

The Making of Modern America:
Migratory Flows in the Age of Mass Migration

Oriana Bandiera (LSE)

Imran Rasul (UCL)

Martina Viarengo (LSE/Harvard)

March 2011

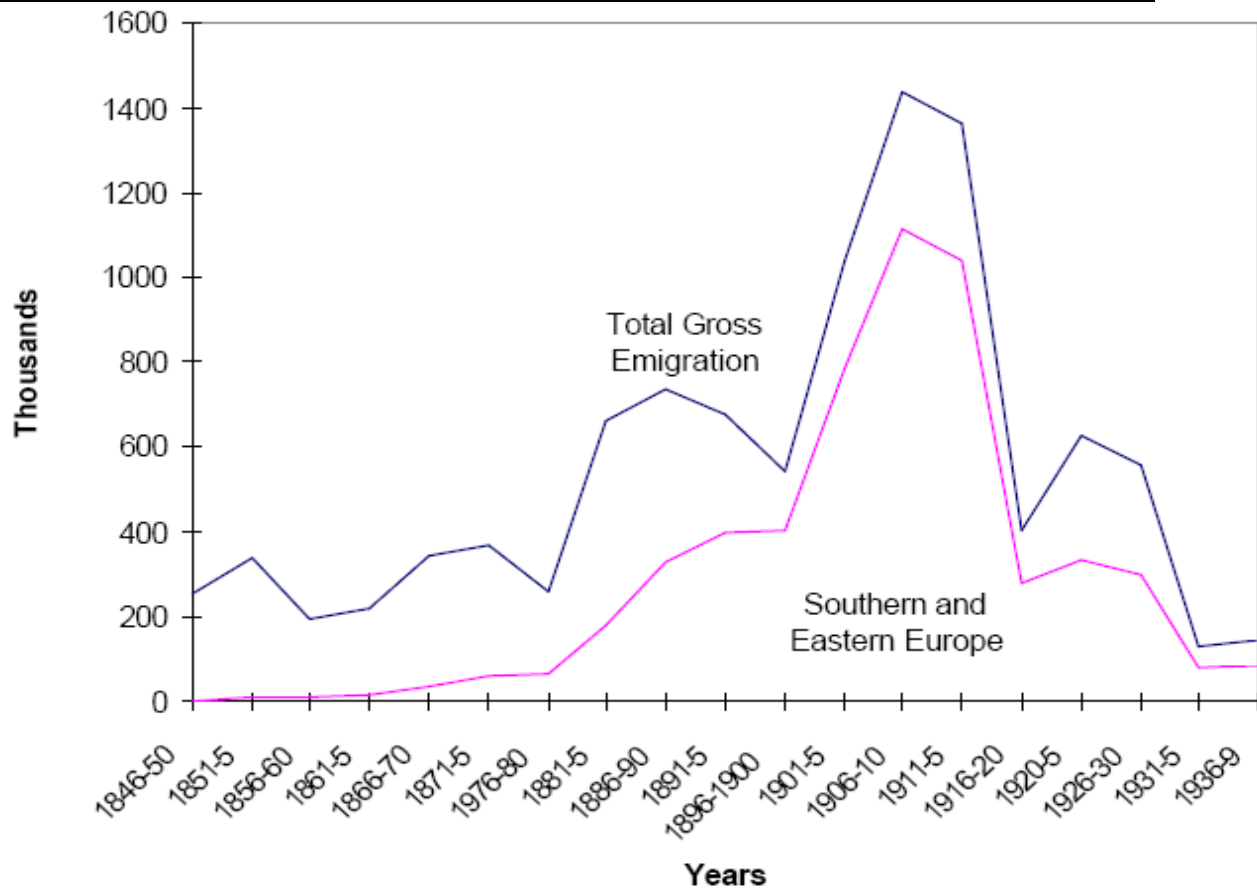
The Melting Pot Society

- America as the ultimate melting pot: number and percentage of the population that is foreign born (2000 Census: 31mn, 12%)
 - 20% of children in primary and secondary school have at least one foreign-born parent [Jamieson et al 2001]
 - 40% of immigrants in OECD are resident in the US [Hansen 2008]
 - diversity among immigrants and historical inflows of immigrants make America almost unique as a host country
 - we study the time period that laid the foundations for modern America: the 'Age of Mass Migration' at the turn of the 20th century
-

Period Under Study

- between 1892 and 1924, over 24mn immigrants were processed through the main port of entry into the US: Ellis Island, NYC
 - flows are sufficiently large to have significant and permanent impacts both on origin and US economies
 - 100mn living descendants of these Ellis Island immigrants, constituting around 40% of the US population today
 - historically, these represent the peak years of immigration into the US
 - recurrent theme: changing composition of immigrants in terms of country of origin, over this period
 - [Figure 1]
-

Figure A1A: Gross Intercontinental Migration From Europe: 1846-1939 (annual averages)



Research Questions

- document mass migration into and out of America at the turn of the twentieth century
 - combining administrative data from the primary port of entry (Ellis Island, New York) and census data from 1880, 1900, 1910, & 1920
 - for each year of arrival 1892-24, we measure immigration into EI, by nationality-gender-age (*nag*) cohorts
 - under a set of assumptions we scale this up to measure immigration into US by *nag* cohort
 - *stock estimates*: for the decades 1900-10, 1910-20, we estimate the total number of immigrant outflows from the US by *nag* cohort
 - *flow estimates*: for each year of arrival and *nag* cohort, we estimate the percentage of immigrants who had left the US by each census date
-

Why Care About Immigrant Outflows?: Immigrant Behavior

- the process of out-migration is central to nearly all issues in the economics of immigration
 - underpins analysis of **immigrant behavior in host** country
 - probability to emigrate and time horizon affects investment into own human capital, children's human capital, saving, and social networks [Galor and Stark 1990, Dustmann 1997, 1999, 2007, Cortes 2004]
 - acquisition of human capital, e.g. language skills, in part determines **substitutability** with native labor
-

Why Care About Immigrant Outflows?: Host Country Effects

- underpins effects of immigrants on **natives in host** country labor markets
 - hotly debated topic with mixed evidence on wage outcomes:
 - negative effect: Borjas et al [1996], Borjas [1999, 2003]
 - little impact: Card [1990, 2001, 2005], Dustmann et al [2005]
 - this debate in contemporary data often hinges in the substitutability of immigrant and native labor [Angrist and Kugler 2003, Ottaviano and Peri 2006, Borjas et al 2008]
 - out-migration might contribute to this substitutability, e.g. acquisition of human capital by immigrants that stay
-

Why Care About Immigrant Outflows?: Origin Country Effects

- understanding brain drain and brain gain to **origin** country [Dustmann and Weiss 2007, Docquier and Rapoport 2008, Mayr and Peri 2008]
 - out-migration might create information networks benefiting origin country [Agrawal et al 2008, Kerr 2008]
 - out-migration fosters development in origin countries [Stark *et al* 1997, Dos Santos and Vinay 2003, Mayr and Peri 2008]
 - general view that historic mass migration had large beneficial effects on origin countries [Hatton and Williamson 2005]
-

Why Care About Immigrant Outflows?: Lack of Contemporary Statistics

- estimated to be over 110mn individuals resident outside their home country today [UN 2005]
 - official statistics related to **out-migration** are of poor quality or non-existent [Thomas 1973, Warren and Peck 1980, Keeling 2006]
 - processing immigrants at US ports declined after implementation of Immigration Act 1924: allowed processing at overseas embassies
 - US abandoned conducting a systematic exit review in 1957
 - sending country records worse than receiving country data [Willcox 1979]
 - UK also does not record out-migration: estimates based on limited surveys [International Passenger Survey, Quarterly National Household Survey]
-

Why Care About Immigrant Outflows?: Lack of Historical Statistics

- no official statistics on out-migration prior to 1907: from late 1800s steamship companies furnished unofficial data about numbers of departing passengers
 - official statistics use these **incomplete** data to estimate emigrant traffic prior to 1907
 - we try to improve on these aggregates, and provide evidence by cohort, using administrative records from EI and census data
 - as an indication of the extent of measurement error in official statistics, our **lower bound** estimate implies the actual number of immigrant arrivals into NYC to be **18% (50%)** higher than official statistics between 1900-10 (1910-20)
-

Comparison to Other Periods of Mass Migration

- between 1900 and 1920 we estimate around 25mn arrivals into US
 - there are no comparably large migratory flows in history: Latinos into US: 8mn in 1980s [Suarez-Orozco and Sommer 2000]; A8 into UK: 2004-6 .6mn [Okolski 2006]
 - flows are sufficiently large to have significant and permanent impacts both on origin and US economies
 - historic episodes of forced migration or relaxation of severe constraints:
 - India-Pakistan-Bangladesh partition: 18million outflow over a four year period [Bharadwaj *et al* 2008]
 - Russian migrants into Israel: 12% population increase between 1990-4 [Friedberg 2001]
 - East to West Germany: .6mn in 1989-90 [Heiland 2004]
-

Understanding Return Migration

- hard to explain in simple income-maximizing models in the presence of large wage differentials [Sjastaad 1962, Harris and Todaro 1970, Gibson and McKenzie 2009]
 - indeed, there is a vast literature estimating substantial gains from migration [Hansen 2008]
 - Abramitzky *et al* [2009]: among Norwegian migrants between 1865 and 1900, compare within brother pairs of stayer and leaver to find mean rates of return to migration of 93% (42%) for those coming from rural (urban) areas in Norway
-

Models of Return Migration

- planned as part of an optimal life cycle residential location sequence [Borjas and Bratsberg 1996, Stark *et al* 1997, Dustmann and Weiss 2007]
 - time varying complementarities between consumption and location [Hill 1987, Djajic and Milbourne 1988, Keeling 2006]
 - target income earner models [Piore 1979, Mesnard 2004, Yang 2006]
 - erroneous beliefs on/shocks to outcomes in US [Pessino 1991, Borjas and Bratsberg 1996]
 - this paper: **measuring** migratory flows
 - future work: develop a model along the lines of Borjas and Bratsberg [1996] to **explain** migratory flows
-

Data Source 1: Administrative Records from Ellis Island

- American Family Immigration History Centre holds database of 24 million individuals whose names appear on the original ship passenger manifests for the Port of New York between 1892 and 1924
 - data entered by 12,000 volunteers of the Church of Jesus Christ of Latter-day Saints [5.6 million man-hours]
 - the Port of New York accounted for around 71% of all immigrant arrivals to the United States between 1892 and 1924
 - we describe in detail the various assumptions needed to **scale-up** this number to estimate total inflows into the US
 - Note: have checked for duplicate records (less than .1%)
-

Ship Manifests

- data originally recorded in ship passenger manifests: Manifest of Immigrants Act, March 2nd 1819 required that from Jan 1st 1820, the master of every vessel entering a US port list each passenger “taken on at any foreign port” by name, gender, age, occupation, calling, and country of origin
 - by 1913 passenger lists expanded to encompass beliefs about politics, marriage, health information, literacy, destination [Smith 1996]
 - manifests collected by the Immigration Service from 1891 generally included:
(i) onboard aliens; (ii) foreign born US citizens [Barde and Bobonis 2006]
 - hence passenger lists are **complete** covering all ships and passengers
 - [Figure 1]
-

Figure 1: Passenger Ship Manifest from March 3rd, 1903

IMMIGRATION SERVICE, Form 1560 B.

SALOON, CABIN, AND STEERAGE ALIENS MUST BE COMPLETELY MANIFESTED. THIS SHEET IS FOR STEERAGE PASSENGERS.

List 60/78

LIST OR MANIFEST OF ALIEN PASSENGERS FOR THE U. S. IMMIGRATION OFFICER AT PORT OF ARRIVAL.

Required by the regulations of the Secretary of the Treasury of the United States, under Act of Congress approved March 3, 1903, to be delivered to the U. S. Immigration Officer by the Commanding Officer of any vessel having such passengers on board upon arrival at a port in the United States.

103 S. *Roma* sailing from *Trapani, Italy, Starboard*, 1905 Arriving at Port of *New York, N.Y.*, 1905.

No. on List.	NAME IN FULL.	Age.		Married or Single.	Calling or Occupation.	Able to Read, Write.	Nationality, (Country of last permanent residence.)	Race or People.	Last Residence, (Province, City, or Town.)	Final Destination, (State, City, or Town.)	Whether having a ticket to such final destination.	By whom was passage paid?	Whether in possession of \$50, and if not, how much?	Whether ever before in the United States; and if so, where and when?	Whether going to join a relative or friend, and if so, what relative or friend, and his name and complete address.	Whether in possession of a Poly-an Ancestral.	Whether ever before in the United States; and if so, where and when?	Condition of Health, Mental and Physical.	Deformed or Crippled, Nature, length of time, and cause.
		Yrs.	Mos.																
1	<i>Tasarelli Antonio</i>	<i>22</i>	<i>6</i>	<i>Single</i>	<i>Hand laborer</i>		<i>Italy</i>	<i>South Italian</i>	<i>Trapani</i>	<i>New York</i>		<i>By himself</i>	<i>100</i>		<i>Brother of [unclear] [unclear] Pa.</i>			<i>Good</i>	
2	<i>Amansharis</i>	<i>5</i>	<i>7</i>	<i>Single</i>	<i>Hand laborer</i>		<i>Italy</i>	<i>South Italian</i>	<i>Trapani</i>	<i>New York</i>		<i>By himself</i>	<i>100</i>		<i>Sons Ruth Gyranza</i>			<i>Good</i>	
3	<i>Castro Elisio</i>	<i>21</i>	<i>6</i>	<i>Single</i>	<i>Hand laborer</i>		<i>Italy</i>	<i>South Italian</i>	<i>Trapani</i>	<i>New York</i>		<i>By himself</i>	<i>100</i>		<i>Brother of [unclear] [unclear] Pa.</i>			<i>Good</i>	
4	<i>Orzoff Eusebio</i>	<i>25</i>	<i>6</i>	<i>Single</i>	<i>Hand laborer</i>		<i>Italy</i>	<i>South Italian</i>	<i>Trapani</i>	<i>New York</i>		<i>By himself</i>	<i>100</i>		<i>Brother of [unclear] [unclear] Pa.</i>			<i>Good</i>	

Notes: The passenger ship manifest shown was accessed from <http://www.ellisland.org/search> on April 24th 2010. Fields indicated in solid (dashed) boxes are available (are not available) in the electronic format of the administrative records.

Data Source 2: US Census 1880, 1900-10-20

- IPUMS census samples: 1880 (100%), 1900 (5%), 1910 (1.4%), 1920 (1%)
[Ruggles et al 2009]
 - in these years, around 14% of the sample are foreign born: nationality of birth and year of arrival recorded
 - nationality of birth recorded even if individual has obtained US citizenship by census date (31%)
 - only keep the foreign born population, so drop children of immigrants born in the US: 20% of the 6.4mn sample are US born but have at least one foreign born parent
 - over census years 1900-20, this gives 916,773 foreign born individuals
-

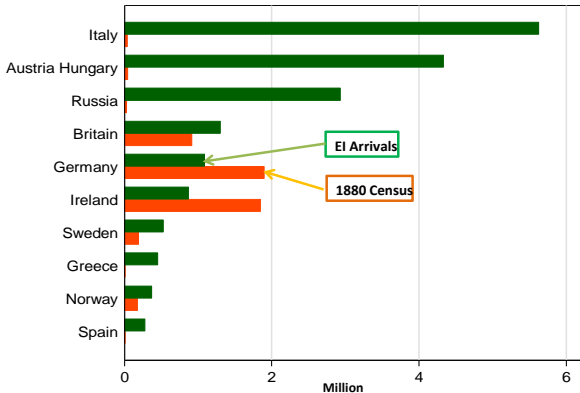
Descriptives

- aggregate arrivals into Ellis Island, 1892-1924
 - age distribution by gender
 - time series: inflows and migrant ages
 - [Figure 2]
-

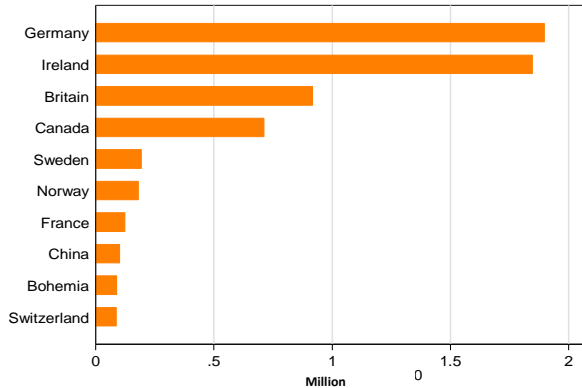
Figure 2: Descriptive Evidence from Administrative Records

A. Total Immigrants, by Nation of Birth

(i) Ellis Island Arrivals 1892-1924, and Population in 1880 US Census

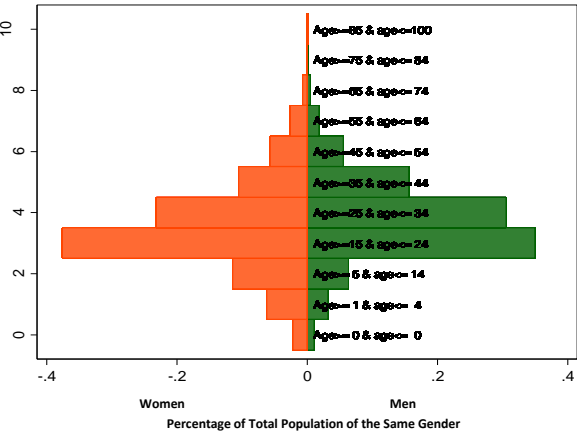


(ii) Foreign Born Population in 1880 US Census

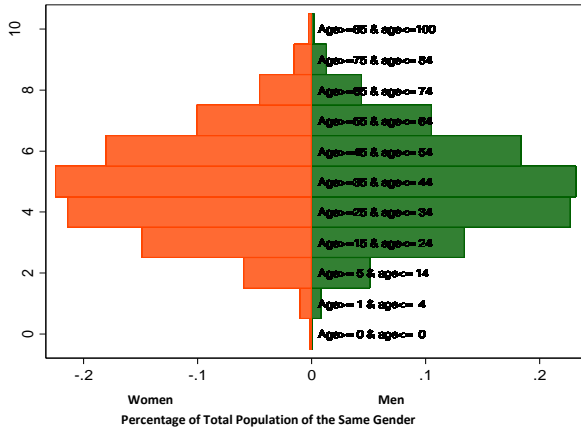


B. Age Distribution, by Gender

(i) Ellis Island Arrivals 1892-1924

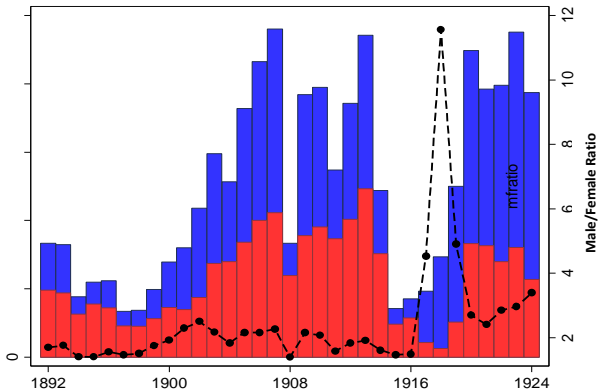


(ii) Foreign Born Population in 1880 US Census

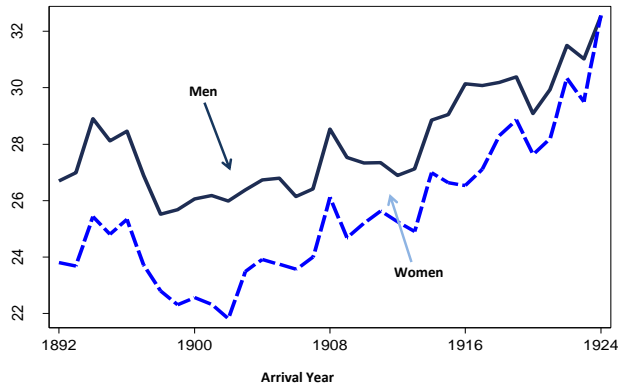


C. Time Series

(i) Arrivals by Gender and Year



(ii) Average Age by Gender and Year



Cohorts in Administrative Records from Ellis Island

- information on the passenger's age, gender, nationality, date of arrival, whether new arrival or returnee
 - also data (unexploited) on marital status, ethnicity, place of residence, name of vessel, port of departure, other ports at which stopped
 - aggregated to measure inflow in each year of each cohorts defined by nationality-gender-age (*nga*)
 - potential cohorts: 118 nationalities \times 11 age groups (0, 1-4, 5-14, 25-34, ..., 85+) \times 2 genders = 2596 cohorts
 - of these, 1525 cohorts have at least one immigrant arrival between 1900 and 1920
-

Accounting Methodology

- established methodology to calculate out-migration estimates: combine census data with INS data on legal immigration [Warren and Peck 1980] and focusing on **one** entry cohort [Jasso and Rosenzweig 1982]
 - studies based on more recent data have to distinguish between temporary and permanent migrants on entry (e.g. students, extended business travellers)
 - temporary visitors have no option but to return and so should not be included in out-migration calculations
 - no such issues arise for the time period we study when immigrants had the possibility to permanently reside in the US
-

Accounting For Out-migration

- foreign-born population in two consecutive years is given by,

$$P_{t+1}^{ng,a+1} = (1 - D_t^a)P_t^{nga} + I_{t+1}^{ng,a+1} - E_{t+1}^{ng,a+1},$$

- P_t^{nga} : stock of individuals in cohort nga in year t
 - D_t^a : mortality rate of individuals of age a in year t
 - $I_{t+1}^{ng,a+1}$: immigrant inflow of cohort nga between years t and $t + 1$
 - $E_{t+1}^{ng,a+1}$: emigrant outflow of cohort nga between years t and $t + 1$
-

Aggregate Out-migration Flows

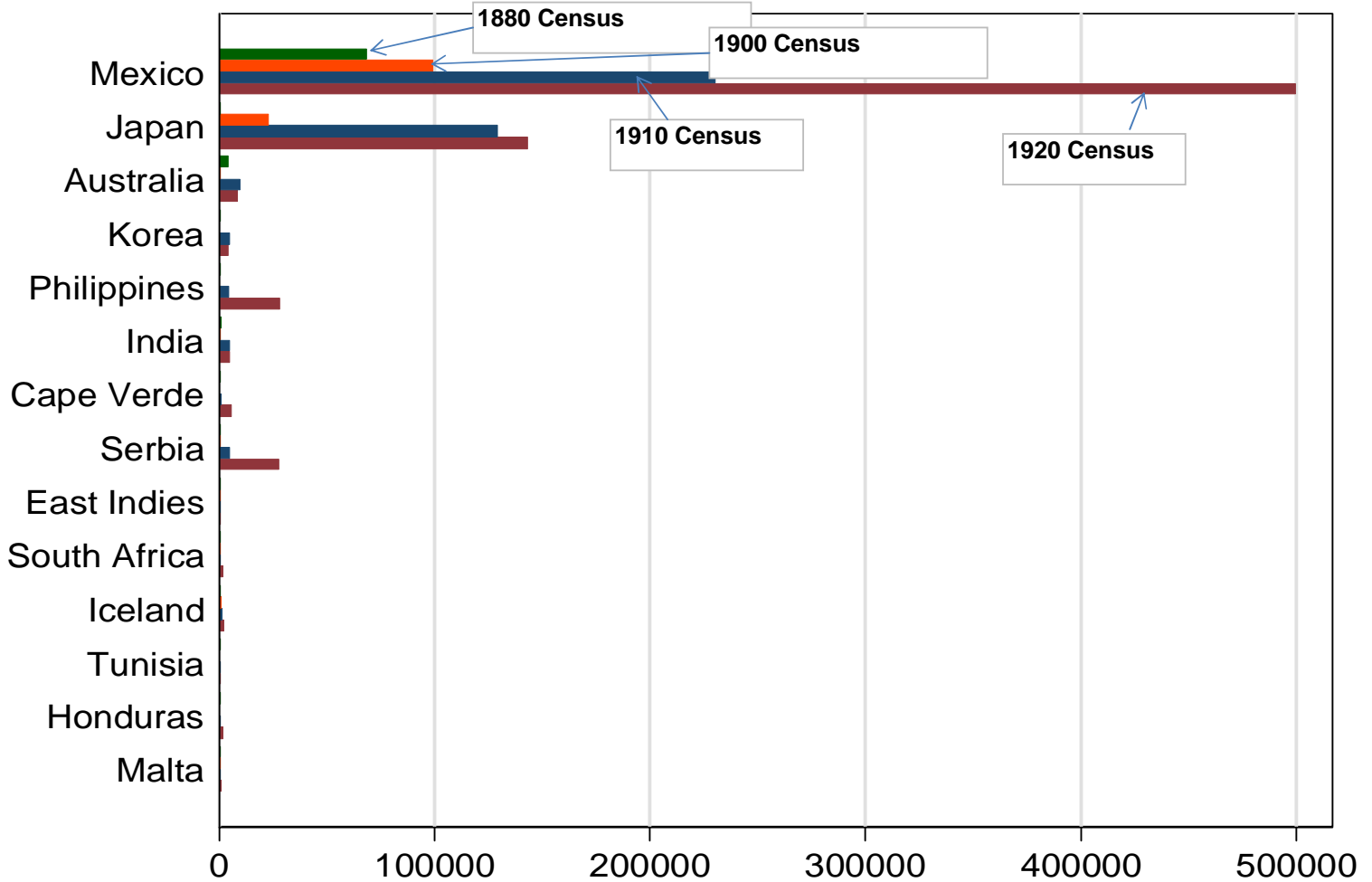
$$\sum_{k=1}^{k=9} \prod_{j=k}^{j=9} (1 - D_{t+j}^{a+j}) (E_{t+k}^{ng,a+k}) = \prod_{j=0}^{j=9} (1 - D_{t+j}^{a+j}) P_t^{nga} - P_{t+10}^{ng,a+10} + I_{t+10}^{ng,a+10} + \sum_{k=1}^{k=9} \prod_{j=k}^{j=9} (1 - D_{t+j}^{a+j}) (I_{t+k}^{ng,a+k})$$

- LHS is number of emigrants in cohort nga in census year $t + 10$ that:
 - emigrated since previous census in year t
 - are still alive somewhere
 - we cannot identify whether they have returned to their home country or moved to some third country
- measured over two census decades: 1900-10, 1910-20

Adjustments

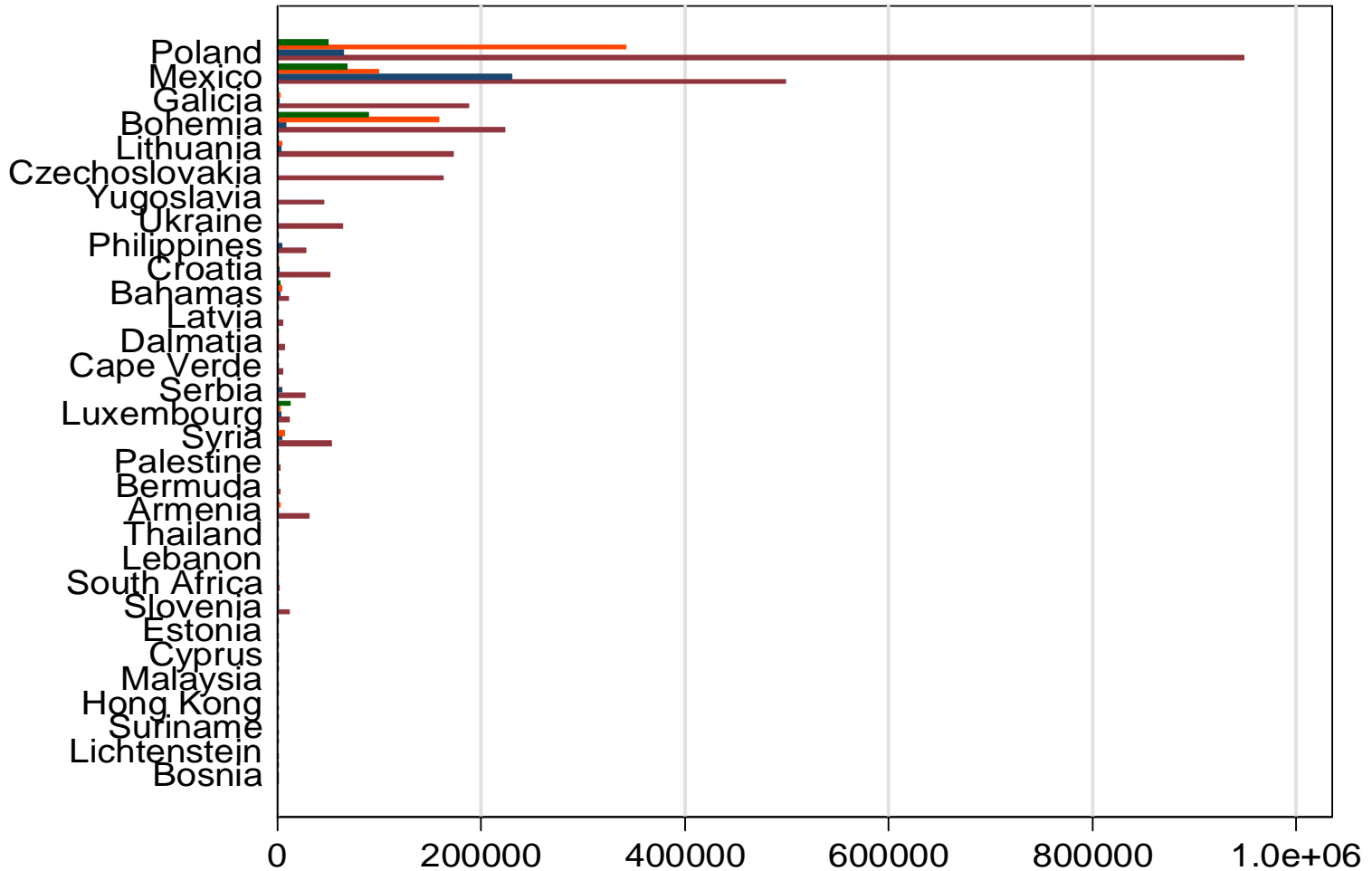
- $I_{t+1}^{ng,a+1}$ refers to total (new) immigrant inflows **into US**
 - to scale up to this number from the EI database we need to make some adjustments:
 - missing *nga* data from from EI database
 - expulsion or death at Ellis Island
 - other ports of entry into US
 - inflows from Canada and Mexico and illegals
 - US citizenship
 - most adjustments go in one direction so that we typically **underestimate** immigration, hence out-migration
-

Figure A1a: Census Populations in 1880, 1900, 1910 and 1920, for Countries With Implied Negative Emigrant Numbers 1900-1910



Notes: The sample of countries for the Figure in each decade is those countries for whom the total estimated number of emigrant departures is found to be negative after making corrections for missing values, excluded immigrants, weighting for the ratio of official arrivals into the US to those into New York each year, inflows from Canada, and foreign born individuals with US citizenship. For these countries, the figure then shows the population in the US based on Census data from 1880, 1900, 1910 and 1920.

Figure A1b: Census Populations in 1880, 1900, 1910 and 1920, for Countries With Implied Negative Emigrant Numbers 1910-1920



Notes: The sample of countries for the figure in each decade is those countries for whom the total estimated number of emigrant departures is found to be negative after making corrections for missing values, excluded immigrants, weighting for the ratio of official arrivals into the US to those into New York each year, inflows from Canada, and foreign born individuals with US citizenship. For these countries, the figure then shows the population in the US based on Census data from 1880, 1900, 1910 and 1920.

Mortality Rates

- out-migration formula assumes death rates are the same in the source and destination country
 - for US from 1900-39, have race-gender-age-year specific death rates: races are white and other, age groups are 0, 1-4, 5-14, 15-24, ..., 75-84, 85+
 - use these age ranges to define cohort dimension a
 - use white mortality rates [Warren and Peck 1980, Jasso and Rosenzweig 1982], black mortality rates, and nationality specific mortality rates [Tizzano 1965, Nugent 2002, Mitchell 2007]
-

Official (INS) Estimates of Immigrant Arrivals

- for officially reported statistics, scholars rely on the work of Imre Ferenczi, which was compiled under the auspices of the International Labor Office in the early 1920s and published in 1929 by the NBER (Willcox 1929)
 - Ferenczi-Willcox (1929) brings together all official data series on post-1820 international migration then available in published form or supplied by governments
 - records alien arrivals to the United States from 1820 to 1924 by nationality
 - no account taken of mortality, so in order to replicate we assume $D_t^a = 0$
-

Official Estimates of Immigrant Arrivals into US

- historians have documented various reasons for mis-reporting immigrant arrivals into US [Hutchinson 1958, Keeling 2006]:
 - careless collection of ship manifests by port officials; failure to forward Passenger Abstracts quarterly to the Department of State; failure of State Department clerks to include all Passenger Abstracts in their annual statistical reports [Hutchinson 1958]
 - immigrant aliens traveling in first and cabin class were not consistently counted as immigrants [25% underestimate?]
-

Official Estimates of Immigrant Arrivals into NYC

- focus first on arrivals into NYC so that direct comparisons can be made
 - administrative records suggest total arrivals are 3-7% higher than officially recorded between 1900-10, and 30% higher in 1910-20
 - US citizens appear less well officially recorded in 1900-10, accurately done in 1910-20
 - address concern that administrative records also count temporary migrants is addressed by examining non-immigrant arrivals
 - [Tables 1-3]
-

Table 1: Official Statistics and Administrative Record Measures of Migrant Inflows into **New York, by Decade**

	(1) Official Statistic Ferenzi-Willcox [1929]	(2) Preferred Estimate	(3) Lower Bound	(4) Ratio: Preferred Estimate to Official Statistic (Col 2/Col 1)	(5) Ratio: Lower Bound Estimate to Official Statistic (Col 3/Col 1)
<u>1900-1910</u>					
Total immigrant foreign-born arrivals	7431670	8968628	8792771	1.21	1.18
Total non-immigrant foreign-born arrivals	713749	-	-	-	-
Total foreign-born arrivals	8145419	8968628	8792771	1.10	1.08
Total US citizen arrivals	1546237	1200336	1150045	0.78	0.74
Total arrivals	9691656	10168964	9942816	1.05	1.03
<u>1910-1920</u>					
Total immigrant foreign-born arrivals	4416448	7054163	6624076	1.60	1.50
Total non-immigrant foreign-born arrivals	856931	-	-	-	-
Total foreign-born arrivals	5273379	7054163	6624076	1.34	1.26
Total US citizen arrivals	1214658	1249759	1147865	1.03	0.95
Total arrivals	6488037	8303922	7771941	1.28	1.20

Table 2: Official Statistics and Administrative Record Measures of Migrant Inflows for the US, by Decade

	(1) Official Statistic Ferenzi-Willcox [1929]	(2) Preferred Estimate	(3) Lower Bound	(4) Ratio: Preferred Estimate to Official Statistic (Col 2/Col 1)	(5) Ratio: Lower Bound Estimate to Official Statistic (Col 3/Col 1)
<u>1900-1910</u>					
Total immigrant foreign-born arrivals	9719358	13712006	8792771	1.41	0.90
Total non-immigrant foreign-born arrivals	994168				
Total foreign-born arrivals	10713526	13712006	8792771	1.28	0.82
Total US citizen arrivals	2040674	1570248	1150045	0.77	0.56
Total arrivals	12754200	15282254	9942816	1.20	0.78
<u>1910-1920</u>					
Total immigrant foreign-born arrivals	6659210	18511266	6624076	2.78	0.99
Total non-immigrant foreign-born arrivals	1540972				
Total foreign-born arrivals	8200182	18511266	6624076	2.26	0.81
Total US citizen arrivals	2111460	2426712	1147865	1.15	0.54
Total arrivals	10311642	20937978	7771941	2.03	0.75

Notes: The official statistics in Column 1 of Tables 1 and 2 are from Ferenzi-Willcox [1929]. For all other statistics derived from Ellis Island Administrative records, these are based on the total number of immigrant arrivals (new and returnee). For statistics related to arrivals into New York City in Table 1, the preferred estimate figure in Column 2 is based on corrections for missing data, other potential errors in recorded nationalities, and exclusions. The lower bound estimate in Column 3 is based on the raw administrative statistics from which no adjustments are made. For statistics related to arrivals and departures into the US in Table 2, the preferred estimate in Column 2 is based on corrections for missing data and other potentially mis-coded nationalities, expulsion or death, inflows from other sea ports, and inflows over land via Canada and Mexico. The lower bound estimate in Column 3 is based on the raw administrative statistics from which no adjustments are made. We assume the census takes place on July 1st each census year and so use mid-year inflows of immigrant numbers, for census years 1900, 1910 and 1920. We make the corresponding adjustment to official statistics to compare these series with our estimates.

Table 3A: Official Statistics and Administrative Record Measures of Migrant Flows for the US, by Decade

	(1) Official Statistic Ferenzi-Willcox [1929]	(2) Preferred Estimate	(3) Lower Bound	(4) Ratio: Preferred Estimate to Official Statistic (Col 2/Col 1)	(5) Ratio: Lower Bound Estimate to Official Statistic (Col 3/Col 1)
<u>1900-1910</u>					
Total immigrant foreign-born arrivals	9719358	13712006	8792771	1.41	0.90
Total migrant departures from US	3377618	10429231	7191956	3.09	2.13
Implied out-migration rate for US	0.348	0.761	0.818	2.19	2.35
<u>1910-1920</u>					
Total immigrant foreign-born arrivals	6659210	18511266	6624076	2.78	0.99
Total migrant departures from US	2372071	18048715	8828942	7.61	3.72
Implied out-migration rate for US	0.356	0.975	1.333	2.74	3.74

Checking With Census Data

- our estimated rates broadly correspond to aggregate numbers of foreign born citizens reported in US census data
 - in 1900 13.6 of the US population was foreign born – rose to 14.7% in 1910, fell to 13.2% in 1920
 - absolute increase between 1900-10 from 10.3mn to 13.5mn (31%)
 - absolute increase between 1910-20 from 13.5mn to 13.9mn (3%) – despite inflow of millions [Table 1]
-

Other Historic Estimates of Out-migration Rates

- Ferenczi-Willcox (1929) based out-migration rate of around .3: widely cited figure in existing literatures in economics, demography and economic history [Hatton and Williamson 2005]
 - variation across countries: up to half among Spanish, Italians and Argentines; lower for Russians, Irish and Scandinavians
 - Kuznets and Rubin [1954] also try to correct official statistics and find rates around .5
 - [Table 3B]
-

Table 3B: Kuznets-Rubin [1954] and Administrative Record Measures of Migrant Flows for the US, by Decade**(1) Kuznets-Rubin
[1954]****(2) Preferred
Estimate****(3) Lower
Bound****(4) Ratio: Preferred
Estimate to Official
Statistic (Col 2/Col 1)****(5) Ratio: Lower Bound
Estimate to Official
Statistic (Col 3/Col 1)****1900-1910****Total immigrant foreign-born arrivals**

9447500

13712006

8792771

1.45

0.93

Total migrant departures from US

4230000

10429231

7191956

2.47

1.70

Implied out-migration rate for US

0.448

0.761

0.818

1.70

1.83

1910-1920**Total immigrant foreign-born arrivals**

7400000

18511266

6624076

2.50

0.90

Total migrant departures from US

3963000

18048715

8828942

4.55

2.23

Implied out-migration rate for US

0.536

0.975

1.333

1.82

2.49

Notes: The official statistics in Column 1 of Table 3A are from Ferenczi-Willcox [1929]. In Table 3B the comparison is made to the Kuznets-Rubin [1954] correction of these official statistics. For all other statistics derived from Ellis Island Administrative records, these are based on the total number of immigrant arrivals (new and returnee). The preferred estimate in Column 2 is based on corrections for missing data and other potentially mis-coded nationalities, expulsion or death, inflows from other sea ports, and inflows over land via Canada and Mexico. The lower bound estimate in Column 3 is based on the raw administrative statistics from which no adjustments are made. The implied out-migration rate is the total number of migrant departures divided by the total number of immigrant arrivals into the US in the same decade. We assume the census takes place on July 1st each census year and so use mid-year inflows of immigrant numbers, for census years 1900, 1910 and 1920. We make the corresponding adjustment to official statistics to compare these series with our estimates.

Out-migration Rates of Specific Cohorts

- breaking down by various groups see out-migration rates of between 50-85% between 1900-10, and between 70-100% for 1910-20
 - within 1900-10 decade, do find higher out-migration rates for Italy and Spain, lower for Russia
 - Germany and Austria-Hungary have net outflows during the decade covering WW1
 - Canadian rates more in line with historic and contemporary out-migration rates
 - [Table 4]
-

Table 4: Out-migration Rate Estimates by Cohort and Decade

Cohort	1900-1910			1910-1920		
	(1) Total immigrant arrivals	(2) Total immigrant departures	(3) Implied out-migration rate	(4) Total immigrant arrivals	(5) Total immigrant departures	(6) Implied out-migration rate
Aggregate based on total immigrant arrivals	13712006	10429231	0.761	18511266	18048715	0.975
Aggregate based on first time immigrant arrivals	13317559	10034791	0.754	13863483	13400968	0.967
White mortality rate	14154747	8951276	0.632	14582711	11836158	0.812
Other mortality rate	14107050	8239115	0.584	14621274	10900747	0.746
Nationality specific mortality rate	14211636	8750683	0.616	14631129	11595469	0.793

Nationality	1900-1910			1910-1920		
	(1) Total immigrant arrivals	(2) Total immigrant departures	(3) Implied out-migration rate	(4) Total immigrant arrivals	(5) Total immigrant departures	(6) Implied out-migration rate
Top Ten Nationalities Based on Immigrant Arrivals into NYC Between 1892-1924	11996584	7181314	0.599	11002964	10765310	0.978
Rank 1: Italy	3372036	2438093	0.723	2721625	2281362	0.838
Rank 2: Austria-Hungary	2869037	1584087	0.552	878582	1338729	1.524
Rank 3: Russia	2024757	825060	0.407	1116179	1085628	0.973
Rank 4: Great Britain	964993	645387	0.669	2914890	2772513	0.951
Rank 5: Germany	1164191	608361	0.523	646595	787111	1.217
Rank 6: Ireland	644574	477324	0.741	556334	522350	0.939
Rank 7: Sweden	397799	236088	0.593	442348	398894	0.902
Rank 8: Greece	352056	263177	0.748	462087	393909	0.852
Rank 9: Norway	277015	159524	0.576	479516	453094	0.945
Rank 10: Spain	144674	121729	0.841	628530	562680	0.895
Other: Canada	307064	120745	0.393	363390	157123	0.432

Notes: All statistics derived from Ellis Island Administrative records are based on the number of new immigrant arrivals, except in the first row that is based on the total number of immigrant arrivals (new and returnee). These preferred estimates are based on corrections for missing data and other potentially mis-coded nationalities, expulsion or death, inflows from other sea ports, and inflows over land via Canada and Mexico. In the first two rows, a survival rate of one is assumed. In the third and fourth rows, survival rates of whites and "other" race are used. In the fifth row, nationality specific mortality rates are used for Italy, Austria-Hungary, Great Britain, Russia, Germany, Ireland, Sweden, Spain, France, Norway, Denmark, Finland, Belgium, Romania and Switzerland. For all other nationalities, white mortality rates are assumed. For gender and age specific cohorts, white mortality rates are assumed. In the lower panel, for the country specific cohorts, country specific mortality rates are used. The ten countries chosen (plus Canada) are those from which the most immigrant arrivals originate from into Ellis Island over the period 1892-1924.

More Recent Estimates of Out-migration Rates from US

- official US Bureau of the Census estimate of out-migration between 1981-90 is 22% (cannot distinguish between those with permanent or temporary visas)
 - using Census and INS data: Warren and Peck [1980] estimate out-migration rate between 1960-70 of 18%, higher for women, and higher for older age groups; using similar data Borjas and Bratsburg [1996] estimate 17.5% out-migration rate within 10 years, much of it within the first 5 years
 - Jasso and Rosenzweig [2002]: combine INS administrative records at entry for 1971 cohort of legal permanent immigrants with their subsequent naturalization to estimate 10 year out-migration rates of between 30 and 50%
 - Mayr and Peri [2008]: use differences in census years (1980-2000) and simulations suggest 25% return after 20 years, increasing with education, varies by nationality
-

Summary

- demographic accounting exercise **overturns conventional wisdom** on aggregate out-migration rates out of US
 - out-migration rates two to three times higher than official estimates suggest: almost complete turnover of the population between 1910-20
 - potentially important selection into remaining in US
 - the extent of out-migration is likely to have qualitatively large impacts on immigrant behavior, natives in the US, and the economies of sending countries
-

Conclusion: Why Care About Out-migration?

- lack of accurate official statistics on out-migration
 - underpins analysis of **immigrant behavior in host** country
 - underpins effects of immigrants on **natives in host** country labor markets
 - important mechanism for **sending** countries to benefit
-

Why is Historic Information Useful?

- pre-WW1, migration took place without visas, quotas, asylum status, green cards, illegals or security barriers
 - from 1917 onwards, experience legislative changes related to those eligible to enter US, financial costs of entry, and quotas
 - a view that a return to an era of free movement of labor, as with trade, would raise world welfare [Rodrik 2002, Rosenzweig 2005, Pritchett 2006, Kremer and Watt 2006]
 - a 3% increase in labor migration would result in half the gains associated with complete trade liberalization [Walmsley and Winters 2003]
 - removal of all barriers to migration between OECD and non-OECD countries would boost world output between 92 and 172% [Klein and Ventura 2004]
-

Estimating Return Migration Flows

- denote I_{kyt} the immigrant arrivals of type k (age, gender, nationality) who arrived in year y and who are estimated to be alive by census date t
- denote P_{kyt} the number of foreign born of type k who arrived in year y observed in census year t
- return migration rate: share of immigrants of cohort k who arrived in year y and left by census date t ,

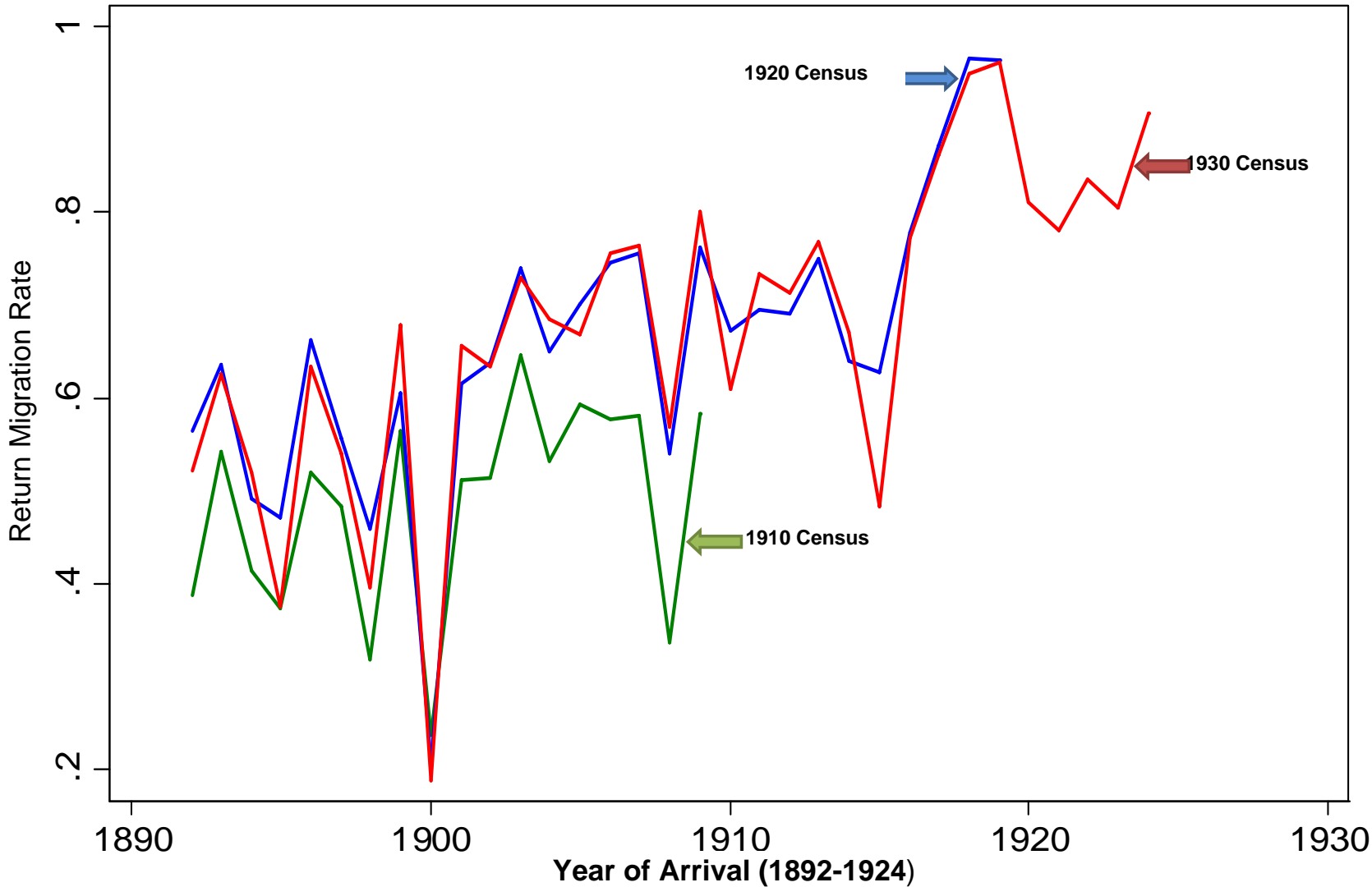
$$r_{kyt} = 1 - \frac{P_{kyt}}{I_{kyt}}$$

- if I_{kyt} is accurately measured, then $0 \leq \frac{P_{kyt}}{I_{kyt}} \leq 1$, thus $0 \leq r_{kyt} \leq 1$
-

Return Migration Flows: Descriptives

- graph return migration rates by year of arrival aggregated across all cohorts, for each of the census years (t)
 - on average over this period, 64% of immigrants do not stay permanently in the US
 - large fluctuations by year of arrival: generally upward trend as entry requirements become stricter; highest rates of return among cohorts that arrived 1917-1919
 - large variations in return flows by nationality: higher for Southern Europeans (Italy, Spain) than Russia
 - dramatic reduction in return migration among those from Austria-Hungary and Germany that arrived during WW1: weaker intentions to return
 - [Figure 6, Table 4: Cols 1-2b]
-

Figure 6: Return Migration Rates, by Arrival Year



Notes: The return migration rate for each year of arrival cohort is computed as one minus the ratio of the number of immigrants of that cohort present in the census to the number of immigrants of that cohort from Ellis Island administrative records who are estimated to be alive at census date. This is done for three census years: 1910, 1920 and 1930. These estimates for immigrant and emigrant numbers are based on corrections for missing values, excluded immigrants, and weighting for the ratio of official arrivals into the US to those into New York each year. White mortality rates are assumed. The sample is restricted to the ten countries that have the highest number of (unadjusted) arrivals into Ellis Island between 1892 and 1899.

Table 4: Return Migration Estimates by Nationality and Year of Arrival

Mean and standard deviation in parentheses

Nationality	(1) All Years of Arrival 1892-1919	(2a) Pre-War Period 1911-1914	(2b) WW1 1915-1918
Austria-Hungary	.628 (.301)	.811 (.090)	.381 (.408)
Great Britain	.528 (.333)	.617 (.235)	.828 (.141)
France	.669 (.337)	.695 (.279)	.828 (.210)
Germany	.428 (.307)	.639 (.165)	.238 (.387)
Ireland	.546 (.338)	.634 (.241)	.706 (.353)
Italy	.646 (.218)	.692 (.153)	.597 (.273)
Norway	.540 (.338)	.588 (.327)	.841 (.192)
Russia	.473 (.275)	.598 (.177)	.402 (.348)
Spain	.787 (.285)	.547 (.350)	.747 (.261)
Sweden	.476 (.328)	.556 (.297)	.757 (.243)

Notes: The return migration rate for each nationality cohort is computed as one minus the ratio of the number of immigrants of that cohort present in the census to the number of immigrants of that cohort from Ellis Island administrative records who are estimated to be alive at census date. The sample is restricted to the ten nationalities that have the largest inflows into Ellis Island between 1892 and 1899. These estimates are based on corrections to immigrant and emigrant numbers for missing values, excluded immigrants, and weighting for the ratio of official arrivals into the US to those into New York each year. White mortality rates are assumed. In Columns 3a to 3d, we report averages over all cohorts of a given nationality. The number of individuals of the same nationality is the number of individuals on the same ship at arrival that are of the same nationality. The other networks are similarly defined.

Table 5: Correlates of Return Migration

Dependent Variable: Return Migration Rates

Robust standard errors in parentheses

	(1) Demographics	(2) Time	(3) Countries
Gender [male=1]	0.0801*** (0.00660)	0.0801*** (0.00646)	0.0801*** (0.00603)
Aged 15-24	0.141*** (0.0104)	0.141*** (0.0101)	0.141*** (0.00949)
Aged 25-34	0.293*** (0.0101)	0.293*** (0.00984)	0.293*** (0.00924)
Aged 35-44	0.323*** (0.0107)	0.323*** (0.0104)	0.323*** (0.00972)
Aged 45+	0.313*** (0.0115)	0.313*** (0.0112)	0.313*** (0.0106)
Arrival year		0.00800*** (0.000450)	0.00801*** (0.000425)
Log(number of years in US)		0.0236*** (0.00512)	0.0236*** (0.00488)
Austria-Hungary			0.199*** (0.0147)
Great Britain			0.100*** (0.0139)
France			0.241*** (0.0142)
Ireland			0.118*** (0.0136)
Italy			0.218*** (0.0120)
Norway			0.111*** (0.0142)
Russia			0.0444** (0.0137)
Spain			0.358*** (0.0141)
Sweden			0.0477*** (0.0141)
Adjusted R-squared	0.166	0.202	0.304
Observations	7998	7998	7998

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The return migration rate for each nationality-age-gender cohort is computed as one minus the ratio of the number of immigrants of that cohort present in the census to the number of immigrants of that cohort from Ellis Island administrative records who are estimated to be alive at census date. These estimates for immigrant and emigrant numbers are based on corrections for missing values, excluded immigrants, and weighting for the ratio of official arrivals into the US to those into New York each year. White mortality rates are assumed, and robust standard errors are reported. The sample is restricted to the ten countries that have the highest number of (unadjusted) arrivals into Ellis Island between 1892 and 1899. The WW1 dummy is equal to one for the period 1915-18.