Econ 373: US Economic History

Taylor Jaworski

Fall 2023

▶ Between 1850 and 1914 roughly 55M people moved from Old World to New World, with about 30 million of these moving to United States

- ▶ Between 1850 and 1914 roughly 55M people moved from Old World to New World, with about 30 million of these moving to United States
- ▶ Hatton & Williamson (1998) provide an overview of literature on mass migration that mirrors work of O'Rourke & Williamson related to trade

- ▶ Between 1850 and 1914 roughly 55M people moved from Old World to New World, with about 30 million of these moving to United States
- ▶ Hatton & Williamson (1998) provide an overview of literature on mass migration that mirrors work of O'Rourke & Williamson related to trade
- ▷ Abramitzky & Boustan (2017) emphasize revisions to literature in last two decades to reflect contribution of "credibility revolution"
- Abramitzky & Boustan are also working on a book that promises to be very good

- ▶ Between 1850 and 1914 roughly 55M people moved from Old World to New World, with about 30 million of these moving to United States
- ▶ Hatton & Williamson (1998) provide an overview of literature on mass migration that mirrors work of O'Rourke & Williamson related to trade
- ▶ Abramitzky & Boustan (2017) emphasize revisions to literature in last two decades to reflect contribution of "credibility revolution"
- Abramitzky & Boustan are also working on a book that promises to be very good
- ▶ Early literature focused on placing Age of Mass Migration in context of changes in the broader Atlantic World, more recent work focused on relationship to research in labor or development economics

ightharpoonup I would highlight two broad themes in the economic history of immigration:

- ▶ I would highlight two broad themes in the economic history of immigration:
- 1. Political Economy: the integration of immigrants into society, politics, and culture; the consequences of emigration on the home country or region

- ▶ I would highlight two broad themes in the economic history of immigration:
- 1. Political Economy: the integration of immigrants into society, politics, and culture; the consequences of emigration on the home country or region
- 2. The Economics of Immigration: assimilation; the returns to migration; self-selection; the impact of the immigrants on natives; and immigrants and innovation

- ▶ I would highlight two broad themes in the economic history of immigration:
- 1. Political Economy: the integration of immigrants into society, politics, and culture; the consequences of emigration on the home country or region
- 2. The Economics of Immigration: assimilation; the returns to migration; self-selection; the impact of the immigrants on natives; and immigrants and innovation
- ➤ These themes are not wholly distinct, but we can think about economic historians and history-inclined economists as dominating #1 and economists in labor and development (and perhaps trade) dominating #2

Exit, Voice, and Loyalty

- ➤ The economist Albert Hirschman wrote the classic book Exit, Voice, and Loyalty, which
 provides a framework for studying decision-making within organizations or countries in
 response to decline, oppression, etc
- ➤ The book's title posits three responses (i.e., exit, voice, or loyalty) and then works through the implications of these choices—using big social science frameworks like this to motivate your work can be useful
- ▶ Karadja & Prawitz (2019) use Hirschman's framework explicitly to understand the causes and direction of political change in Sweden during the Age of Mass Migration
- ▶ Less explicitly, Dippel & Heblich (2021) consider the role that refugees from the 1848-49 German Revolutions played in the Civil War era in the United States

▶ Migration of any kind is a selective process

- ▶ Migration of any kind is a selective process
- ▶ In terms of the classic Roy (1951) model this means that decisions about where to locate (or what to do) are *optimizing* decisions and reflect individual characteristics and/or constraints of the economic environment—and some of these factors may be unobserved

- Migration of any kind is a selective process
- ▶ In terms of the classic Roy (1951) model this means that decisions about where to locate (or what to do) are *optimizing* decisions and reflect individual characteristics and/or constraints of the economic environment—and some of these factors may be unobserved
- ▶ In terms of attributes we can similarly think about selection as reflecting where in the origin location skill distribution migrants come from where the relative skill flow will reflect differences in the returns to skill between the origina and destination

- Migration of any kind is a selective process
- ▶ In terms of the classic Roy (1951) model this means that decisions about where to locate (or what to do) are *optimizing* decisions and reflect individual characteristics and/or constraints of the economic environment—and some of these factors may be unobserved
- ▷ In terms of attributes we can similarly think about selection as reflecting where in the origin location skill distribution migrants come from where the relative skill flow will reflect differences in the returns to skill between the origina and destination
- ▶ I suggest starting with Borjas (1987) for a now old formulation of the Roy model in the context of immigration, Chiquiar & Hanson (2005) provide a more recent formulation

- During the Age of Mass Migration, given the prevailing distribution of wages on either side of the Atlantic we would expect:
- immigration from the European "core" to be neutrally selection

- During the Age of Mass Migration, given the prevailing distribution of wages on either side of the Atlantic we would expect:
- \rightarrow immigration from the European "core" to be neutrally selection \rightarrow so why do people move?

- During the Age of Mass Migration, given the prevailing distribution of wages on either side of the Atlantic we would expect:
- \downarrow immigration from the European "core" to be neutrally selection \rightarrow so why do people move?
- immigration from the European "periphery" to be negatively selected

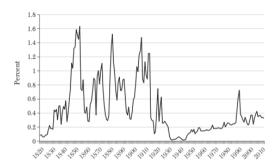
- During the Age of Mass Migration, given the prevailing distribution of wages on either side of the Atlantic we would expect:
- \downarrow immigration from the European "core" to be neutrally selection \rightarrow so why do people move?
- immigration from the European "periphery" to be negatively selected

- ▶ In historical perspective, it is interesting to consider "refugee" migrations that result from famine, persecution, etc, and the role this can play in reshaping countries
- 4 with respect to politics, diversity of demand, the stock of human capital, the technological frontier

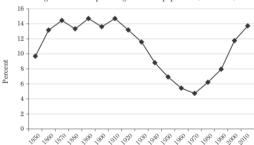
- During the Age of Mass Migration, given the prevailing distribution of wages on either side of the Atlantic we would expect:
- \downarrow immigration from the European "core" to be neutrally selection \rightarrow so why do people move?
- immigration from the European "periphery" to be negatively selected

- ▶ In historical perspective, it is interesting to consider "refugee" migrations that result from famine, persecution, etc, and the role this can play in reshaping countries
- 4 with respect to politics, diversity of demand, the stock of human capital, the technological frontier
- ➤ The role that migration costs have played is not really considered in detail (with the
 exception of noting changes in transatlantic shipping costs over time), the role of bilateral
 frictions is worth some attention

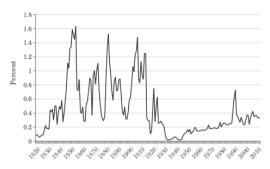
- ▶ A few stylized facts (related to selection) for the Age of Mass Migration in the US:
- ↓ time-series: flow and stock



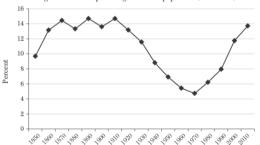
Panel B. Forign-born stock as percentage of the US population (1850–2010)

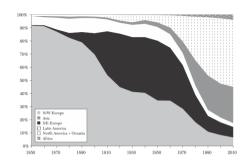


- ▶ A few stylized facts (related to selection) for the Age of Mass Migration in the US:
- time-series: flow and stock
- distribution across sending country

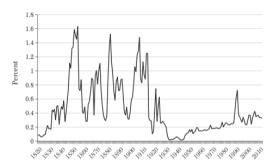


Panel B. Forign-born stock as percentage of the US population (1850–2010)

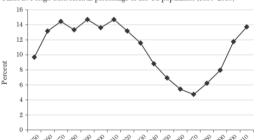


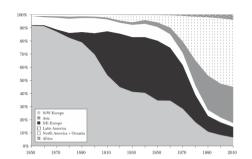


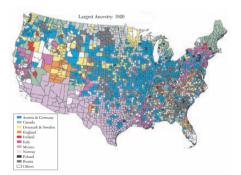
- ▶ A few stylized facts (related to selection) for the Age of Mass Migration in the US:
- time-series: flow and stock
- distribution across sending country
- distribution across US county and ethnic enclaves



Panel B. Forign-born stock as percentage of the US population (1850–2010)







- ▶ A few stylized facts (related to selection) for the Age of Mass Migration in the US:
- time-series: flow and stock
- distribution across sending country
- 4 distribution across US county and ethnic enclaves
- ▶ Abramitzky, Boustan, & Eriksson (2014) use advances in census linking to revisit the literature on assimilation and selection based on country of origin

- ▶ A few stylized facts (related to selection) for the Age of Mass Migration in the US:
- time-series: flow and stock
- distribution across sending country
- 4 distribution across US county and ethnic enclaves
- ▶ Abramitzky, Boustan, & Eriksson (2014) use advances in census linking to revisit the literature on assimilation and selection based on country of origin
- Collins & Zimran also draw on census linking to understand the determinants of assimilation from 1850 to 1880 and non-assimilation from 1900 to 1930

- ▷ A few stylized facts (related to selection) for the Age of Mass Migration in the US:
- by time-series: flow and stock
- distribution across sending country
- 4 distribution across US county and ethnic enclaves
- ▶ Abramitzky, Boustan, & Eriksson (2014) use advances in census linking to revisit the literature on assimilation and selection based on country of origin
- Collins & Zimran also draw on census linking to understand the determinants of assimilation from 1850 to 1880 and non-assimilation from 1900 to 1930
- ▶ Let's first do a quick primer on census linking (more generally applicable)

- ▶ Many kinds of historical data do not have unique identifiers that allow researchers to follow individuals over time
- work by Raj Chetty along with researchers Opportunity Insights is an example where such identifiers are available but need to be suppressed in publicly available data and replication files

- ▶ Many kinds of historical data do not have unique identifiers that allow researchers to follow individuals over time
- work by Raj Chetty along with researchers Opportunity Insights is an example where such identifiers are available but need to be suppressed in publicly available data and replication files

- ▶ A literature in economic history (due to Joe Ferrie) in the 1990s sought to create individual-level panel data to study immigration and mobility in the United States
- Ferrie's insight was that an individual's first and last name together with a few immutable characteristics could be used to identify, but there was a technological constraint on the size of samples

- ▶ Many kinds of historical data do not have unique identifiers that allow researchers to follow individuals over time
- work by Raj Chetty along with researchers Opportunity Insights is an example where such identifiers are available but need to be suppressed in publicly available data and replication files

- ▶ A literature in economic history (due to Joe Ferrie) in the 1990s sought to create individual-level panel data to study immigration and mobility in the United States
- Ferrie's insight was that an individual's first and last name together with a few immutable characteristics could be used to identify, but there was a technological constraint on the size of samples
- ▶ Today, the constraint is not technological but statistical: how to create a representative panel of individuals for the particular question of interest

- ▶ Criteria for an algorithm for census linking:
 - 1. Minimize false matches
 - 2. Maximize true matches
 - 3. Represent the population of interest
 - 4. Feasible to implement under current technological constraints

- ▶ Criteria for an algorithm for census linking:
 - 1. Minimize false matches
 - 2. Maximize true matches
 - 3. Represent the population of interest
 - 4. Feasible to implement under current technological constraints
- ▶ There are different approaches in the literature, I'll discuss a common one here and provide references to the literature at the end of this section

- ▶ Criteria for an algorithm for census linking:
 - 1. Minimize false matches
 - 2. Maximize true matches
 - 3. Represent the population of interest
 - 4. Feasible to implement under current technological constraints
- ▶ There are different approaches in the literature, I'll discuss a common one here and provide references to the literature at the end of this section
- ▶ Caveat: the commonly used approaches have tended to work best (or only) for white men for the United States, matching women and Black Americans has proved challenging

1. Clean names in datasets A and B to remove non-alphabetic characters, account for common mis-spellings and nicknames

1. Clean names in datasets A and B to remove non-alphabetic characters, account for common mis-spellings and nicknames

2. Restrict to people who are unique by first and last name, implied birth year, and place of birth (either state or country) in dataset A

1. Clean names in datasets A and B to remove non-alphabetic characters, account for common mis-spellings and nicknames

2. Restrict to people who are unique by first and last name, implied birth year, and place of birth (either state or country) in dataset A

3. For each record in dataset A, look for records in dataset B that match on first name, last name, place of birth, and exact birth year

1. Clean names in datasets A and B to remove non-alphabetic characters, account for common mis-spellings and nicknames

2. Restrict to people who are unique by first and last name, implied birth year, and place of birth (either state or country) in dataset A

3. For each record in dataset A, look for records in dataset B that match on first name, last name, place of birth, and exact birth year

 \triangleright From here there are three possibilities: (i) take unique matches, (ii) drop multiple matches, (iii) expand birth year cutoff ± 1 , ± 2 and take only unique matches

- ▶ Variants include:
- 4 using NYSIIS (New York State Identification and Intelligence System) standardized names
- using Jaro-Winkler string distance to measure the similarity of two strings
- using middle names/initials as a linking characteristic for those who have them
- using a fully automated probabilistic approach, formally balancing similarity across many measures
- using a machine-learning algorithm that takes as an input "hand-linked" records
- using FamilySearch to augment automated matches

- ▶ Variants include:
- 4 using NYSIIS (New York State Identification and Intelligence System) standardized names
- using Jaro-Winkler string distance to measure the similarity of two strings
- using middle names/initials as a linking characteristic for those who have them
- busing a fully automated probabilistic approach, formally balancing similarity across many measures
- using a machine-learning algorithm that takes as an input "hand-linked" records
- using FamilySearch to augment automated matches
- ▶ In population-to-population matching automated methods perform well, but it is worth considering the performance (i.e., exclusion restrictions) for any particular application

Practical Aspects of Census Linking

▶ Literature (how to):

- → Bailey et al, "How Well Do Automated Methods Linking Perform?" *JEL* (2020)
- Abramitzky et al, "Automated Linking of Historical Data," *JEL* (forthcoming)
- ▶ Price et al, "Combining Family History and Machine Learning to Link Historical Records"

Practical Aspects of Census Linking

▶ Literature (how to):

- → Bailey et al, "How Well Do Automated Methods Linking Perform?" *JEL* (2020)
- Abramitzky et al, "Automated Linking of Historical Data," *JEL* (forthcoming)
- ▶ Price et al, "Combining Family History and Machine Learning to Link Historical Records"

▶ Literature (critical):

Bodenhorn et al, "Sample-Selection Biases and the Industrialization Puzzle," JEH (2017)

Practical Aspects of Census Linking

▶ Literature (how to):

- → Bailey et al, "How Well Do Automated Methods Linking Perform?" *JEL* (2020)
- Abramitzky et al, "Automated Linking of Historical Data," *JEL* (forthcoming)
- Price et al, "Combining Family History and Machine Learning to Link Historical Records"

▶ Literature (critical):

Bodenhorn et al, "Sample-Selection Biases and the Industrialization Puzzle," JEH (2017)

Resources:

- ↓ IPUMS.org for complete count census data, 1850-1940

- ▶ Abramitzky, Boustan, & Eriksson (ABE) construct data covering native-born and immigrant workers to the United States from 1900, 1910, and 1920 censuses
- including immigrants from 16 sending countries and information on individual characteristics

▶ The paper's objective is quantify "assimilation," i.e., the convergence immigrants to native-born workers with respect to some characteristic

- ▶ Abramitzky, Boustan, & Eriksson (ABE) construct data covering native-born and immigrant workers to the United States from 1900, 1910, and 1920 censuses
- including immigrants from 16 sending countries and information on individual characteristics

➤ The paper's objective is quantify "assimilation," i.e., the convergence immigrants to native-born workers with respect to some characteristic → occupational income

- ▶ Abramitzky, Boustan, & Eriksson (ABE) construct data covering native-born and immigrant workers to the United States from 1900, 1910, and 1920 censuses
- including immigrants from 16 sending countries and information on individual characteristics

- ➤ The paper's objective is quantify "assimilation," i.e., the convergence immigrants to native-born workers with respect to some characteristic → occupational income
- the key empirical issue is that assimilation has to be constructed from data that includes comparisons of people who arrived in different years or the possibility of return migration

- ▶ Abramitzky, Boustan, & Eriksson (ABE) construct data covering native-born and immigrant workers to the United States from 1900, 1910, and 1920 censuses
- 4 including immigrants from 16 sending countries and information on individual characteristics

- ➤ The paper's objective is quantify "assimilation," i.e., the convergence immigrants to native-born workers with respect to some characteristic → occupational income
- the key empirical issue is that assimilation has to be constructed from data that includes comparisons of people who arrived in different years or the possibility of return migration

- ▶ ABE's solution is construct a panel of individuals using census linking
- the result is a sample of 20,225 immigrants and 1,650 native-born workers

Evaluating Census Linking

SAMPLE SIZES AND MATCH RATES BY PLACE OF BIRTH

Country	1900 Number in Universe (1)	Number Matched (2)	Match Rate, Total (3)	1900 Number, Unique (4)	Match Rate Unique (5)				
	A. 1900 Source: IPUMS								
Austria	4,835	339	.070	4,677	.072				
England	7,438	664	.089	6,175	.107				
France	11,615	728	.063	9,139	.079				
Germany	19,855	2,248	.113	16,733	.134				
Ireland	9,737	861	.088	6,323	.136				
Italy	7,624	811	.106	7,042	.115				
Norway	3,541	425	.120	2,822	.151				
Russia	5,804	644	.111	5,203	.124				
Sweden	6,164	559	.091	4,070	.137				
US natives	10,000	1,650	.165	8,345	.197				
	B. 1900 Source: Ancestry.com								
Belgium	6,060	545	.090	5,962	.091				
Denmark	34,594	1,980	.058	17,425	.114				
Finland	23,843	828	.035	22,197	.037				
Portugal	12,585	584	.046	8,362	.070				
Scotland	53,091	4,349	.082	15,529	.280				
Switzerland	22,276	3,311	.149	20,588	.161				
Wales	17,767	1,342	.076	9,876	.135				

Sample is restricted to white men ages 18 to 35 in 1900 and living outside the South

▶ Immigrants are restricted to have arrived between 1880 and 1900

 Comparing matching based on samples to complete counts from Ancestry.com

Evaluating Census Linking

0	0			n	Th	Th
SAMPLE	SIZES	AND	MATCH	RATES BY	PLACE OF	BIRTH

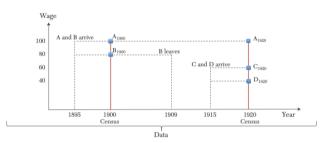
Country	1900 Number in Universe (1)	Number Matched (2)	Match Rate, Total (3)	1900 Number, Unique (4)	Match Rate Unique (5)				
	A. 1900 Source: IPUMS								
Austria	4,835	339	.070	4,677	.072				
England	7,438	664	.089	6,175	.107				
France	11,615	728	.063	9,139	.079				
Germany	19,855	2,248	.113	16,733	.134				
Ireland	9,737	861	.088	6,323	.136				
Italy	7,624	811	.106	7,042	.115				
Norway	3,541	425	.120	2,822	.151				
Russia	5,804	644	.111	5,203	.124				
Sweden	6,164	559	.091	4,070	.137				
US natives	10,000	1,650	.165	8,345	.197				
	B. 1900 Source: Ancestry.com								
Belgium	6,060	545	.090	5,962	.091				
Denmark	34,594	1,980	.058	17,425	.114				
Finland	23,843	828	.035	22,197	.037				
Portugal	12,585	584	.046	8,362	.070				
Scotland	53,091	4,349	.082	15,529	.280				
Switzerland	22,276	3,311	.149	20,588	.161				
Wales	17,767	1,342	.076	9,876	.135				

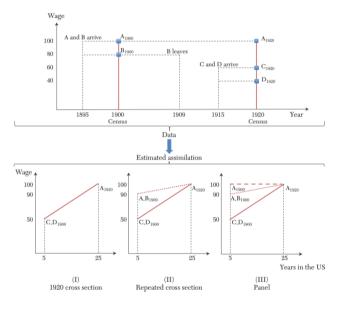
Sample is restricted to white men ages 18 to 35 in 1900 and living outside the South

▶ Immigrants are restricted to have arrived between 1880 and 1900

 Comparing matching based on samples to complete counts from Ancestry.com

▶ What is the advantage of panel data relative to the alternatives?





▶ The main analysis compares occupational mobility of native-born and immigrant workers:

$$y_{ijmt} = \gamma_{t-m} + \mu_m + \theta_t + \theta_j + f(age_{it}) + \varepsilon_{ijmt}$$

 \triangleright i: individual, j: country of origin, m: year of arrival, t: census year

▶ The main analysis compares occupational mobility of native-born and immigrant workers:

$$y_{ijmt} = \gamma_{t-m} + \mu_m + \theta_t + \theta_j + f(age_{it}) + \varepsilon_{ijmt}$$

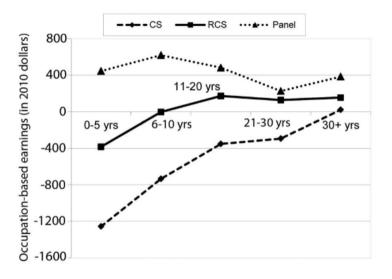
- \triangleright i: individual, j: country of origin, m: year of arrival, t: census year
- $\triangleright \gamma_{t-m}$ measures convergence of immigrants relative to natives

▶ The main analysis compares occupational mobility of native-born and immigrant workers:

$$y_{ijmt} = \gamma_{t-m} + \mu_m + \theta_t + \theta_j + f(age_{it}) + \varepsilon_{ijmt}$$

- ▷ i: individual, j: country of origin, m: year of arrival, t: census year
- $\triangleright \gamma_{t-m}$ measures convergence of immigrants relative to natives

ightharpoonup ABE first present γ_{t-m} from cross-section, repeated cross-section, and panel data



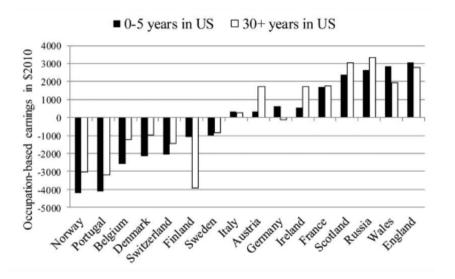
▶ The main analysis compares occupational mobility of native-born and immigrant workers:

$$y_{ijmt} = \gamma_{t-m} + \mu_m + \theta_t + \theta_j + f(age_{it}) + \varepsilon_{ijmt}$$

- \triangleright i: individual, j: country of origin, m: year of arrival, t: census year
- $\triangleright \gamma_{t-m}$ measures convergence of immigrants relative to natives

 \triangleright ABE first present γ_{t-m} from cross-section, repeated cross-section, and panel data

 \triangleright ABE then present results by origin country and 0-5/30+ years in the United States



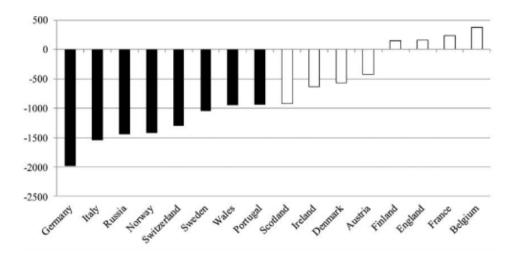
▶ There is no evidence of assimilation with increased duration in the United States

▶ There is some evidence of assimilation for immigrant groups starting out the "farthest behind" native-born workers

> There is no evidence of assimilation with increased duration in the United States

▶ There is some evidence of assimilation for immigrant groups starting out the "farthest behind" native-born workers

▶ Looking at differences in "early" versus "late" skill levels



> There is no evidence of assimilation with increased duration in the United States

▶ There is some evidence of assimilation for immigrant groups starting out the "farthest behind" native-born workers

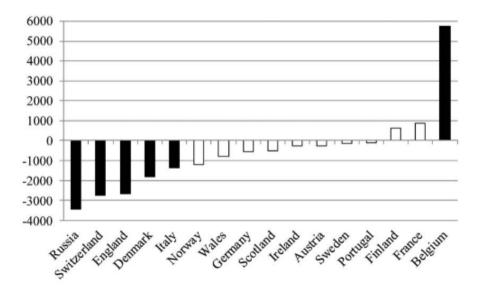
▶ Looking at differences in "early" versus "late" skill levels: countries like Germany, Italy, and Russia saw the largest decrease in skills upon arrival

> There is no evidence of assimilation with increased duration in the United States

▶ There is some evidence of assimilation for immigrant groups starting out the "farthest behind" native-born workers

▶ Looking at differences in "early" versus "late" skill levels: countries like Germany, Italy, and Russia saw the largest decrease in skills upon arrival

▶ Looking at selection of return migrants



> There is no evidence of assimilation with increased duration in the United States

▶ There is some evidence of assimilation for immigrant groups starting out the "farthest behind" native-born workers

▶ Looking at differences in "early" versus "late" skill levels: countries like Germany, Italy, and Russia saw the largest decrease in skills upon arrival

▶ Looking at selection of return migrants: countries like Russia, Switzerland, and England saw negative selection of return migrants

▶ Immigrants to the United States during the Age of Mass Migration did not have lower skill occupations upon arrival, moved up the occupational ladder at same rate as natives

- ▶ Immigrants to the United States during the Age of Mass Migration did not have lower skill occupations upon arrival, moved up the occupational ladder at same rate as natives
- ▶ The sending country is particularly important for predicting success of immigrants in the United States

- ▶ Immigrants to the United States during the Age of Mass Migration did not have lower skill occupations upon arrival, moved up the occupational ladder at same rate as natives
- ▶ The sending country is particularly important for predicting success of immigrants in the United States
- ➤ This suggests considerable scope for quickly integrating immigrants into the American society (contrary to concerns at the time)

- ▶ Immigrants to the United States during the Age of Mass Migration did not have lower skill occupations upon arrival, moved up the occupational ladder at same rate as natives
- ▶ The sending country is particularly important for predicting success of immigrants in the United States
- ▶ This suggests considerable scope for quickly integrating immigrants into the American society (contrary to concerns at the time)
- ▶ But...also suggests scope for labor market competition between immigrants and natives
- perhaps consistent with evidence from Goldin (1993) on the timing, geography, and political economy of immigration restriction in the early twentieth century

▶ Given progress in census linking since ABE, both in terms of our understanding of matching methodologies and availability of complete count census, this work is worth revisiting

▶ Given progress in census linking since ABE, both in terms of our understanding of matching methodologies and availability of complete count census, this work is worth revisiting

- Collins & Zimran draw on newer data to cover both the "early" and "later" periods of the Age of Mass Migration, and come to two main conclusions:
- European immigrants substantially upgraded their occupational status relative to natives in the nineteenth century, but not in the early twentieth century

▶ Given progress in census linking since ABE, both in terms of our understanding of matching methodologies and availability of complete count census, this work is worth revisiting

- Collins & Zimran draw on newer data to cover both the "early" and "later" periods of the Age of Mass Migration, and come to two main conclusions:
- European immigrants substantially upgraded their occupational status relative to natives in the nineteenth century, but not in the early twentieth century
- This difference was due to collapsing differences in occupational distributions between immigrants and natives, not the changing composition of sending countries

