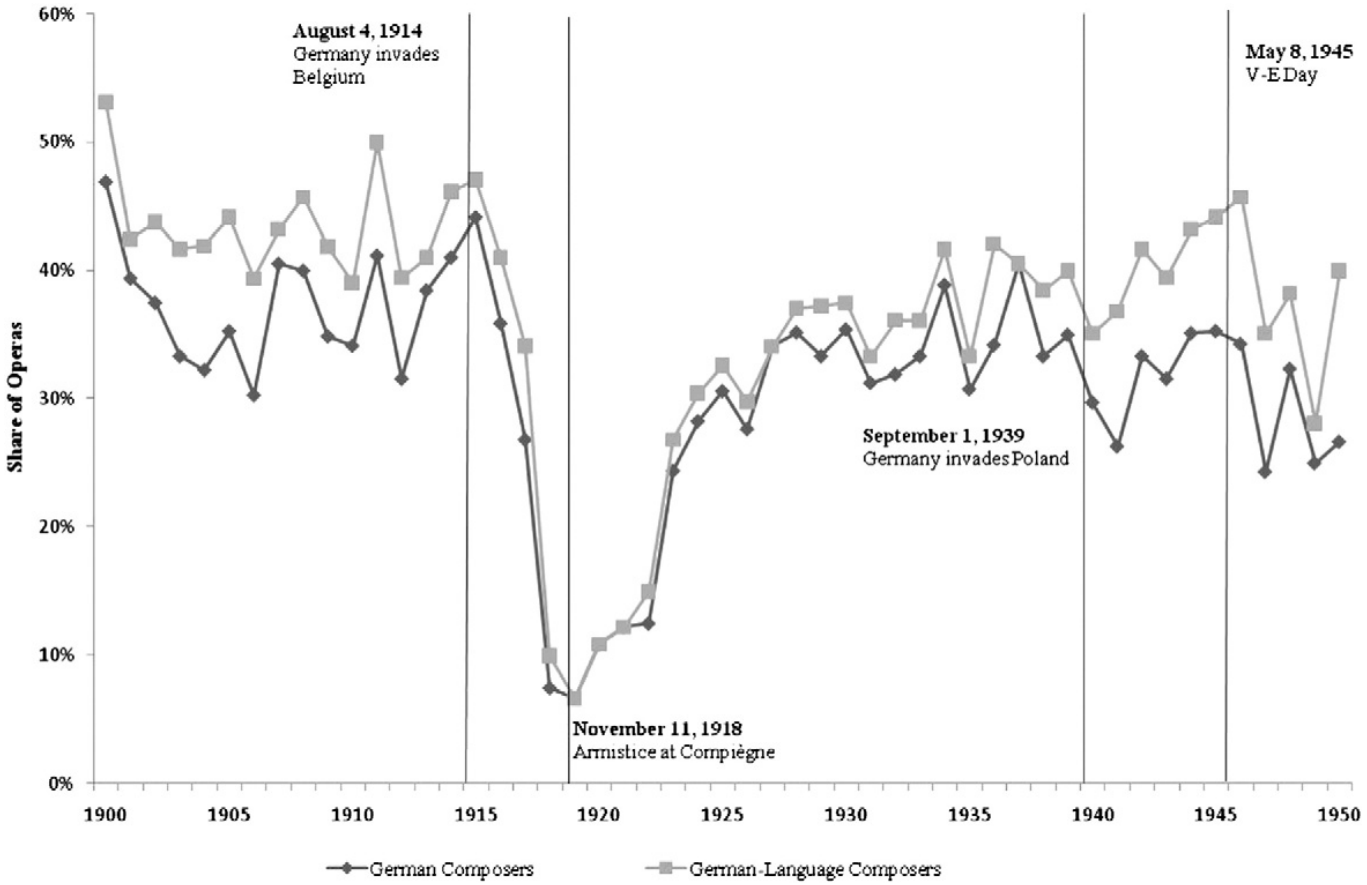


Fig. 1. The share of German-language operas from 1900 to 1950. *Notes*: Data on operas are collected from historical schedules of performances in the online archives of the Metropolitan Opera in New York. German-language composers include Austrian and Bohemian composers.

The German American Experience

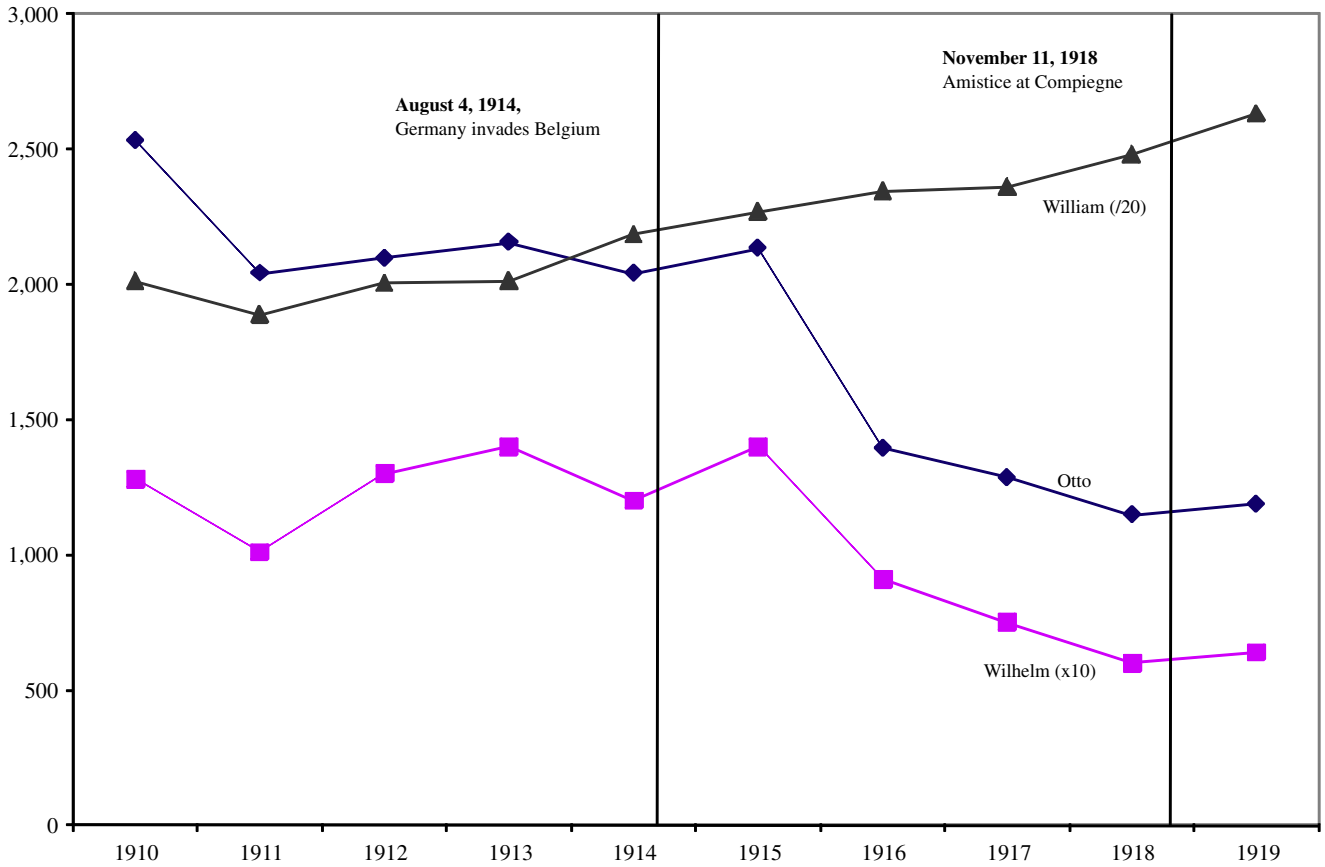


Fig. 2. Boys named Otto, Wilhelm, and William from 1910 to 1919. *Notes*: Data are constructed by counting the number of children with the name Otto or Wilhelm born between 1910 and 1919 and recorded in the United States Census of 1920. To scale the series in one graph, the number of Wilhelms is multiplied by 10 and the number of Williams is divided by 20.

The German American Experience

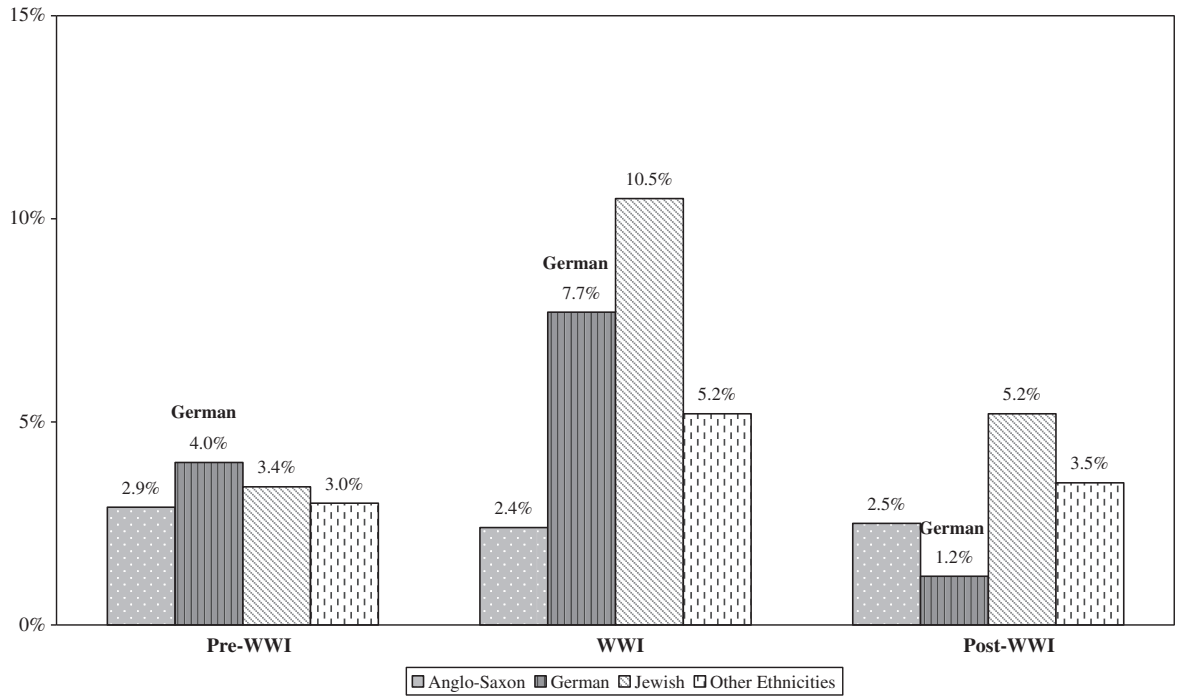


Fig. 5. Rejected applicants by ethnicity, 1883–1936. *Notes*: Data on names and election outcomes are collected at the archives of the NYSE. Names are matched to ethnicities by a commercial algorithm that uses linguistic rules and location-specific naming practices.

The German American Experience

Table 3
Means of rejection rates and blackballs by ethnicity, *Pre-war* versus *War*.

	Panel A: Rejections						Panel B: Number of blackballs					
	(Non-Jewish) German			Jewish German			(Non-Jewish) German			Jewish German		
	War	Pre-War	Difference	War	Pre-War	Difference	War	Pre-War	Difference	War	Pre-War	Difference
German American	0.077	0.040	0.037	0.087	0.024	0.063	0.966	0.391	0.575	0.809	0.190	0.619
	(0.016)	(0.014)	(0.021)	(0.020)	(0.018)	(0.027)	(0.179)	(0.109)	(0.210)	(0.226)	(0.137)	(0.264)
Anglo-Saxon	0.024	0.029	−0.005	0.024	0.029	−0.005	0.304	0.345	−0.041	0.304	0.345	−0.041
	(0.006)	(0.005)	(0.008)	(0.006)	(0.005)	(0.008)	(0.073)	(0.040)	(0.084)	(0.071)	(0.039)	(0.081)
Difference	0.053	0.011	0.042*	0.063	−0.005	0.068**	0.662	0.046	0.616***	0.505	−0.155	0.660**
	(0.017)	(0.015)	(0.023)	(0.021)	(0.019)	(0.028)	(0.194)	(0.116)	(0.226)	(0.236)	(0.143)	(0.277)

Notes: The p-value for the difference in differences in rejection rates is 0.068 for non-Jewish German Americans and 0.016 for Jewish German Americans. The p-value for the difference in differences in blackballs is 0.007 for non-Jewish German Americans and 0.017 for Jewish German Americans. German American refers to all U.S. citizens with German-sounding names. Data on admissions decisions and on the names of applicants were collected from the NYSE Archives. Names are matched to ethnicities by a commercial algorithm that takes advantage of linguistic rules and location-specific naming practices. This algorithm groups Germans Jews together with other Jewish Americans. To identify German Jews, Jewish applicants are assigned to the most frequent country of origin for immigrants with their last name in the arrival records of ships entering New York between 1850 and 1950. Standard errors in parentheses are based on a linear probability regression of rejection probabilities on ethnicities.

* $p<0.10$.
** $p<0.05$.
*** $p<0.01$.

What Aids Assimilation?

- Moser's study points to particularly large levels of taste-based discrimination against immigrant groups
- One way to get around this taste-based discrimination is to try to appear more 'American'
- We can find evidence of this in naming practices
- Let's take a quick look at Abramitzky, Boustan and Erickson (2016)

Measuring Names

$$R = \frac{\frac{n_i^F}{n_{\text{total}}^F}}{\frac{n_i^N}{n_{\text{total}}^N}}$$

- R: relative probability that a name is held by a foreign person
- n_i^F: number of foreign-born individuals with name
- n_{total}^F: total number of foreign-born individuals
- n_i^N: number of native-born individuals with name
- n_{total}^N: total number of native-born individuals

Measuring Names

$$F = 100 \cdot \frac{\frac{n_i^F}{n_{\text{total}}^F}}{\frac{n_i^F}{n_{\text{total}}^F} + \frac{n_i^N}{n_{\text{total}}^N}}$$

- F: foreignness index
- Goes to 100 if only foreign-born individuals have a certain name
- Goes to 0 if no foreign-born individuals have a certain name
- Goes to 50 if same percentage of foreign-born and native-born individuals have the name

What Aids Assimilation?

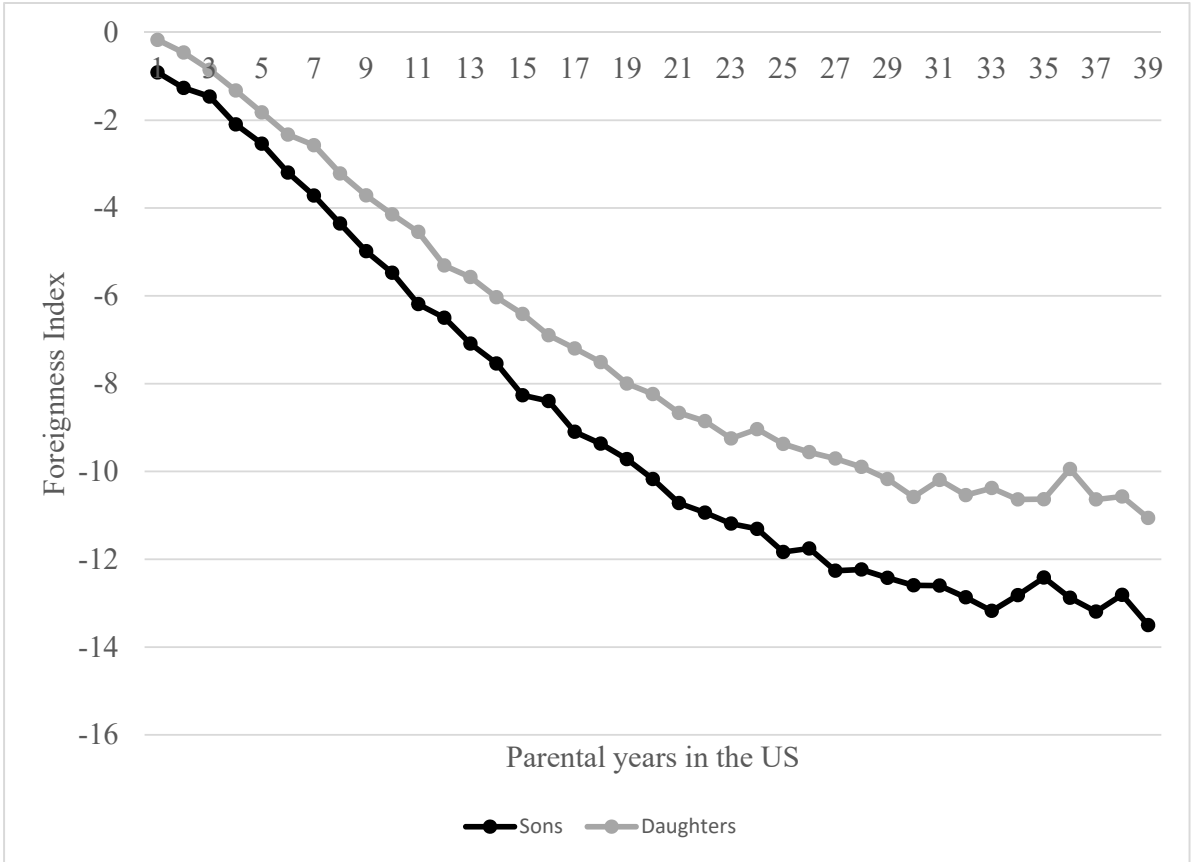
Table 1: Examples of foreign, neutral, and native names (1900-1920 birth cohorts)

Most foreign (F-Index >0.90)	Most neutral (0.5 < F-Index < 0.52)	Most native (F-Index <0.025)
	A. Male names	
Vito	Orlando	Gaylord
Mario	Benjiman	Doyle
Hyman	Murray	Clay
Pasquale	Otto	Lowell
Isidor	Theodor	Dale
Nick	Herman	Wayne
	B. Female names	
Sonia	Margaret	Bethany
Antoinette	Deborah	Merlene
Concetta	Helene	Garnet
Johanna	Kathleen	Arlyce
Molly	Beatrice	Joellen
Carmela	Fay	Opal

Notes: Names with 100 or more observations selected for having high/lowest/most neutral F-index values in 1920 complete-count Census for the birth cohorts of 1900-20.

What Aids Assimilation?

Figure 4: Immigrants selected less foreign names for children after spending time in US, (Dependent variable = F-index)



Notes: The graph reports coefficients from estimates of Equation 1, a regression of the F-index on a set of dummy variables for years that the household head had spent in the US by the time of the child's birth. Regressions also include dummy variables for child's age and a set of household fixed effects. Data from 1920 complete-count Census. Sample includes children aged 0-18 who were born in a non-southern state and are living with their parents. Households must have a foreign-born head and the spouse (mother) must be less than 43 years old (N (sons) = 2,130,352; N (daughters) = 2,081,724).

What Aids Assimilation?

Table 6: The effect of name foreignness on education, earnings and unemployment

	No household fixed effects			With household fixed effects	
	(1) Baseline	(2) Add controls	(3) Add F-index at 20	(4) Full Sample	(5) Brothers 1-2 years apart
	Panel A				
	Dependent variable: Highest grade (Mean = 10.26)				
F-index	-0.009***	-0.009***	---	-0.006***	-0.008***
	(0.0001)	(0.0001)		(0.0003)	(0.0007)
N	1,054,765	972,211		972,211	168,515
	Panel B				
	Dependent variable: =1 if unemployed *x* 100 (Mean = 9.5)				
F-index	0.026***	0.026***	0.015***	0.027***	0.017***
	(0.001)	(0.002)	(0.005)	(0.003)	(0.008)
F-index at 20			0.012**		
			(0.005)		
N	988,383	910,936	910,936	910,936	157,531
	Panel C				
	Dep. Variable: Annual earnings (Mean=$21,057)				
F-index	-10.61***	-12.65***	-2.67	-6.51**	-13.67**
	(0.910)	(0.958)	(2.88)	(2.70)	(61.48)
F-index at 20			-11.04***		
			(3.00)		
N	673,810	620,413	620,413	620,413	107,045

Note: Sample includes men matched between 1920 and 1940 complete-count Censuses. Men must be 3-18 in 1920, born outside the South and living at home with parents in 1920 in a household whose head was foreign-born. Panel C is further restricted to men with non-zero earnings who were not self-employed in 1940. All regressions control for a vector of dummies for child's age in 1940. Columns 2-5 control for parental years in the US and child's rank in the birth order. Columns 4-5 add household fixed effects.