The linked lives of Japanese-Americans before World War II

Naomi Harada Thyden
Minnesota Department of Children, Youth and Families

Evan Roberts
History of Medicine & Population Studies, University of Minnesota

UNIVERSITY OF MINNESOTA

LIFE COURSE CENTER

Overview

- Project motivation
  - Linking War Relocation Authority records to 1950 census and MLP

- MLP and small immigrant populations
  - MLP methods
  - Considerations for small populations
  - Which immigrant groups are best comparator for Japanese?

- Japanese-American lives in the inter-war era
  - Low MLP linkage rates

UNIVERSITY OF MINNESOTA

LIFE COURSE CENTER

1. Project motivation

Lillian Harada, died at 70
David Harada, died at 66

Joan Thyden, died at 89
Wayne Thyden, alive at 95

LIFE COURSE
CENTER

UNIVERSITY OF MINNESOTA

Asian Americans and the U.S.

We Are Not Silent | BY LUCY FELDMAN

CONFRONTING AMERICA'S LEGACY OF ANTI-ASIAN VIOLENCE BY CADY LANG

- History of exclusion and discrimination
- "Asian hate"
- Perpetual foreigners
  - Model minority myth
  - Calls for Asian American health research
    Abuelazam (2022)
    Rubin, Hires, Sauma & Hau (2022)
    Yi (2020)
    Muramatsu & Chin (2022)
    Sabado-Liwag et al. (2022) Health Affairs
    Yi et al. (2022)
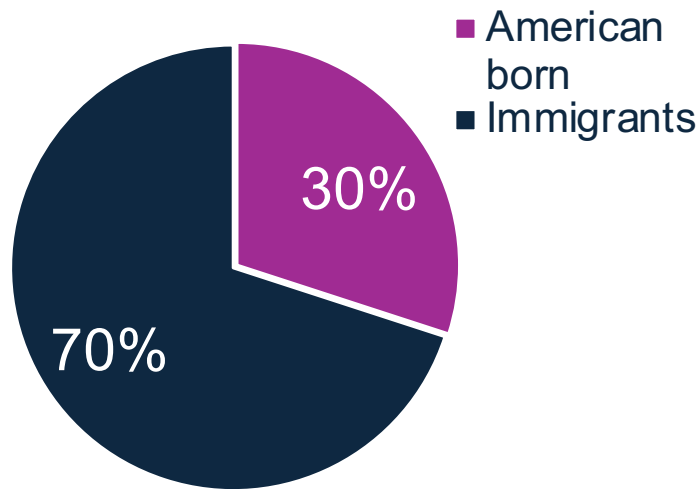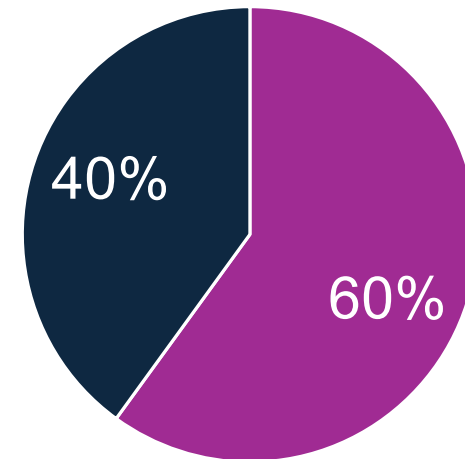
LIFE COURSE CENTER

UNIVERSITY OF MINNESOTA

The healthy immigrant effect and Japanese Americans

All Asian Americans

- ■ American born
- ■ Immigrants

30%

70%

Japanese Americans

40%

60%

UNIVERSITY OF MINNESOTA

LIFE COURSE CENTER

Japanese and Japanese American forced removal and incarceration, 1942-1944

➢ Roosevelt administration's Executive Order 9066

➢ 126,000 people of Japanese descent

　➢ 2/3 US-born,1/3 immigrants forbidden from naturalizing

➢ $400 million of lost property

➢ Post-war assimilation and trauma coping mechanisms*

*Nagata (2019)

UNIVERSITY OF MINNESOTA

Japanese Americans and Health

- Japanese Americans have higher rates of pancreatic cancer, obesity, high blood pressure, diabetes, and asthma than non-Hispanic whites.

- Researchers have not found genetic causes and have therefore attributed disparities broadly to 'lifestyle'.

  However, disparities persist after accounting for lifestyle measures.

- There is a need to understand environmental, sociocultural, behavioral, and biological drivers of health disparities among Japanese Americans.

LIFE COURSE
CENTER

UNIVERSITY OF MINNESOTA

Data linkage to describe Japanese Americans pre-incarceration
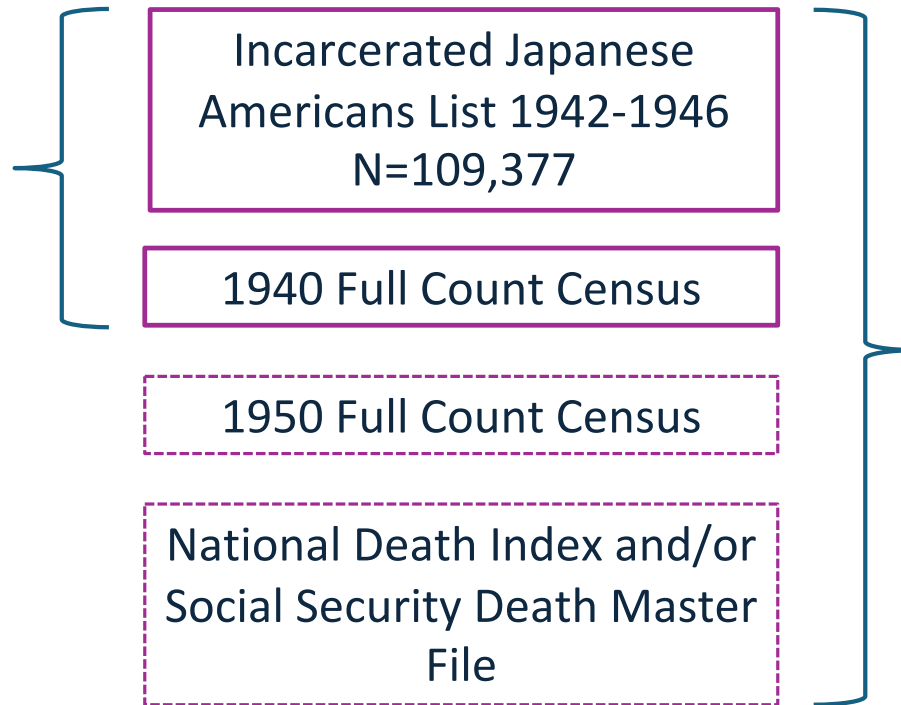
Data linkage to assess
1. Socioeconomic changes from pre- to post-incarceration
2. Mortality effects of incarceration

Incarcerated Japanese Americans List 1942-1946
N=109,377

1940 Full Count Census

1950 Full Count Census

National Death Index and/or Social Security Death Master File

LIFE COURSE CENTER

UNIVERSITY OF MINNESOTA

Bad news ...

- 1950 complete-count census data has missing RACE values for Asians in the contiguous 48
- Worse in California
- FBPL and MBPL are sample line questions in 1950
  - One member per household
- Can't identify third generation Japanese

2. MLP and small immigrant populations

LIFE COURSE CENTER

2.1 MLP Linking Strategy

- Step 1: Individual links
  - Linked based on individual and household characteristics
  - Only men (so far)

- Step 2: Household links
  - Links within households of individuals linked
  - Includes women
  - Lowers thresholds for matches

UNIVERSITY OF MINNESOTA

LIFE COURSE
CENTER

Step 1: Summary

Potential Matches
- Males
- Same place of birth
- Birth year is within +/- 3 years
- Name similarity

Within those matches, predict similarity and confirm match when:

- The predicted similarity exceeds a given threshold (α)

- The predicted similarity of the **most probable** combination (α_1) exceeds the predicted similarity of the **second-most probable** combination (α_2) by at least (β).

Implementing Step 1

- Machine learning algorithm optimized using train-test-split procedure, aiming to maximize Matthew's Correlation Coefficient (MCC).

- We have highly imbalanced class distribution, non-links (negatives) are much more than links (positives).

- MCC is a balanced measure taking all four categories (true positives, false positives, true negative and false negatives) into account.

Step 1 Linking Variables

- **Individual characteristics:** First and last name similarity (JW, Soundex), middle name initial match, birth year difference, second gen immigrant, year of immigration difference, race match, number of potential matches, number of same name individuals in census A

- **Household characteristics:** Father/mother presence, birthplace and JW, spouse presence, birthplace and JW, relatives present, unrelated household members present

- **Contextual characteristics:** Region of birth, common neighbors, geographical distance, street name

LIFE COURSE
CENTER

UNIVERSITY OF MINNESOTA

Rationale for Step 2

- Many individuals are difficult to link due to the failure to appropriately distinguish between individuals in the pool of potential matches.

- Characteristics change or are transcribed incorrectly and appear to change
  - First name changes (John W – William J)
  - (More substantial) inconsistencies in year and place of birth, and gender

Step 2: Summary

Potential Matches

- Household residents with linked individuals (from previous step)
- Birth year with +/- 10 years

Within those matches, predict similarity and confirm match when:

- The predicted similarity exceeds a given threshold (α)
- The predicted similarity of the **most probable** combination (α_1) exceeds the predicted similarity of the **second-most probable** combination (α_2) by at least (β).

Step 1 vs. Step 2

By simultaneously relaxing and restricting the blocking criteria, Step II exploits confident links from Step I to link previously unlinked individuals.

	Step I	Step II
Adjusted Bigram match	Yes	No
Gender	Yes	No
State/country of birth	Yes	No
First/last name JW-score > 0.8	Yes	No
Year of birth	+/-3	+/-10
Census B household members of successfully linked Census A individual residing in same household as focal individual	No	Yes

UNIVERSITY OF MINNESOTA

LIFE COURSE CENTER
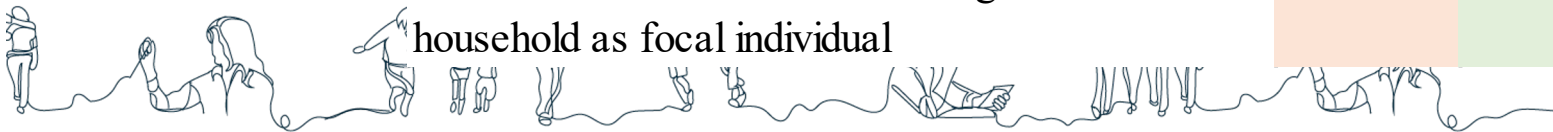
Implementing Step 2

- The changes to the blocking criteria results in a substantial reduction in the mean number of potential matches per Census A individual

- Algorithm calibrated similarly as Step 1, on training data uniquely generated for Step 2

Step 2 Linking Variables

- **Individual characteristics:** First and last name similarity (JW score), first name initial match, middle name initial match, birth year difference, foreign born, birth place match, race match, gender match, marital status and marriage duration match, household position match

Implementing the MLP algorithm

| Census A men selected | → | Linked to universe of potential matches from Census B, following Step 1 blocking criteria | → | Step 1 links identified, using outlined algorithm |

| Unlinked census A household members of confirmed links selected | → | Linked to universe of potential matches from Census B, following Step 2 blocking criteria | → | Step 2 links identified, using outlined algorithm |

2.2 MLP and small immigrant populations

LIFE COURSE CENTER

MLP and small populations

Probability of being linked depends on

Step 1

- Distinctiveness in (name, birth year, birth location, race) space
- Accuracy and stability of name enumeration over time

Step 2

- Family structure —> are people living with consistent kin

Japanese might be better compared to European immigrant groups than Chinese

Immigration to the United States, 1850-1940

Scandinavia includes Denmark, Norway, Sweden, Iceland (**not Finland**)

Y-axis: Thousands of immigrants

X-axis: Year

Legend:
- Japan
- Portugal
- China
- Other NW Europe
- Turkey
- Scandinavia (all)
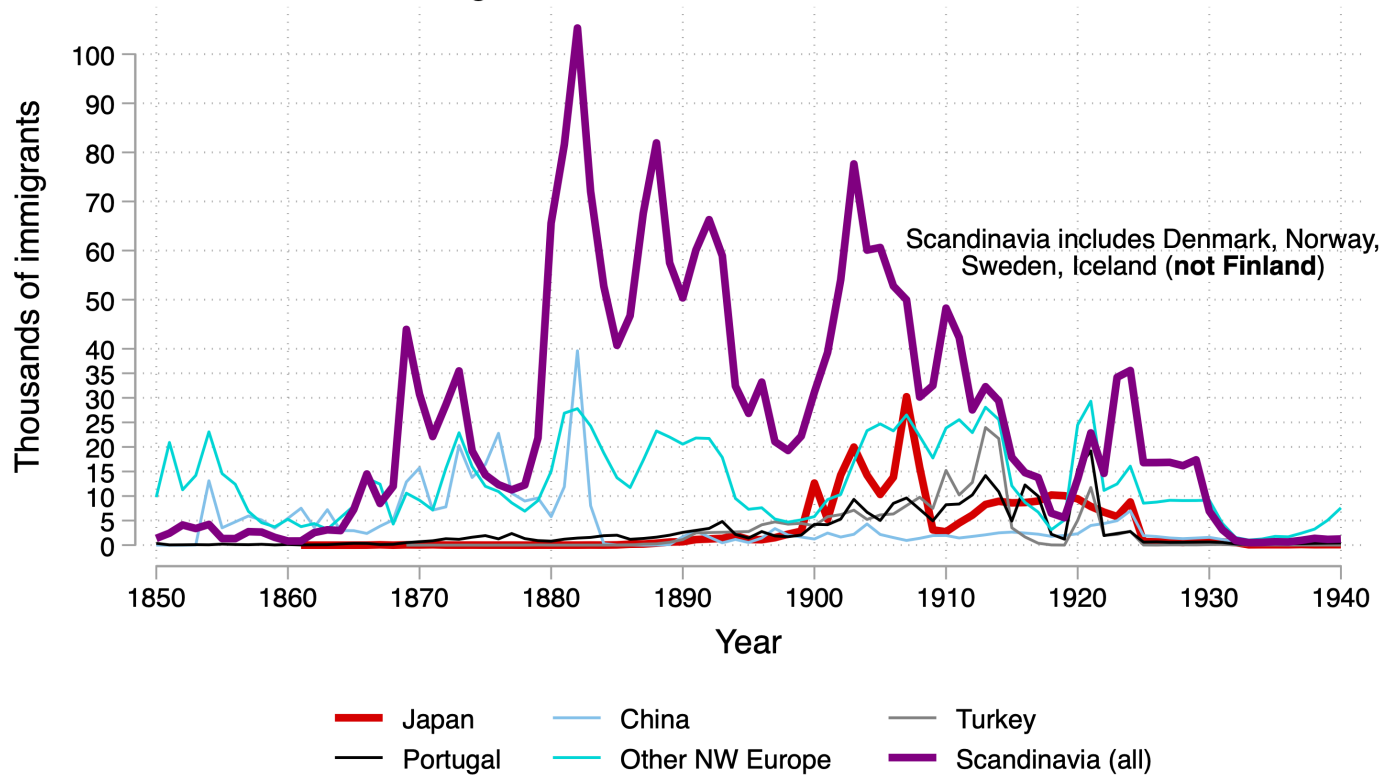
Data from *Historical Statistics of the United States*, Tables Ad106-120 and 136-48
Other Northwestern Europe includes Belgium, France, Luxembourg, Netherlands, Switzerland

LIFE COURSE CENTER

UNIVERSITY OF MINNESOTA

Share of men linked to 1940 by MLP for small immigrant groups
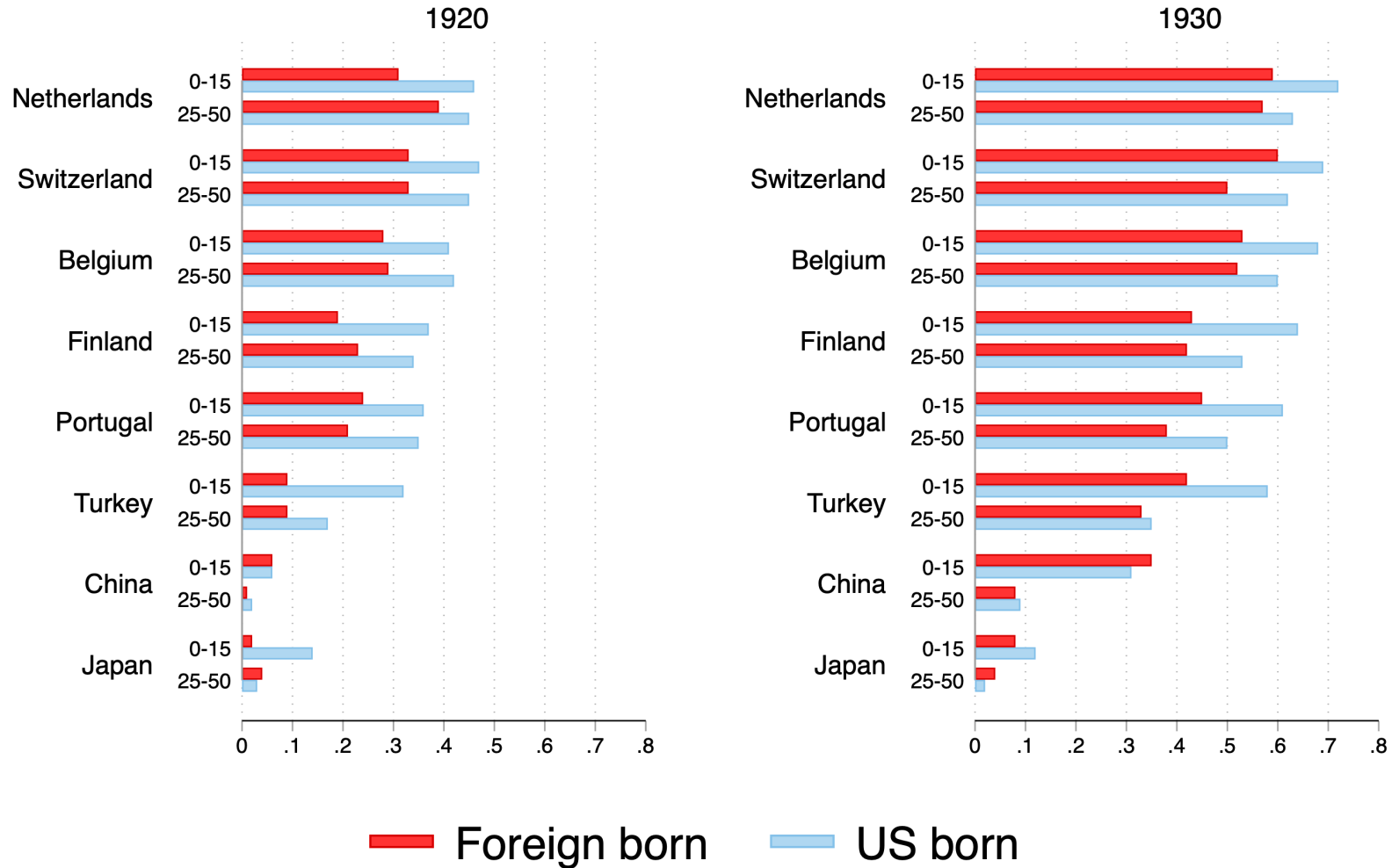
Why so low?

Entropy and possible role of migration

Japan

1900
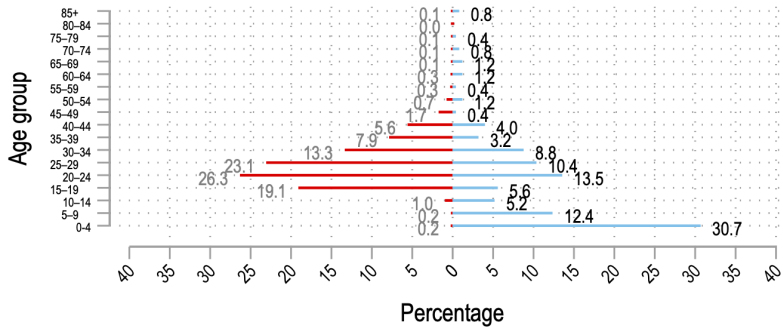
Total population: 25,315 with 25,064 born abroad and 251 in the U.S.

1910
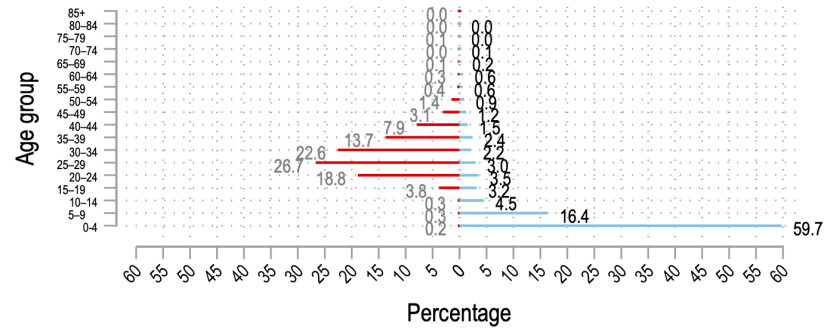
Total population: 63,653 with 60,950 born abroad and 2,703 in the U.S.

1920
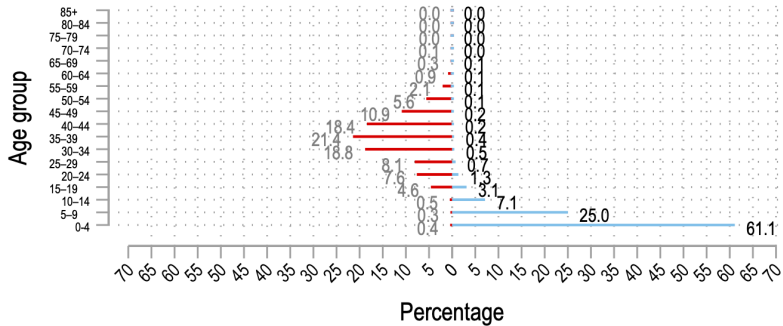
Total population: 74,129 with 58,352 born abroad and 15,777 in the U.S.

1930
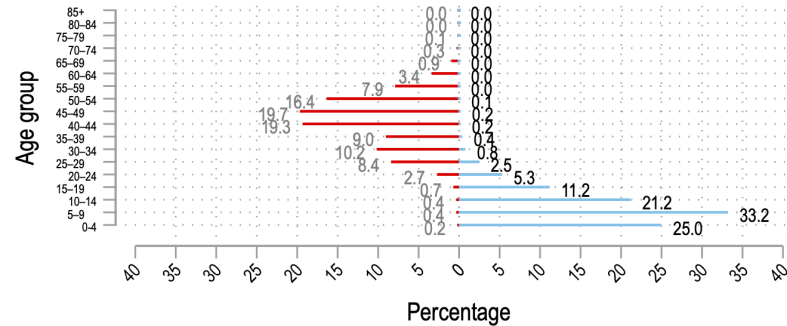
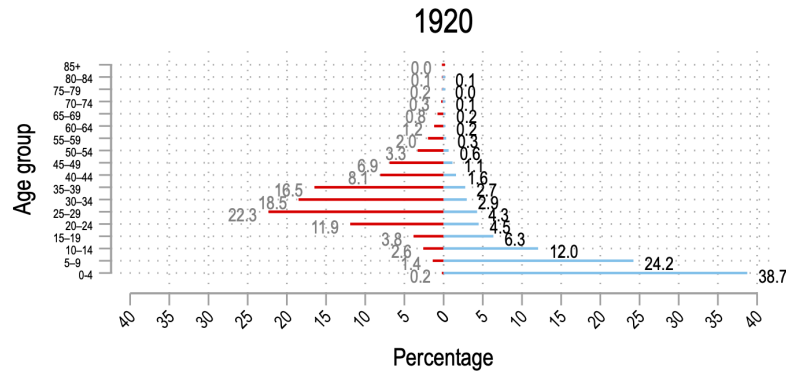Total population: 84,019 with 47,428 born abroad and 36,591 in the U.S.

Foreign born US born

Turkey

1900
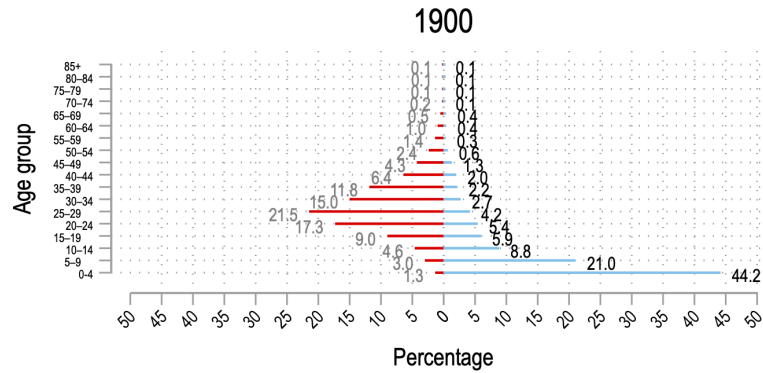
Total population: 7,254 with 5,705 born abroad and 1,549 in the U.S.

1910
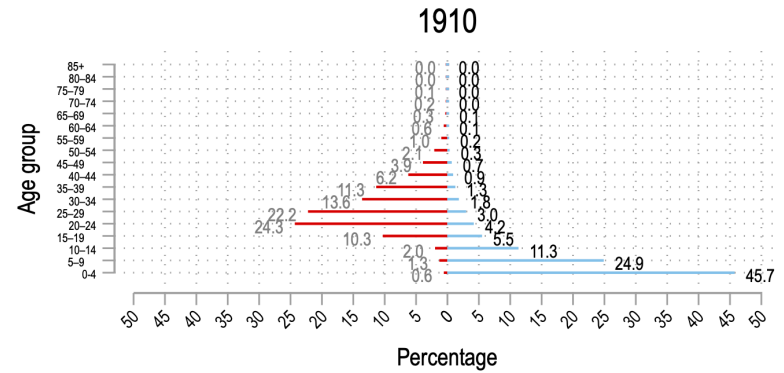
Total population: 63,922 with 53,723 born abroad and 10,199 in the U.S.

1920

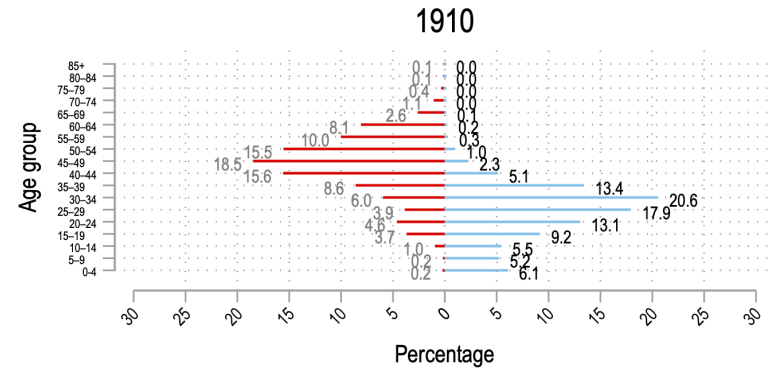Total population: 26,165 with 18,637 born abroad and 7,528 in the U.S.

1930
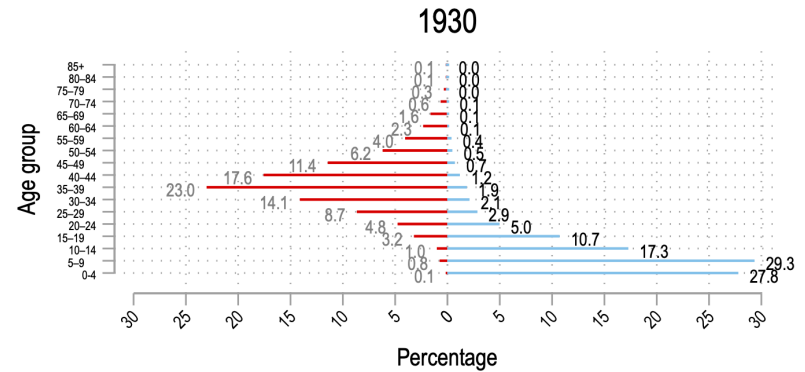
Total population: 48,312 with 29,049 born abroad and 19,263 in the U.S.
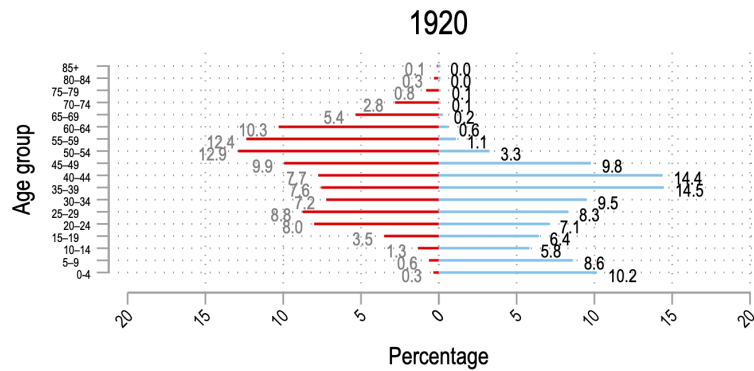
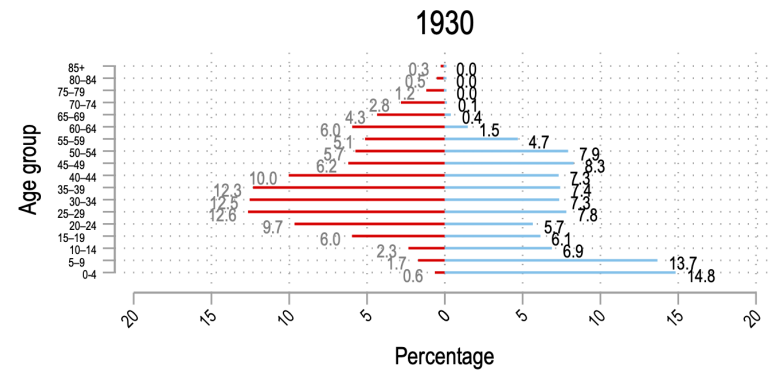Foreign born   US born

China

1900
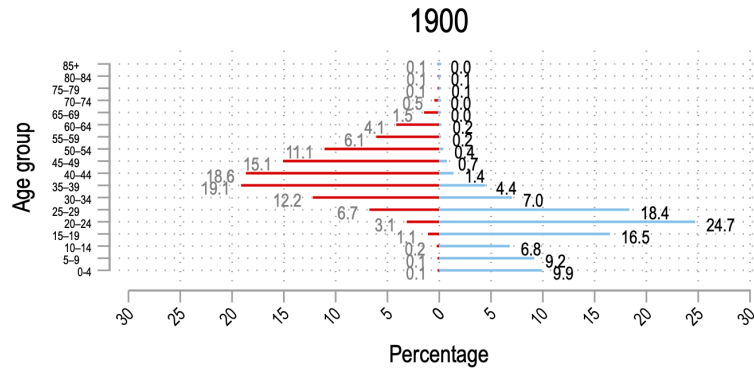
Total population: 86,863 with 79,741 born abroad and 7,122 in the U.S.

1910

Total population: 67,942 with 55,799 born abroad and 12,143 in the U.S.

1920

Total population: 54,653 with 42,158 born abroad and 12,495 in the U.S.
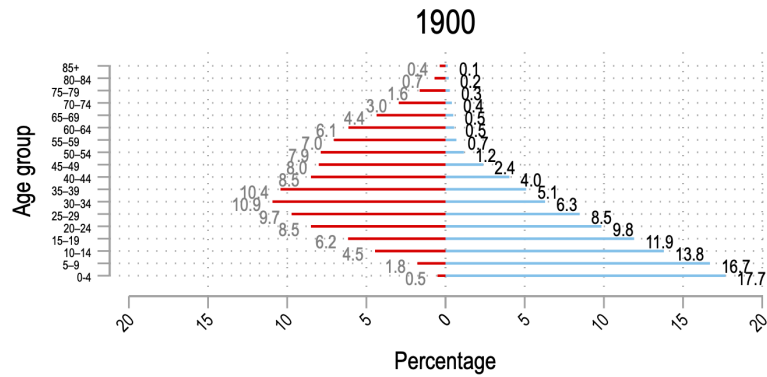
1930

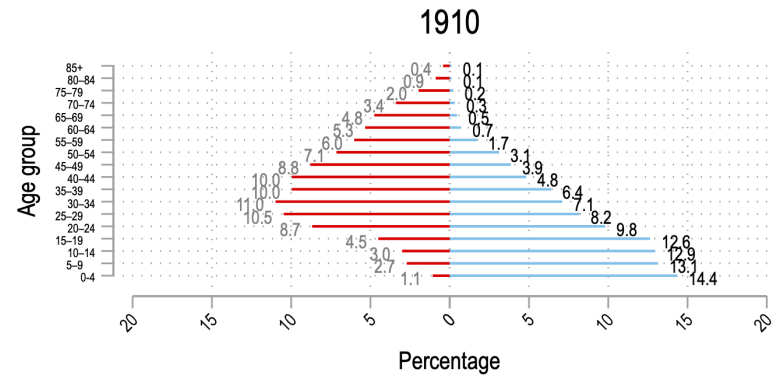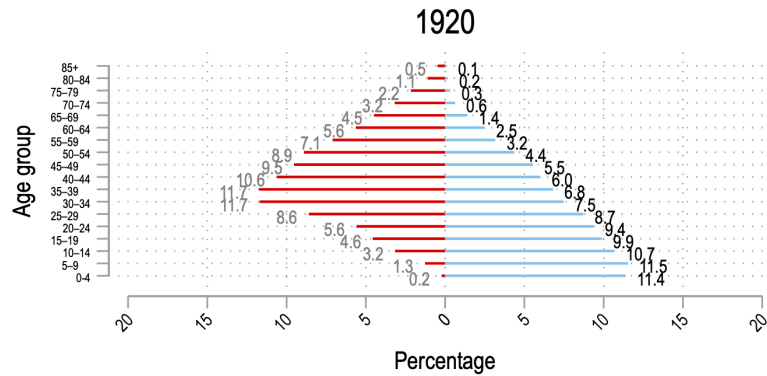Total population: 60,669 with 41,289 born abroad and 19,380 in the U.S.

Foreign born   US born
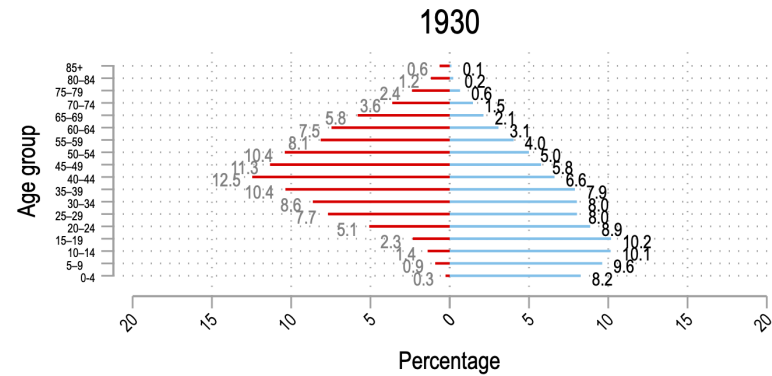
Netherlands

1900

Total population: 130621 with 53,096 born abroad and 77,525 in the U.S.

1910

Total population: 170102 with 67,583 born abroad and 102519 in the U.S.
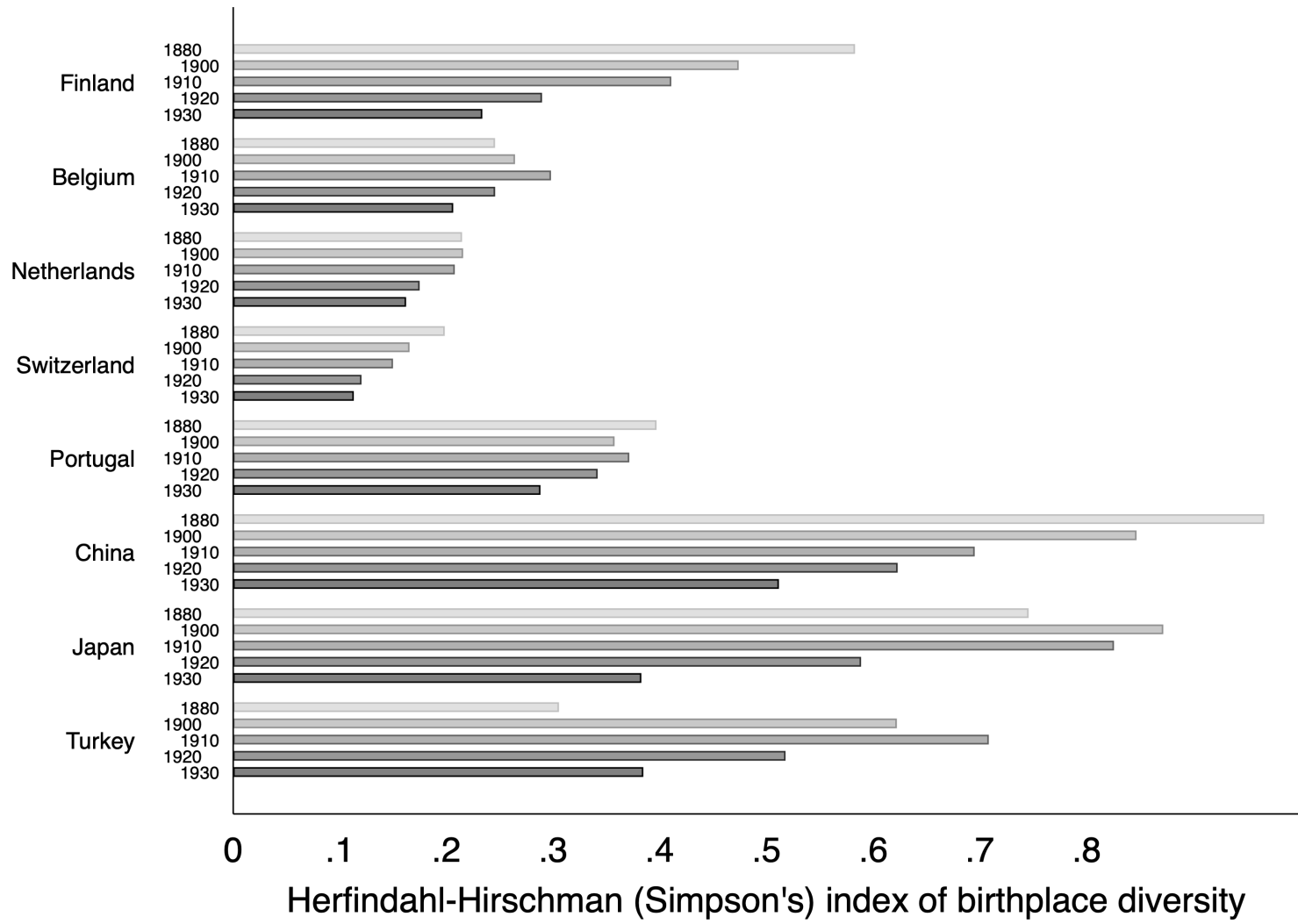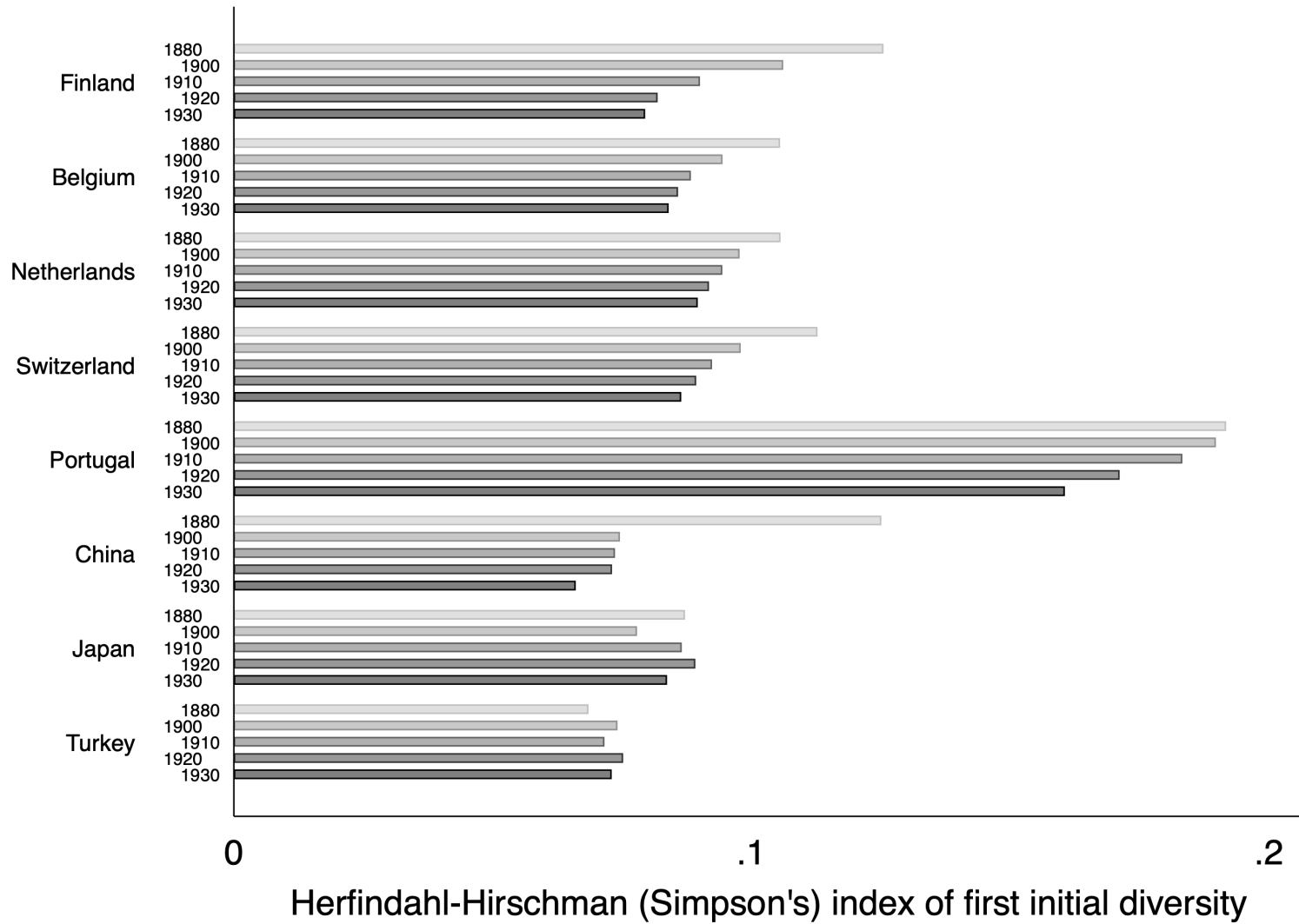
1920

Total population: 212910 with 75,576 born abroad and 137334 in the U.S.
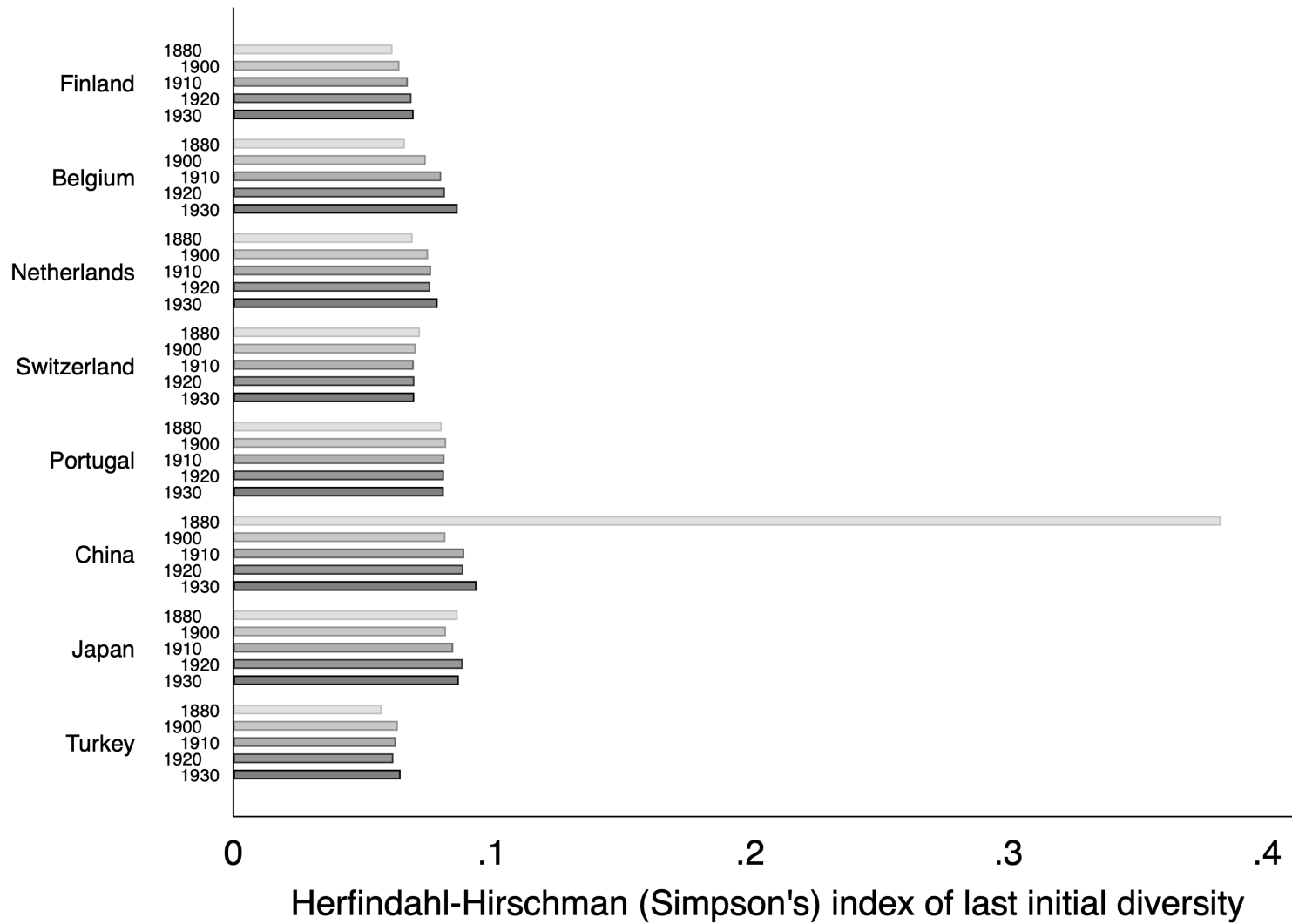
1930

Total population: 230226 with 76,599 born abroad and 153627 in the U.S.

Foreign born   US born

Herfindahl-Hirschman (Simpson's) index of birthplace diversity

Herfindahl-Hirschman (Simpson's) index of first initial diversity

Finland
- 1880
- 1900
- 1910
- 1920
- 1930

Belgium
- 1880
- 1900
- 1910
- 1920
- 1930

Netherlands
- 1880
- 1900
- 1910
- 1920
- 1930

Switzerland
- 1880
- 1900
- 1910
- 1920
- 1930

Portugal
- 1880
- 1900
- 1910
- 1920
- 1930

China
- 1880
- 1900
- 1910
- 1920
- 1930

Japan
- 1880
- 1900
- 1910
- 1920
- 1930

Turkey
- 1880
- 1900
- 1910
- 1920
- 1930

0   .1   .2   .3   .4
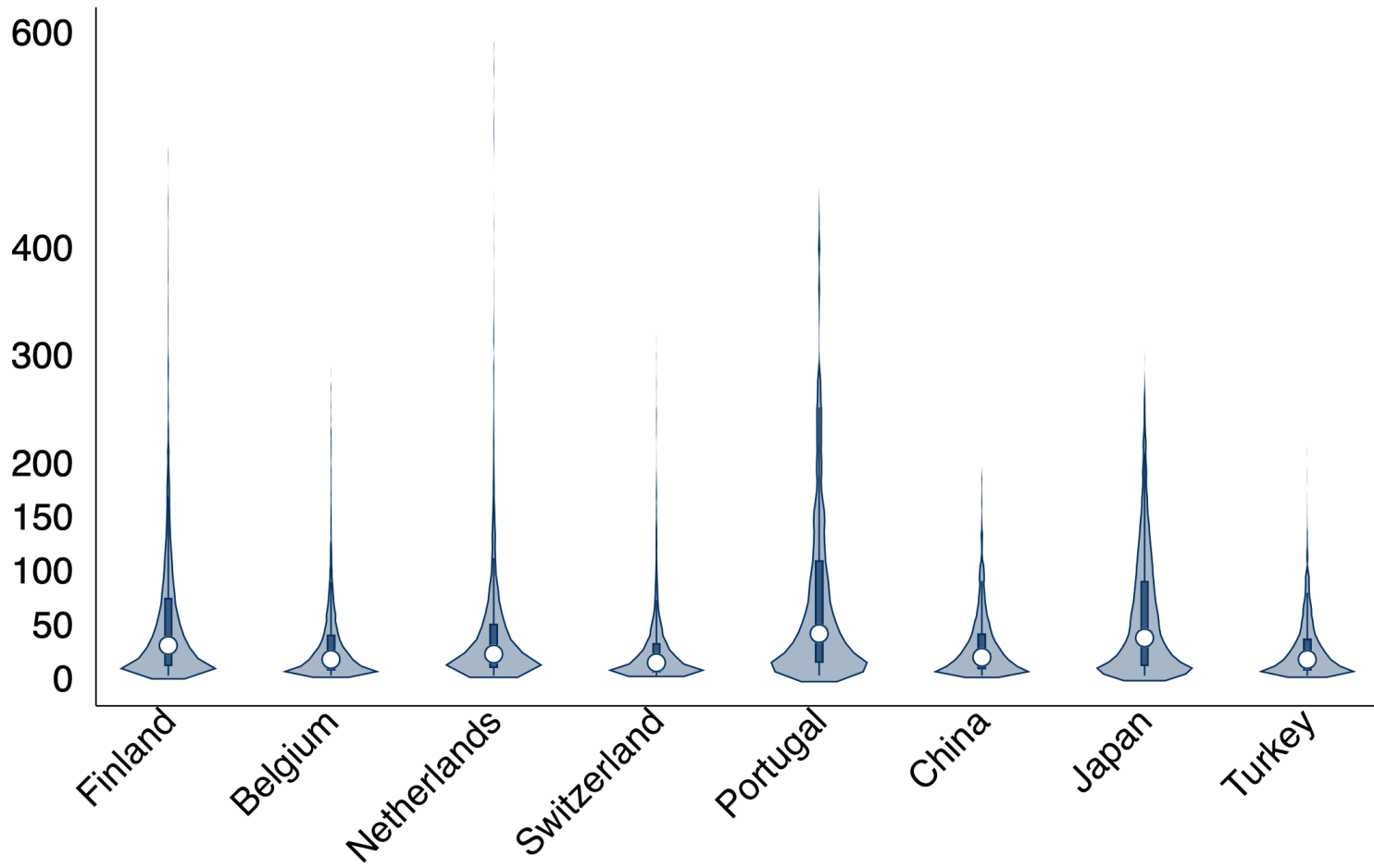
Herfindahl-Hirschman (Simpson's) index of last initial diversity

Distribution of number of comparisons by country of origin, 1930

Men only

Unlikely to be entirely migration

Birthplace	Hawaii, 1940		Contiguous 48, 1940	
Hawaii	0.76	120,718	0.02	3,018
Japan	0.24	37,430	0.37	44,871
California			0.43	53,183
Oregon			0.02	2,447
Washington			0.09	9,889
Other states			0.07	

455 of 127,000 mainland U.S. Japanese had lived in Hawaii five years earlier

Conclusion

- Distinctively low linkage rates for Japanese-Americans in the MLP samples

- Japanese-Americans concentrated in birthplace-birth cohort-name space