

Online Appendix for
Working Their Way Up?
US Immigrants' Changing Labor Market Assimilation in the Age of Mass Migration

William J. Collins and Ariell Zimran
Vanderbilt University & NBER

December 2021

A. Additional tables and figures	42
B. Census linkage	55
C. Results with alternate linkage methods	60
D. Results with imputed occupation codes	93
E. Procedure to recategorize farm family members in 1850	102
F. Results with 20-year linkage spans	103
G. Results with alternate weighting methods	112
H. Results omitting natives with foreign-born fathers	121
I. Results omitting immigrants more than 10 years from arrival	130
J. Results with alternate occupational wealth scores	139
K. Results limited to English speakers	152
L. References	161

Appendix A: Additional Tables and Figures

Table A.1: Summary statistics

	(1)	(2)	(3)	(4)	(5)
	1850–1880	1870–1900	1880–1910	1900–1930	1910–1940
<i>Panel A: Natives</i>					
Initial Rank	0.448 (0.276)	0.411 (0.279)	0.415 (0.271)	0.412 (0.283)	0.462 (0.272)
Final Rank	0.595 (0.203)	0.573 (0.233)	0.570 (0.243)	0.545 (0.254)	0.548 (0.247)
Initial Age	23.636 (3.761)	22.764 (4.399)	22.764 (4.282)	23.018 (4.385)	23.841 (3.698)
Initial Region					
Northeast	0.596	0.448	0.408	0.360	0.339
Midwest	0.371	0.519	0.547	0.562	0.535
West	0.033	0.033	0.045	0.077	0.126
<i>Panel B: Immigrants</i>					
Initial Rank	0.399 (0.258)	0.437 (0.254)	0.436 (0.254)	0.433 (0.253)	0.389 (0.240)
Final Rank	0.525 (0.240)	0.555 (0.234)	0.543 (0.243)	0.508 (0.259)	0.491 (0.265)
Initial Age	24.817 (3.710)	24.691 (4.217)	24.659 (4.344)	24.696 (4.234)	24.789 (3.553)
Initial Region					
Northeast	0.673	0.414	0.375	0.493	0.546
Midwest	0.307	0.531	0.550	0.432	0.330
West	0.020	0.054	0.075	0.075	0.123
Natives	152,008	169,085	253,545	413,687	689,517
Immigrants	21,896	31,127	35,184	60,849	105,172
Total	173,904	200,212	288,729	474,536	794,689

Notes: Standard deviations in parentheses. All statistics weighted by inverse linkage probability. Ranks are based on the average rank measure used for the main results.

Table A.2: Dissimilarity indices

	(1)	(2)
	Initial	Final
1850–1880	0.2262 (0.0030)	0.2054 (0.0042)
Observations	166,576	166,576
1870–1900	0.1598 (0.0041)	0.1173 (0.0047)
Observations	200,310	200,310
1880–1910	0.1426 (0.0044)	0.1082 (0.0044)
Observations	288,849	288,849
1900–1930	0.1487 (0.0036)	0.1216 (0.0035)
Observations	474,777	474,777
1910–1940	0.2315 (0.0012)	0.1729 (0.0018)
Observations	737,371	737,371

Notes: This table presents the dissimilarity indices of Figure 7, which control for age. Observations are weighted to correct for selection into linkage. Robust standard errors in parentheses.

Table A.3: Occupational category differences

	(1) Unskill	(2) Farmer	(3) Craft	(4) Operative	(5) White Collar
<i>Panel A: 1850–1880</i>					
Initial	0.140 ^a (0.004)	−0.202 ^a (0.002)	0.031 ^a (0.003)	0.055 ^a (0.003)	−0.024 ^a (0.002)
Final	0.122 ^a (0.004)	−0.181 ^a (0.004)	0.022 ^a (0.003)	0.061 ^a (0.003)	−0.024 ^a (0.003)
Difference	−0.018 ^a (0.005)	0.021 ^a (0.004)	−0.009 ^b (0.004)	0.007 (0.004)	−0.000 (0.003)
<i>Panel B: 1870–1900</i>					
Initial	0.010 ^b (0.004)	−0.145 ^a (0.003)	0.060 ^a (0.004)	0.090 ^a (0.004)	−0.014 ^a (0.003)
Final	0.038 ^a (0.004)	−0.092 ^a (0.004)	0.030 ^a (0.004)	0.050 ^a (0.003)	−0.025 ^a (0.003)
Difference	0.028 ^a (0.006)	0.053 ^a (0.005)	−0.030 ^a (0.005)	−0.040 ^a (0.005)	−0.011 ^a (0.004)
<i>Panel C: 1880–1910</i>					
Initial	−0.002 (0.004)	−0.120 ^a (0.003)	0.048 ^a (0.003)	0.095 ^a (0.004)	−0.021 ^a (0.003)
Final	0.026 ^a (0.004)	−0.073 ^a (0.004)	0.028 ^a (0.004)	0.053 ^a (0.003)	−0.035 ^a (0.003)
Difference	0.028 ^a (0.005)	0.047 ^a (0.004)	−0.019 ^a (0.004)	−0.041 ^a (0.004)	−0.014 ^a (0.004)
<i>Panel D: 1900–1930</i>					
Initial	−0.011 ^a (0.004)	−0.092 ^a (0.002)	0.058 ^a (0.003)	0.091 ^a (0.003)	−0.046 ^a (0.003)
Final	0.039 ^a (0.003)	−0.082 ^a (0.003)	0.028 ^a (0.003)	0.054 ^a (0.003)	−0.039 ^a (0.003)
Difference	0.050 ^a (0.005)	0.009 ^a (0.003)	−0.030 ^a (0.004)	−0.036 ^a (0.004)	0.007 ^b (0.003)
<i>Panel E: 1910–1940</i>					
Initial	0.114 ^a (0.002)	−0.113 ^a (0.001)	0.007 ^a (0.001)	0.111 ^a (0.002)	−0.119 ^a (0.001)
Final	0.071 ^a (0.002)	−0.104 ^a (0.001)	0.032 ^a (0.002)	0.070 ^a (0.002)	−0.069 ^a (0.002)
Difference	−0.043 ^a (0.002)	0.008 ^a (0.001)	0.025 ^a (0.002)	−0.040 ^a (0.002)	0.050 ^a (0.002)

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.10

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. The first two rows of each panel present the results from estimating equation (1) with the occupational category indicator in the column header. The last row of each panel presents estimates of $\beta_t - \beta_{t-30}$. Farm family members are included in the “unskilled” category.

Table A.4: Unconditional Assimilation

Table A.4(a): Ranking farm family equal to farm laborers

	(1) Avg Rank	(2) log(Occ. Wealth)	(3) log(PH Score)	(4) Occ. Wealth Rank	(5) PH Rank
<i>Panel A: 1850–1880</i>					
Initial	−0.072 ^a (0.002)	−0.423 ^a (0.008)	−0.030 ^a (0.003)	−0.107 ^a (0.002)	−0.037 ^a (0.003)
Final	−0.069 ^a (0.002)	−0.382 ^a (0.009)	−0.051 ^a (0.003)	−0.102 ^a (0.002)	−0.036 ^a (0.002)
Difference	0.003 (0.003)	0.041 ^a (0.011)	−0.020 ^a (0.004)	0.005 ^c (0.003)	0.001 (0.003)
<i>Panel B: 1870–1900</i>					
Initial	−0.007 ^a (0.002)	−0.168 ^a (0.009)	0.041 ^a (0.004)	−0.041 ^a (0.002)	0.027 ^a (0.003)
Final	−0.022 ^a (0.002)	−0.157 ^a (0.010)	−0.003 (0.003)	−0.039 ^a (0.003)	−0.005 ^c (0.002)
Difference	−0.015 ^a (0.003)	0.011 (0.012)	−0.044 ^a (0.004)	0.002 (0.003)	−0.031 ^a (0.003)
<i>Panel C: 1880–1910</i>					
Initial	−0.009 ^a (0.002)	−0.130 ^a (0.009)	0.022 ^a (0.003)	−0.034 ^a (0.002)	0.016 ^a (0.003)
Final	−0.021 ^a (0.002)	−0.134 ^a (0.009)	−0.009 ^a (0.003)	−0.033 ^a (0.003)	−0.009 ^a (0.003)
Difference	−0.012 ^a (0.003)	−0.004 (0.012)	−0.032 ^a (0.004)	0.001 (0.003)	−0.026 ^a (0.003)
<i>Panel D: 1900–1930</i>					
Initial	−0.004 ^b (0.002)	−0.103 ^a (0.007)	0.042 ^a (0.003)	−0.019 ^a (0.002)	0.010 ^a (0.002)
Final	−0.029 ^a (0.002)	−0.134 ^a (0.008)	−0.016 ^a (0.003)	−0.039 ^a (0.002)	−0.019 ^a (0.002)
Difference	−0.025 ^a (0.002)	−0.031 ^a (0.009)	−0.059 ^a (0.003)	−0.020 ^a (0.003)	−0.029 ^a (0.003)
<i>Panel E: 1910–1940</i>					
Initial	−0.086 ^a (0.001)	−0.357 ^a (0.003)	−0.056 ^a (0.001)	−0.092 ^a (0.001)	−0.081 ^a (0.001)
Final	−0.057 ^a (0.001)	−0.225 ^a (0.004)	−0.051 ^a (0.001)	−0.070 ^a (0.001)	−0.043 ^a (0.001)
Difference	0.029 ^a (0.001)	0.132 ^a (0.005)	0.005 ^a (0.002)	0.021 ^a (0.001)	0.038 ^a (0.001)

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.10

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. The first two rows of each panel present the results from estimating equation (1) with the occupational rank or score in the column header. The last row of each table presents estimates of $\beta_t - \beta_{t-30}$.

Table A.4(b): Ranking farm family at the midpoint of farm laborers and farmers

	(1) Avg Rank	(2) log(Occ. Wealth)	(3) log(PH Score)	(4) Occ. Wealth Rank	(5) PH Rank
<i>Panel A: 1850–1880</i>					
Initial	−0.118 ^a (0.002)	−0.586 ^a (0.008)	−0.074 ^a (0.003)	−0.170 ^a (0.002)	−0.066 ^a (0.003)
Final	−0.080 ^a (0.002)	−0.379 ^a (0.009)	−0.050 ^a (0.003)	−0.115 ^a (0.003)	−0.044 ^a (0.002)
Difference	0.038 ^a (0.003)	0.207 ^a (0.011)	0.024 ^a (0.003)	0.055 ^a (0.003)	0.021 ^a (0.003)
<i>Panel B: 1870–1900</i>					
Initial	−0.039 ^a (0.002)	−0.304 ^a (0.009)	0.004 (0.003)	−0.088 ^a (0.002)	0.010 ^a (0.003)
Final	−0.026 ^a (0.002)	−0.164 ^a (0.009)	−0.005 ^c (0.003)	−0.046 ^a (0.003)	−0.007 ^a (0.003)
Difference	0.012 ^a (0.003)	0.139 ^a (0.012)	−0.009 ^b (0.004)	0.042 ^a (0.003)	−0.017 ^a (0.003)
<i>Panel C: 1880–1910</i>					
Initial	−0.030 ^a (0.002)	−0.212 ^a (0.008)	0.000 (0.003)	−0.068 ^a (0.002)	0.007 ^a (0.003)
Final	−0.023 ^a (0.002)	−0.136 ^a (0.009)	−0.010 ^a (0.003)	−0.037 ^a (0.003)	−0.010 ^a (0.003)
Difference	0.007 ^b (0.003)	0.076 ^a (0.011)	−0.010 ^a (0.004)	0.031 ^a (0.003)	−0.017 ^a (0.003)
<i>Panel D: 1900–1930</i>					
Initial	−0.044 ^a (0.002)	−0.267 ^a (0.007)	−0.002 (0.002)	−0.077 ^a (0.002)	−0.010 ^a (0.002)
Final	−0.032 ^a (0.002)	−0.140 ^a (0.008)	−0.018 ^a (0.003)	−0.043 ^a (0.002)	−0.021 ^a (0.002)
Difference	0.012 ^a (0.002)	0.126 ^a (0.009)	−0.016 ^a (0.003)	0.034 ^a (0.003)	−0.011 ^a (0.003)
<i>Panel E: 1910–1940</i>					
Initial	−0.121 ^a (0.001)	−0.490 ^a (0.003)	−0.092 ^a (0.001)	−0.143 ^a (0.001)	−0.100 ^a (0.001)
Final	−0.058 ^a (0.001)	−0.226 ^a (0.004)	−0.051 ^a (0.001)	−0.073 ^a (0.001)	−0.044 ^a (0.001)
Difference	0.063 ^a (0.001)	0.264 ^a (0.005)	0.041 ^a (0.002)	0.070 ^a (0.001)	0.056 ^a (0.001)

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.10

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. The first two rows of each panel present the results from estimating equation (1) with the occupational rank or score in the column header. The last row of each table presents estimates of $\beta_t - \beta_{t-30}$.

Table A.4(c): Ranking farm family equal to farmers

	(1) Avg Rank	(2) log(Occ. Wealth)	(3) log(PH Score)	(4) Occ. Wealth Rank	(5) PH Rank
<i>Panel A: 1850–1880</i>					
Initial	−0.132 ^a (0.002)	−0.641 ^a (0.008)	−0.103 ^a (0.003)	−0.181 ^a (0.002)	−0.083 ^a (0.003)
Final	−0.075 ^a (0.002)	−0.379 ^a (0.009)	−0.049 ^a (0.003)	−0.109 ^a (0.003)	−0.040 ^a (0.002)
Difference	0.057 ^a (0.003)	0.263 ^a (0.011)	0.054 ^a (0.003)	0.072 ^a (0.003)	0.043 ^a (0.003)
<i>Panel B: 1870–1900</i>					
Initial	−0.052 ^a (0.002)	−0.349 ^a (0.009)	−0.020 ^a (0.003)	−0.099 ^a (0.003)	−0.005 ^c (0.003)
Final	−0.026 ^a (0.002)	−0.166 ^a (0.009)	−0.006 ^b (0.003)	−0.045 ^a (0.003)	−0.007 ^a (0.003)
Difference	0.026 ^a (0.003)	0.183 ^a (0.012)	0.014 ^a (0.004)	0.054 ^a (0.003)	−0.002 (0.003)
<i>Panel C: 1880–1910</i>					
Initial	−0.037 ^a (0.002)	−0.239 ^a (0.009)	−0.014 ^a (0.003)	−0.073 ^a (0.002)	−0.002 (0.003)
Final	−0.023 ^a (0.002)	−0.136 ^a (0.009)	−0.010 ^a (0.003)	−0.036 ^a (0.003)	−0.010 ^a (0.003)
Difference	0.014 ^a (0.003)	0.103 ^a (0.011)	0.004 (0.004)	0.037 ^a (0.003)	−0.009 ^a (0.003)
<i>Panel D: 1900–1930</i>					
Initial	−0.059 ^a (0.002)	−0.321 ^a (0.007)	−0.031 ^a (0.002)	−0.090 ^a (0.002)	−0.028 ^a (0.002)
Final	−0.032 ^a (0.002)	−0.143 ^a (0.008)	−0.019 ^a (0.003)	−0.043 ^a (0.002)	−0.021 ^a (0.002)
Difference	0.027 ^a (0.002)	0.179 ^a (0.009)	0.012 ^a (0.003)	0.047 ^a (0.003)	0.007 ^a (0.003)
<i>Panel E: 1910–1940</i>					
Initial	−0.133 ^a (0.001)	−0.535 ^a (0.003)	−0.116 ^a (0.001)	−0.151 ^a (0.001)	−0.114 ^a (0.001)
Final	−0.058 ^a (0.001)	−0.227 ^a (0.004)	−0.052 ^a (0.001)	−0.072 ^a (0.001)	−0.044 ^a (0.001)
Difference	0.075 ^a (0.001)	0.308 ^a (0.005)	0.064 ^a (0.002)	0.079 ^a (0.001)	0.071 ^a (0.001)

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.10

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. The first two rows of each panel present the results from estimating equation (1) with the occupational rank or score in the column header. The last row of each table presents estimates of $\beta_t - \beta_{t-30}$.

Table A.5: Unconditional and conditional assimilation based on average rank

Table A.5(a): Ranking farm family equal to farm laborers

	(1) Unconditional	(2) Conditional
1850–1880	0.002 (0.003)	−0.059 ^a (0.002)
1870–1900	−0.014 ^a (0.003)	−0.020 ^a (0.002)
1880–1910	−0.013 ^a (0.003)	−0.020 ^a (0.002)
1900–1930	−0.025 ^a (0.002)	−0.029 ^a (0.002)
1910–1940	0.029 ^a (0.001)	−0.033 ^a (0.001)

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.10

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. Column (1) presents estimates of equation (2) without controlling for initial rank. Column (2) presents estimates of equation (2) with controls for initial rank.

Table A.5(b): Ranking farm family at the midpoint of farmers and farm laborers

	(1) Unconditional	(2) Conditional
1850–1880	0.037 ^a (0.003)	−0.058 ^a (0.002)
1870–1900	0.013 ^a (0.003)	−0.017 ^a (0.002)
1880–1910	0.006 ^b (0.003)	−0.016 ^a (0.002)
1900–1930	0.011 ^a (0.002)	−0.019 ^a (0.002)
1910–1940	0.062 ^a (0.001)	−0.020 ^a (0.001)

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.10

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. Column (1) presents estimates of equation (2) without controlling for initial rank. Column (2) presents estimates of equation (2) with controls for initial rank.

Table A.5(c): Ranking farm family equal to farmers

	(1) Unconditional	(2) Conditional
1850–1880	0.057 ^a (0.003)	–0.050 ^a (0.002)
1870–1900	0.026 ^a (0.003)	–0.015 ^a (0.002)
1880–1910	0.013 ^a (0.003)	–0.014 ^a (0.002)
1900–1930	0.026 ^a (0.002)	–0.016 ^a (0.002)
1910–1940	0.074 ^a (0.001)	–0.018 ^a (0.001)

Significance levels: ^a $p < 0.01$, ^b $p < 0.05$, ^c $p < 0.10$

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. Column (1) presents estimates of equation (2) without controlling for initial rank. Column (2) presents estimates of equation (2) with controls for initial rank.

Figure A.1: Occupational upgrading by initial occupation, ranking farm family as farm laborers

Figure A.1(a): 1850-1880

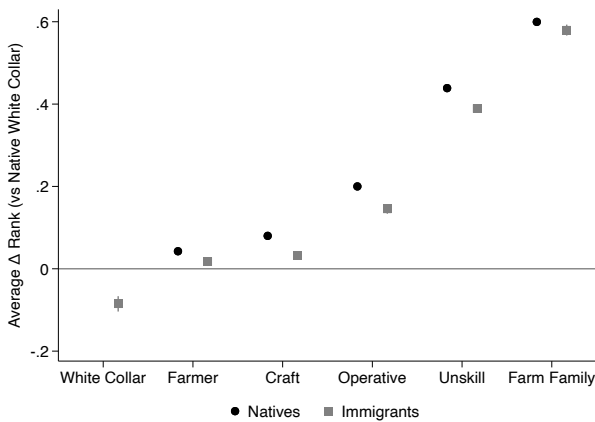


Figure A.1(b): 1870-1900

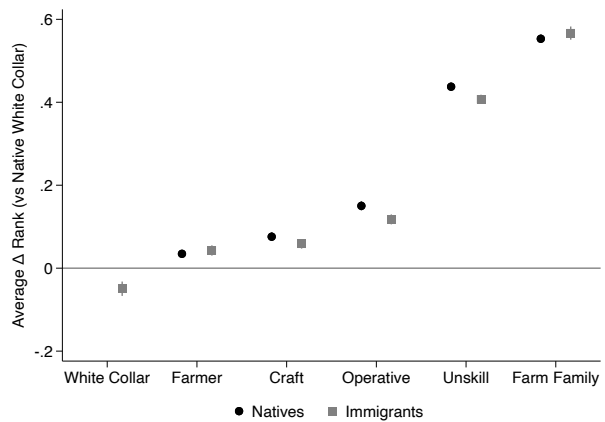


Figure A.1(c): 1880-1910

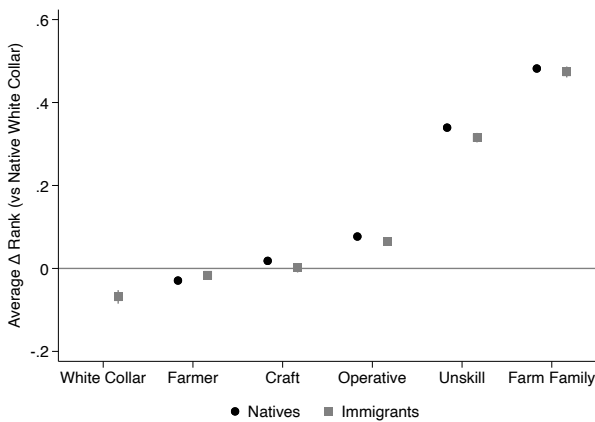


Figure A.1(d): 1900-1930

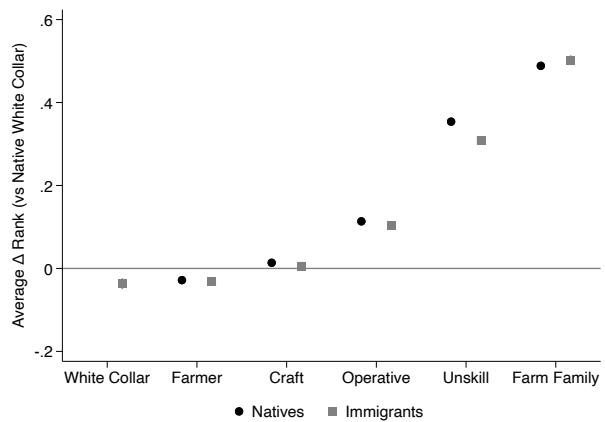
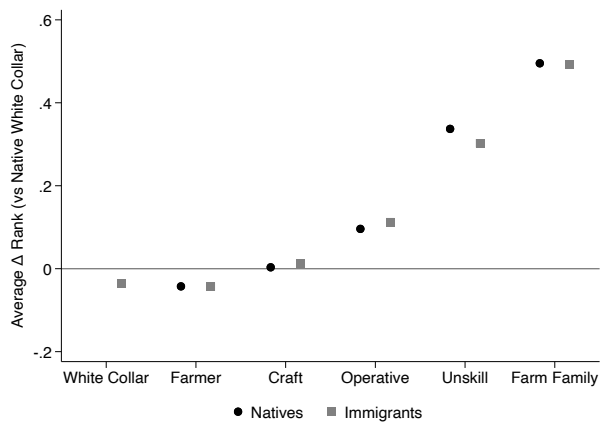


Figure A.1(e): 1910-1940



Notes: These graphs express the average upgrading experienced by immigrants or natives in the each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the level of farm laborers.

Figure A.2: Occupational upgrading by initial occupation, ranking farm family as farmers

Figure A.2(a): 1850-1880

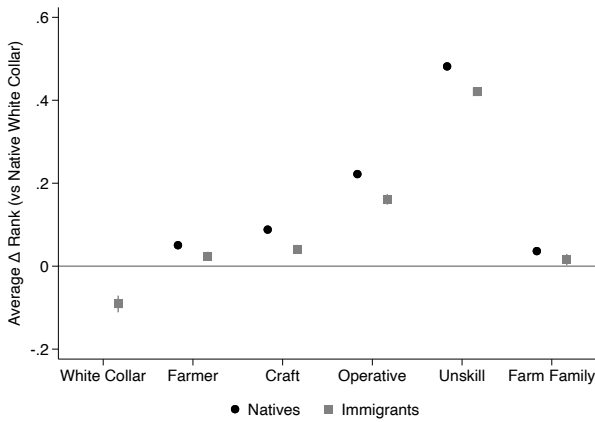


Figure A.2(b): 1870-1900

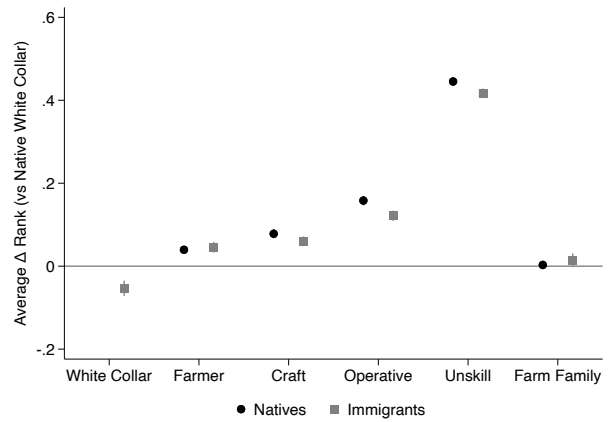


Figure A.2(c): 1880-1910

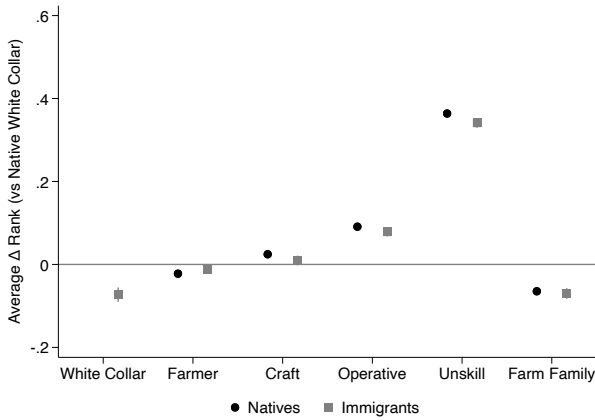


Figure A.2(d): 1900-1930

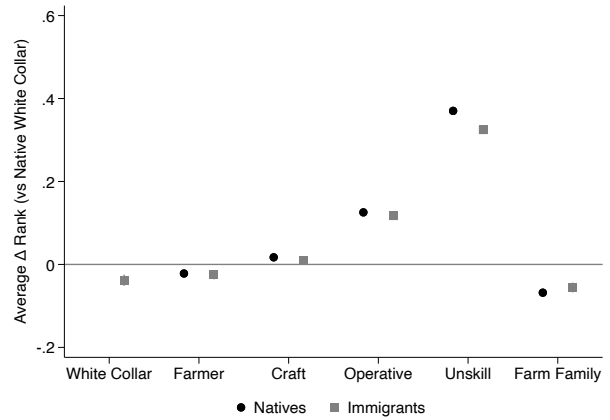
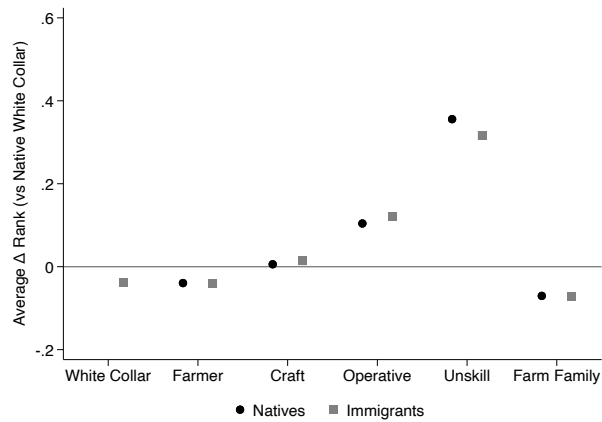
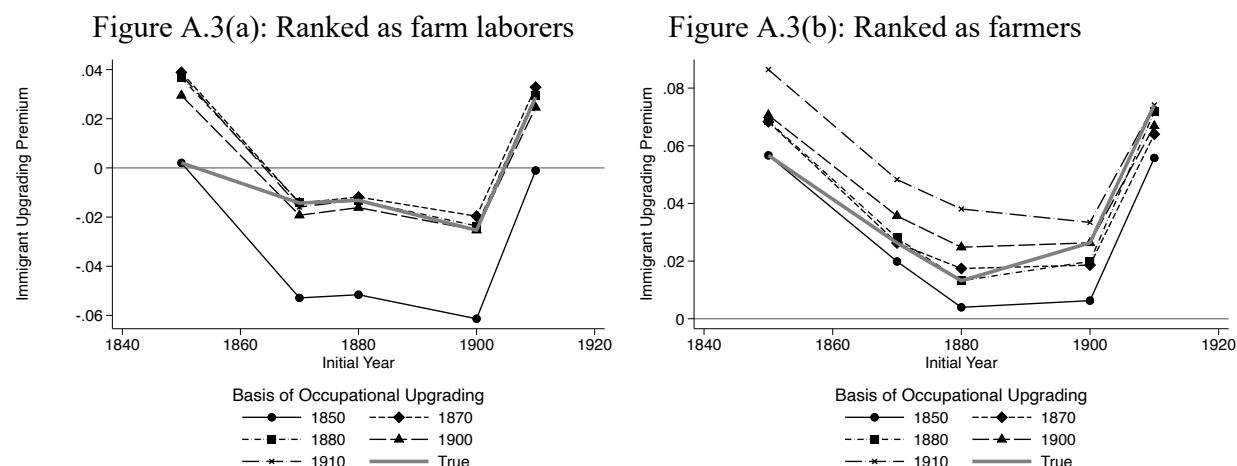


Figure A.2(e): 1910-1940



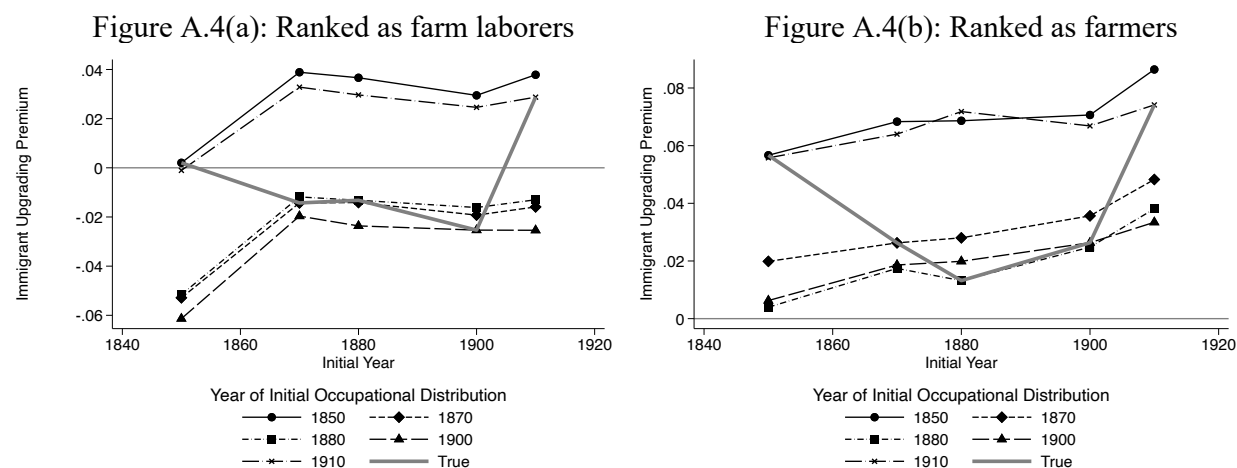
Notes: These graphs express the average upgrading experienced by immigrants or natives in the each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the level of farmers.

Figure A.3: Unconditional assimilation, holding occupational upgrading constant, alternate ranking of farm family members



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x-axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the level indicated in the figure title.

Figure A.4: Unconditional assimilation, holding occupational distributions constant, alternate ranking of farm family members



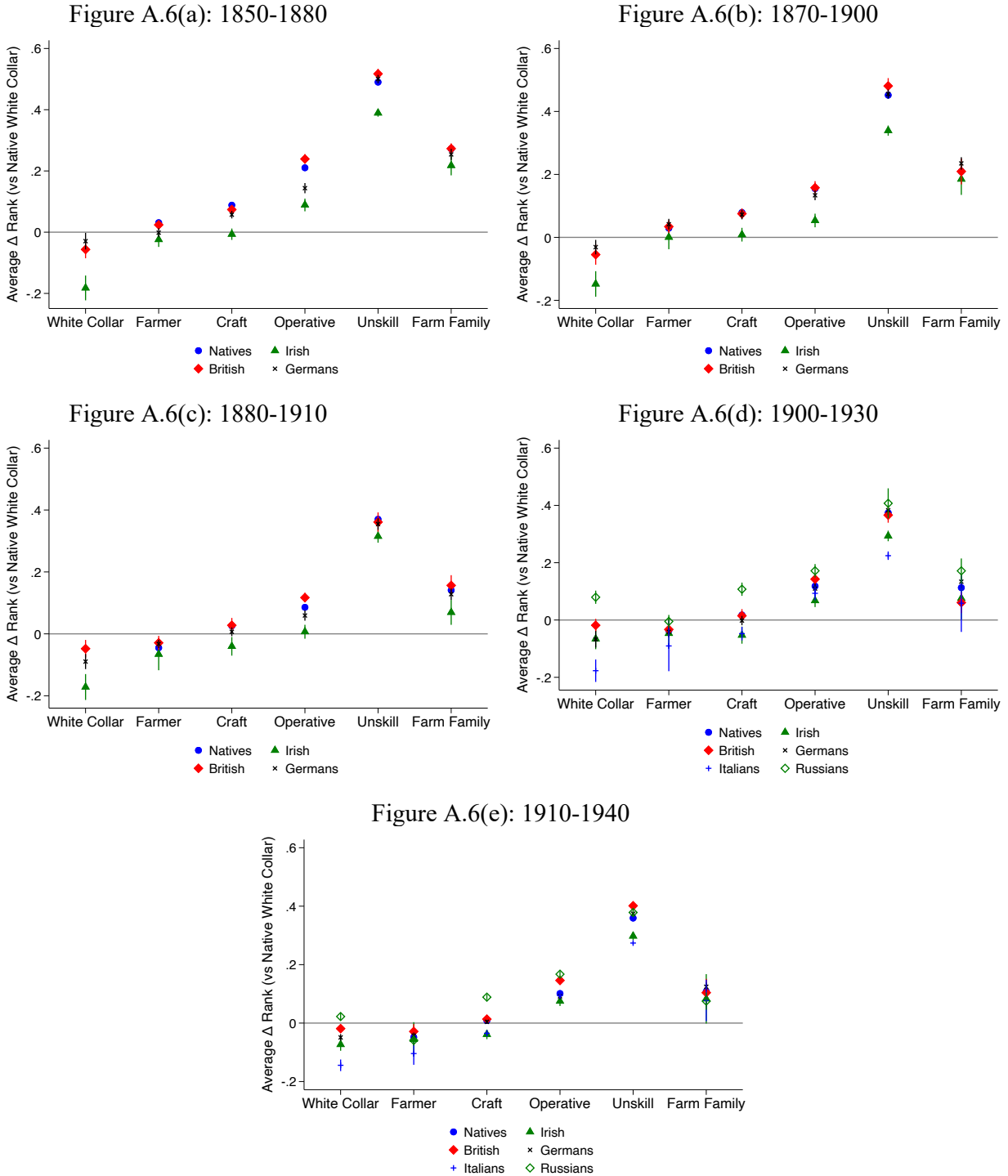
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the level indicated in the figure title.

Figure A.5: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x-axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the level indicated in the figure title.

Figure A.6: Occupational upgrading by initial occupation and source country



Notes: This graph presents estimates analogous to Figure 9, showing the average occupational upgrade by nativity and initial occupation, relative to natives initially in white collar occupations. Unlike Figure 9, however, this figure divides immigrants according to major country-of-origin groups. Farm family members are ranked at the midpoint of farmers and farm laborers.

Appendix B: Census linkage

Census Linkage Procedure

The detailed linkage procedure is as follows. This description is based heavily on the one provided by Collins and Zimran (2019).

1. We extracted males with non-empty first and last names from the full count initial year census and removed punctuation.
2. We divided the listed first name into a given name and middle initial, when one was present.
3. We replaced standard first name abbreviations (e.g., “Wm” was replaced with “William”).
4. We removed any remaining spaces removed from the names.
5. We linked the initial census to itself according to the following criteria.
 - a. The birthplace (country or US state) matches exactly.⁴⁷
 - b. The absolute difference in birth years is less than or equal to 4.
 - c. The first three characters of last name soundex match, and either of the following two conditions are true:
 - i. The last name soundexes are identical and the SAS spelling distance (using the SAS function SPEDIS) is less than or equal to 20.⁴⁸
 - ii. The last name soundexes are non-identical, and the SAS spelling distance is less than or equal to 17.
 - d. First letter of first name matches, and spelling distance between first names is less than or equal to 20.
6. We removed from the sample any individual who had a candidate match in step 5 that was someone other than himself.
7. The remaining men from the initial census and all those from the final census were cleaned according to steps 1, 2, and 4 above (no standardization of name abbreviations was made).
8. We linked males from the initial to the final year according to the same criteria listed in step 5, as well as the following additional criterion: where both records report a middle initial, the middle initials must match for a match to be made.

⁴⁷ For UK-born men (i.e., men born in England, Scotland, or Wales) in 1900-30, we require uniqueness within all UK-born men. In step 8, however, we require a match on the specific birthplace. This distinction is not made in 1850-80 because the data report only the UK as the country of birth.

⁴⁸ The SAS spelling distance is not a symmetric measure. Whenever we use spelling distance, we make a match when $\min\{\text{spedis}_{a,b}, \text{spedis}_{b,a}\} \leq c$, where c is the linkage cutoff. The SPEDIS function is described in detail by Gershteyn (2000).

9. We dropped any case in which more than one final-year individual matched to an initial-year individual or vice versa.

Construction of linkage weights

We construct the weights as follows. First, we designate a base year and a base population at risk for linkage for each sample. Next, we estimate a probit regression of the form

$$P(y_i = 1) = \Phi(\mathbf{x}_i' \beta + \varepsilon_i),$$

where y_i is an indicator taking a value of 1 if individual i was successfully linked between the two censuses of each span and 0 otherwise and \mathbf{x}_i is a vector of observables in the base year. For each linked individual, we compute

$$\hat{p}_i = \Phi(\mathbf{x}_i' \hat{\beta})$$

and construct inverse probability weights of the form $1/\hat{p}_i$.

Summary statistics

Tables B.1 and B.2 present linkage rates for immigrants and natives for each sample. Table B.1 divides the statistics by immigrants and natives and are greater for natives than immigrants in each sample and, for the most part, increase over time. Table B.2 Divides the linkage statistics for immigrants according to country of origin. Figure B.1 compares the observable characteristics of the (unweighted) linked sample to those of the full sample at risk for linkage. These differences between the linked and full sample form the basis of the reweighting procedure discussed above.

Table B.1: Linkage statistics

<i>Link</i>	Natives		Immigrants	
	(1) Start	(2) Linked	(3) Start	(4) Linked
1850–1880	1,401,603	159,162 (0.114)	407,168	23,670 (0.058)
1870–1900	2,604,671	243,074 (0.093)	656,172	42,754 (0.065)
1880–1910	3,500,920	370,280 (0.106)	688,990	49,988 (0.073)
1900–1930	5,708,114	686,072 (0.120)	1,172,600	99,589 (0.085)
1910–1940	5,751,119	921,471 (0.160)	1,964,666	133,301 (0.068)

Notes: Numbers in parentheses indicate the fraction of successful links. For 1850-1880 and 1910-1940, the numbers are relative to the initial year and the sample is limited to non-southern white men aged 18-30 in the base year. For 1870-1900, 1880-1910, and 1900-1930, numbers are relative to the final year and the sample is limited to non-southern white men aged 44-64 in the final year excluding unlinked immigrants arriving after the initial year; in this case, links to men not aged 18-30 in the initial year are not counted as links.

Table B.2: Linkage statistics by nativity

		(1) Norway	(2) Sweden	(3) UK	(4) Ireland	(5) France	(6) Netherlands	(7) Switzerland	(8) Italy	(9) Austria	(10) Germany	(11) Poland	(12) Russia
1850–1880	Start	2,514	824	69,638	204,381	7,894	1,747	2,374	619	124	114,374	744	162
	Linked	164	54	6,794	6,467	772	178	249	15	15	8,816	34	5
	Link Share	(0.065)	(0.066)	(0.098)	(0.032)	(0.098)	(0.102)	(0.105)	(0.024)	(0.121)	(0.077)	(0.046)	(0.031)
1870–1900	Start	24,030	24,567	108,114	183,036	8,764	6,720	10,742	3,974	4,651	253,464	3,724	2,595
	Linked	987	966	10,705	6,749	1,149	765	1,314	228	378	17,861	171	108
	Link Share	(0.041)	(0.039)	(0.099)	(0.037)	(0.131)	(0.114)	(0.122)	(0.057)	(0.081)	(0.070)	(0.046)	(0.042)
1880–1910	Start	37,109	46,510	114,965	126,011	8,903	7,732	11,757	13,948	22,679	254,675	1,097	14,690
	Linked	2,426	2,002	11,639	5,640	1,188	1,086	1,519	886	819	20,042	115	897
	Link Share	(0.065)	(0.043)	(0.101)	(0.045)	(0.133)	(0.140)	(0.129)	(0.064)	(0.036)	(0.079)	(0.105)	(0.061)
1900–1930	Start	50,015	95,575	131,725	110,086	14,070	16,019	16,393	122,522	37,093	244,119	84,593	85,349
	Linked	4,447	4,694	15,442	6,122	1,266	2,578	2,494	9,128	3,962	26,253	6,333	7,043
	Link Share	(0.089)	(0.049)	(0.117)	(0.056)	(0.090)	(0.161)	(0.152)	(0.075)	(0.107)	(0.108)	(0.075)	(0.083)
1910–1940	Start	63,793	91,116	125,374	98,851	13,283	16,090	12,069	375,421	303,181	180,948	7,307	381,893
	Linked	5,552	5,320	16,495	5,984	1,606	3,148	2,252	24,784	11,563	20,077	770	18,136
	Link Share	(0.087)	(0.058)	(0.132)	(0.061)	(0.121)	(0.196)	(0.187)	(0.066)	(0.038)	(0.111)	(0.105)	(0.047)

Notes: Sample restrictions and definitions are the same as in Table B.1.

Figure B.1: Selection into linkage, natives

Figure B.1(a): 1850-1880

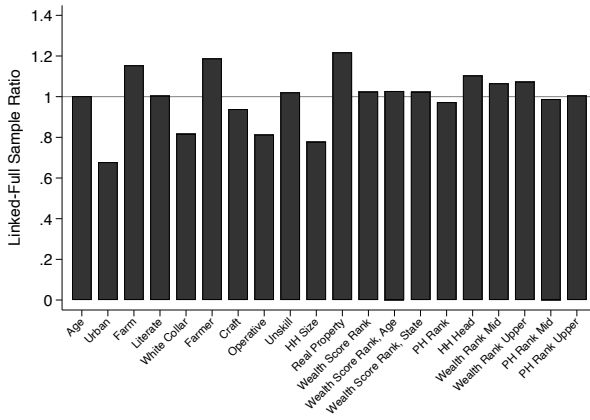


Figure B.1(b): 1870-1900

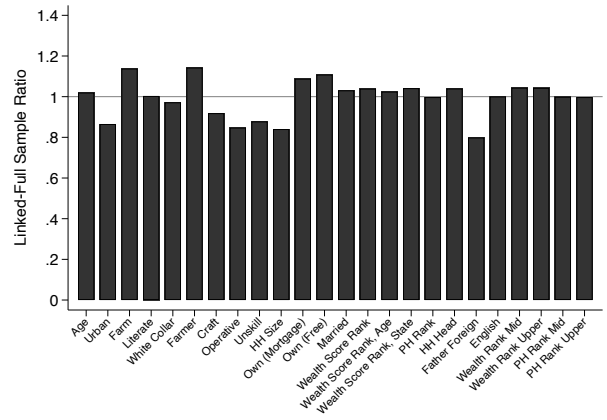


Figure B.1(c): 1880-1910

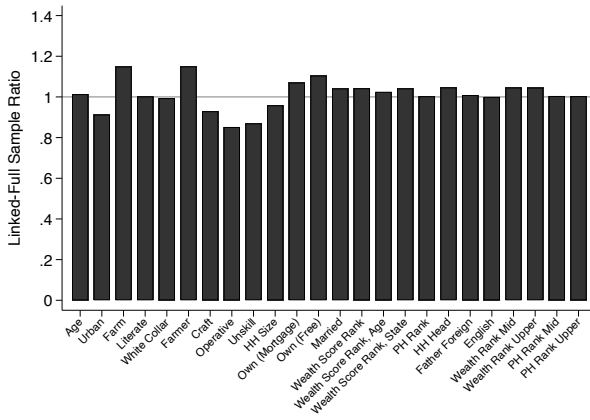


Figure B.1(d): 1900-1930

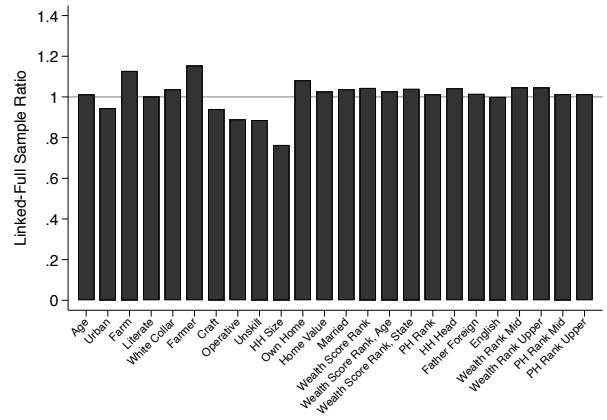
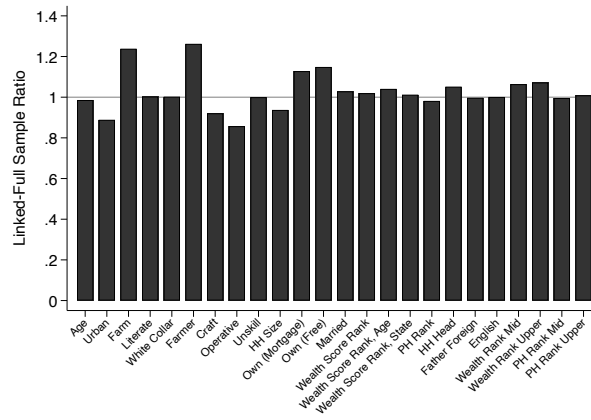


Figure B.1(e): 1910-1940



Notes: Each bar shows the ratio of the value of a variable in the linked sample to the full sample at risk for linkage. A value of 1 indicates that the values of the variable in the linked and full sample are the same—perfect representativeness.

Figure B.2: Selection into linkage, immigrants

Figure B.2(a): 1850-1880

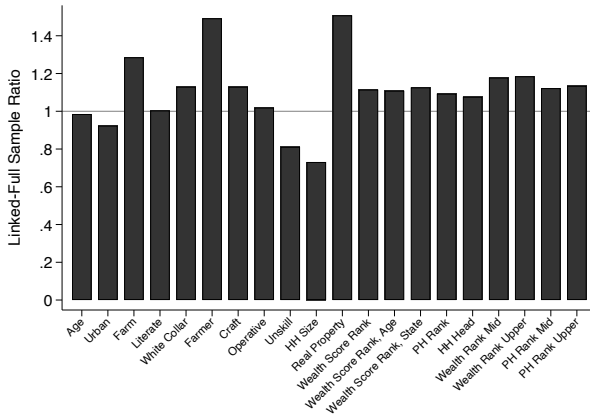


Figure B.2(b): 1870-1900

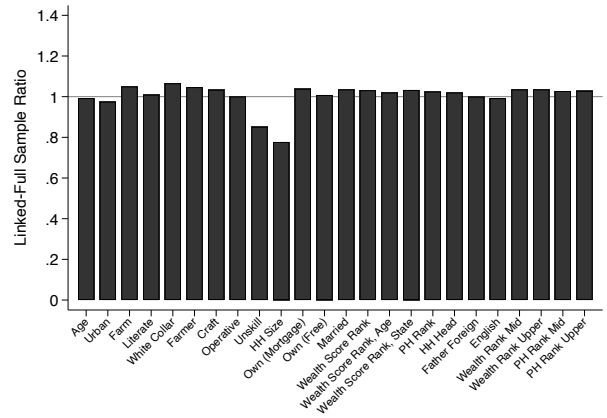


Figure B.2(c): 1880-1910

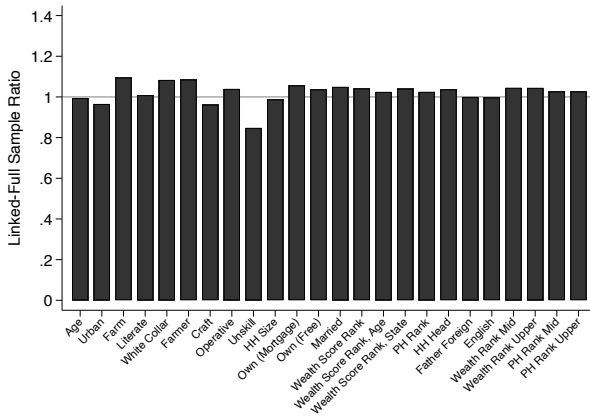


Figure B.2(d): 1900-1930

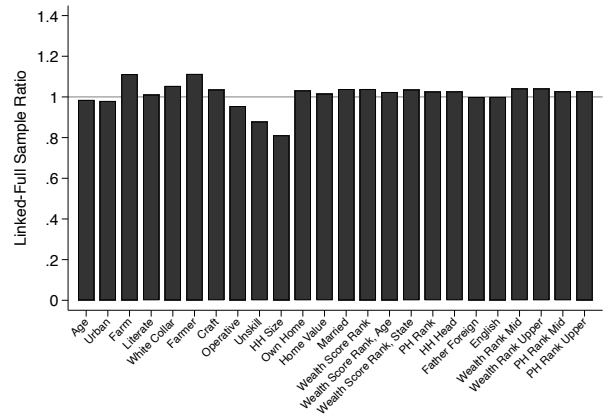
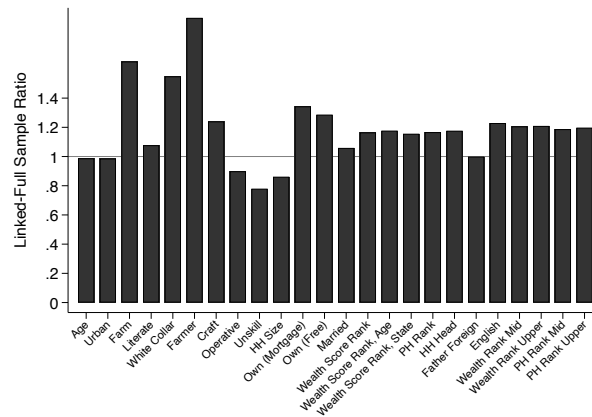


Figure B.2(e): 1910-1940



Notes: Each bar shows the ratio of the value of a variable in the linked sample to the full sample at risk for linkage. A value of 1 indicates that the values of the variable in the linked and full sample are the same—perfect representativeness.

Appendix C: Results with alternate linkage methods

In this Appendix, we repeat the main results with four alternative linkage methods. The first two are the ABE-Exact and ABE-NYSIIS linkage methods, described and provided by Abramitzky et al. (2020). The third is what Zimran (2021) refers to as the “Intersection-of-Matches” method, which accepts only matches made by the method used in the main text and by the two ABE methods. In effect, this limits the sample to cases in which there is an exact match of first and last names and birthplaces and an age-implied birthyear difference of no more than two years, but requires that individuals be unique not just on their exact names, but also that their NYSIIS names be unique and that the names be unique within a band of names defined by the spelling distance cutoffs. The last method is founded on the intersection-of-matches method but keeps only links that can be “corroborated” by observable characteristics not used in the matching. For instance, we require that, if both censuses report parent’s birthplace, the reports match across censuses. The latter two methods are considerably stricter than the two ABE methods and the method used in the main text, which leads to the potential for fewer false matches (Abramitzky et al. 2021a) but also lead to a potentially more selected sample. This selection is likely to be more positive, since it depends on the accurate and consistent report of information 30 years apart, and the somewhat greater assimilation performance of immigrants in these samples is consistent with that suspicion.

The results when using the two ABE methods are very similar to those of the main text. There are slight differences when using the two stricter methods, which could be the result of smaller samples or of greater selectivity, but the main results of a U-shaped assimilation pattern driven by the changing initial occupation distributions are robust.

Figure C.1: Occupational distributions

Figure C.1(a): Initial Year 1850

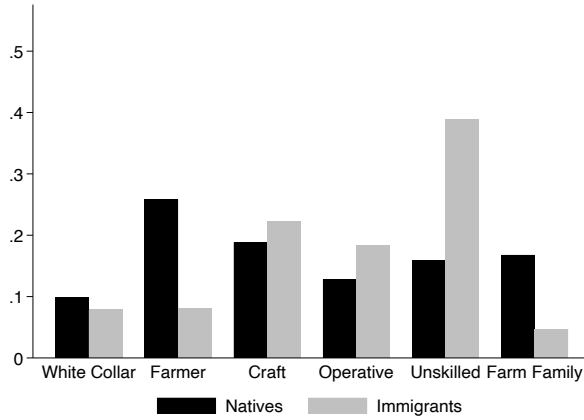


Figure C.1(b): Final Year 1880

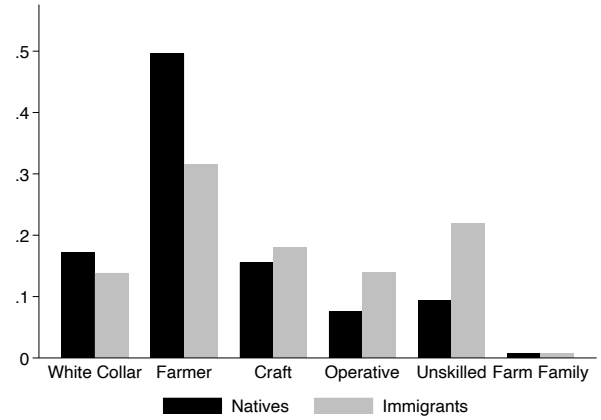


Figure C.1(c): Initial Year 1870

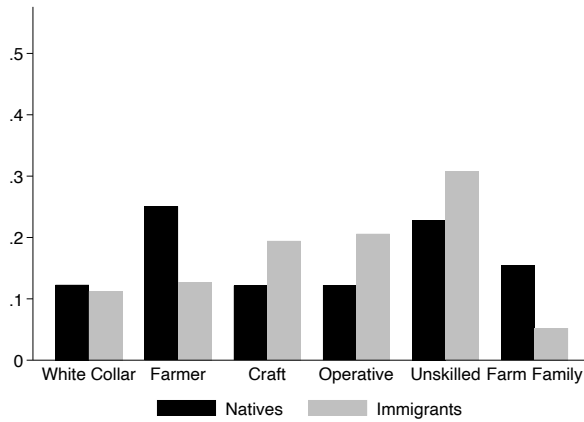


Figure C.1(d): Final Year 1900

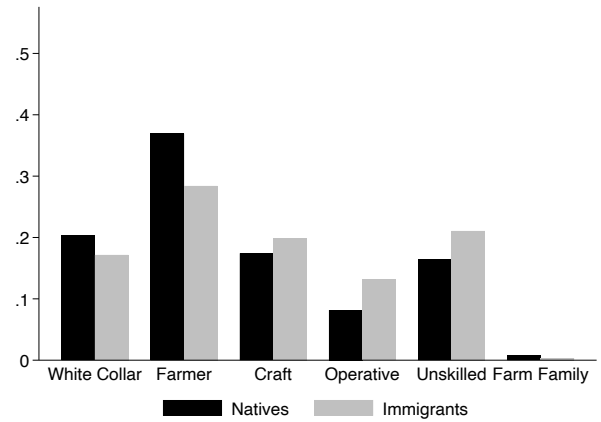


Figure C.1(e): Initial Year 1880

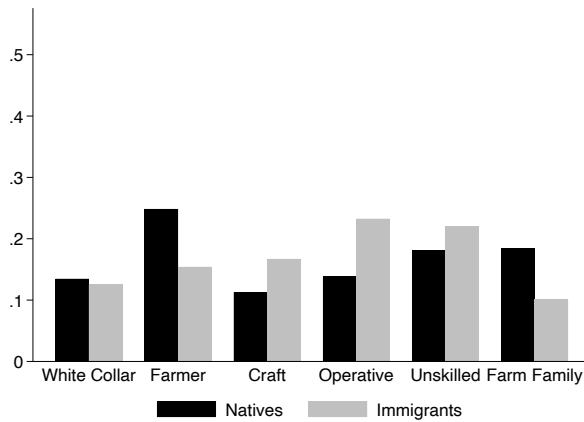


Figure C.1(f): Final Year 1910

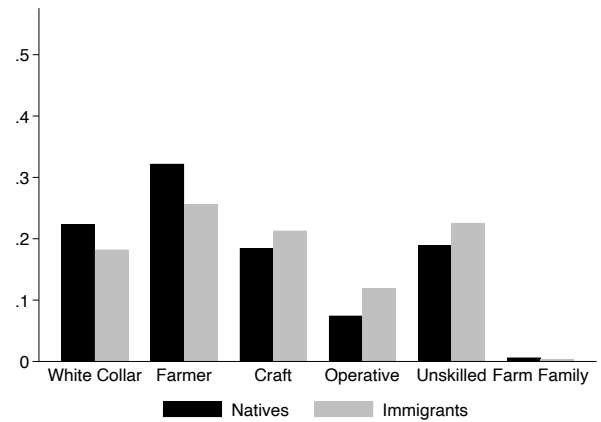


Figure C.1 (continued): Occupational distributions

Figure C.1(g): Initial Year 1900

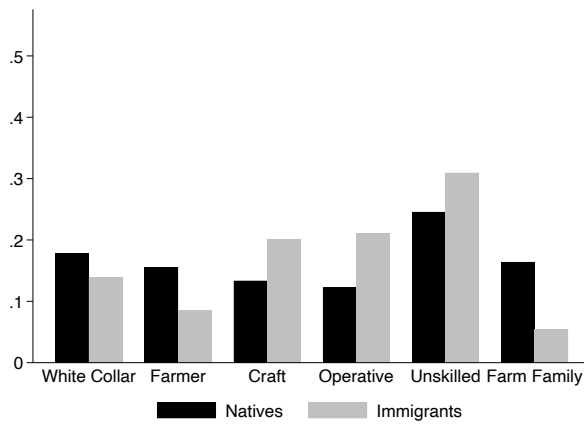


Figure C.1(h): Final Year 1930

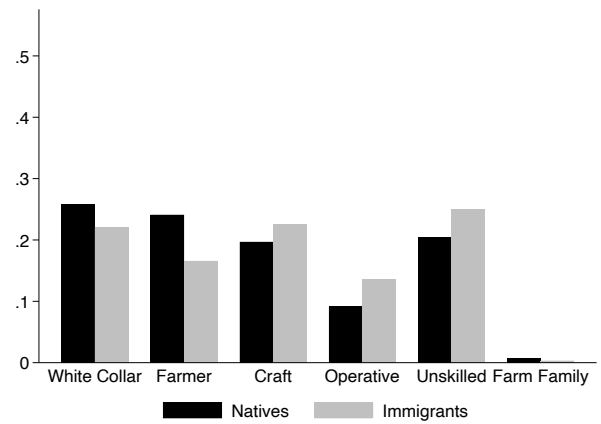


Figure C.1(i): Initial Year 1910

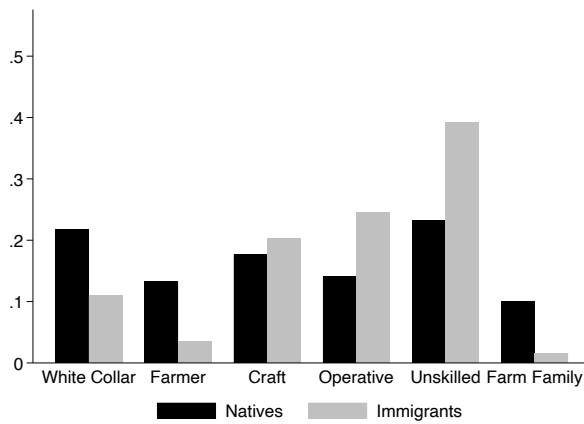
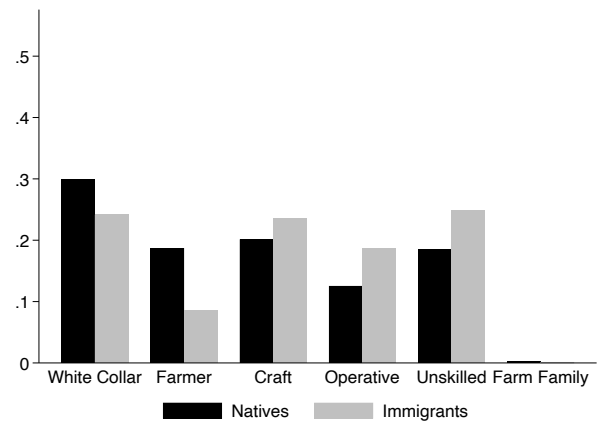


Figure C.1(j): Final Year 1940



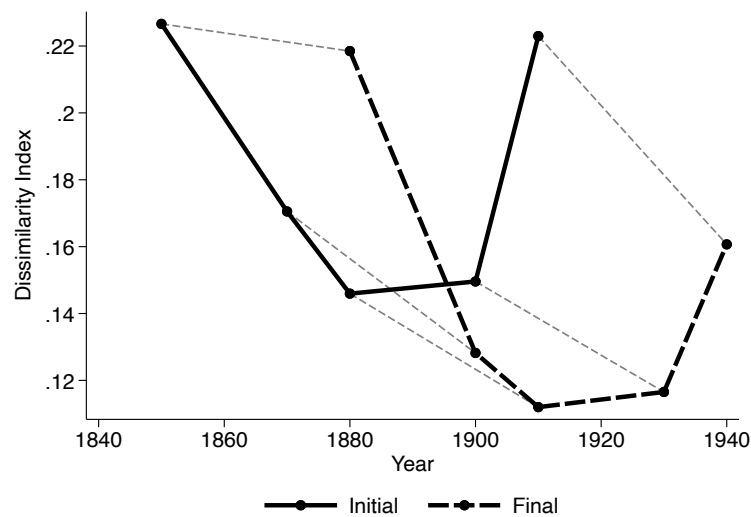
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure C.2: Initial gaps in occupational rank



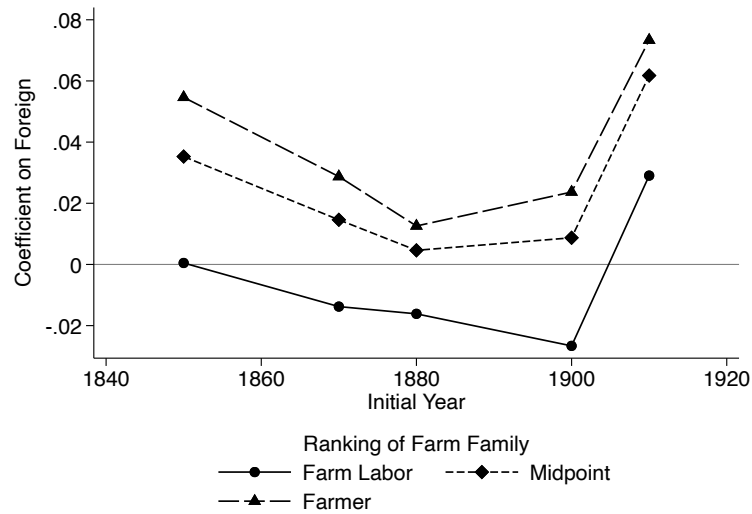
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure C.3: Dissimilarity indices



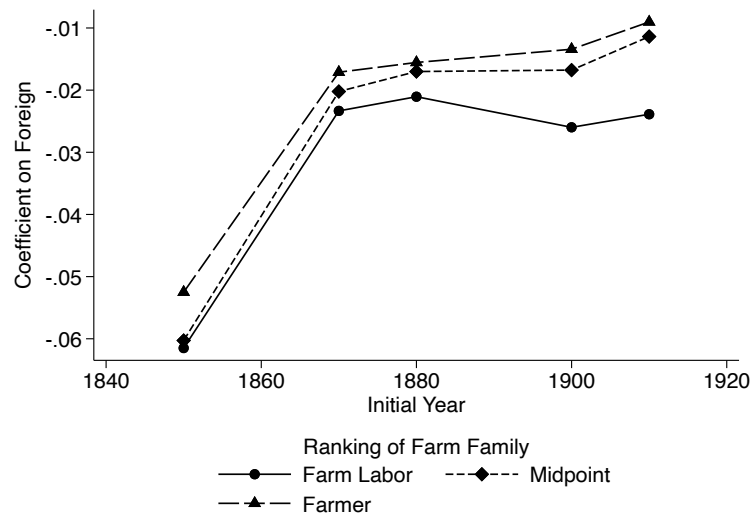
Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure C.4: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.5: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.6: Occupational upgrading by initial occupation

Figure C.6(a): 1850-1880

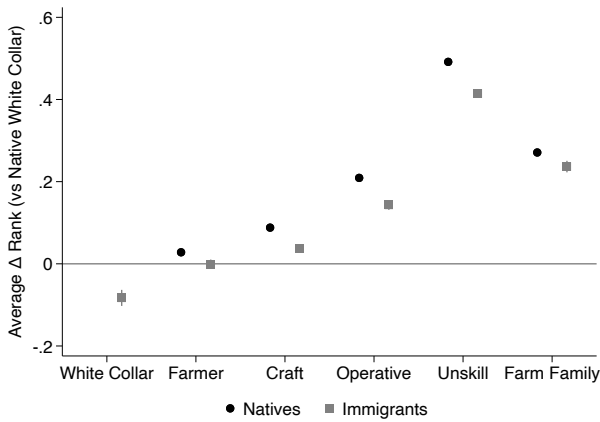


Figure C.6(b): 1870-1900

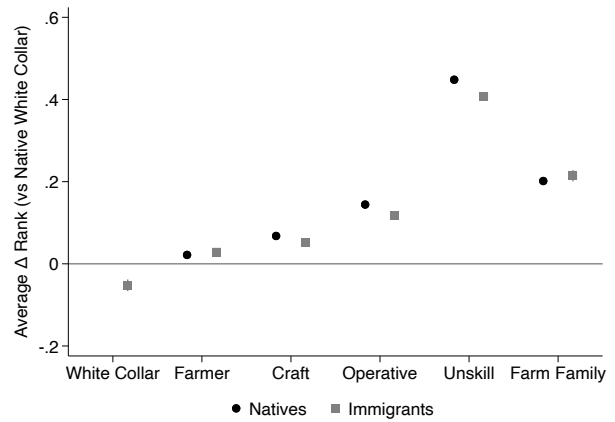


Figure C.6(c): 1880-1910

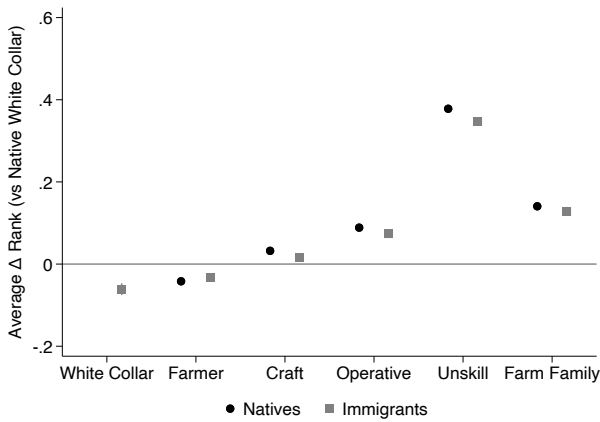


Figure C.6(d): 1900-1930

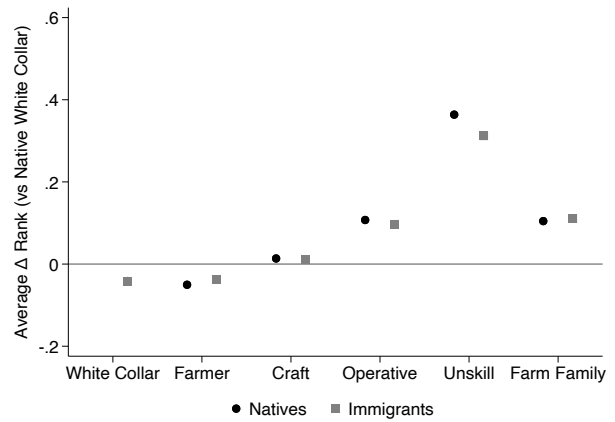
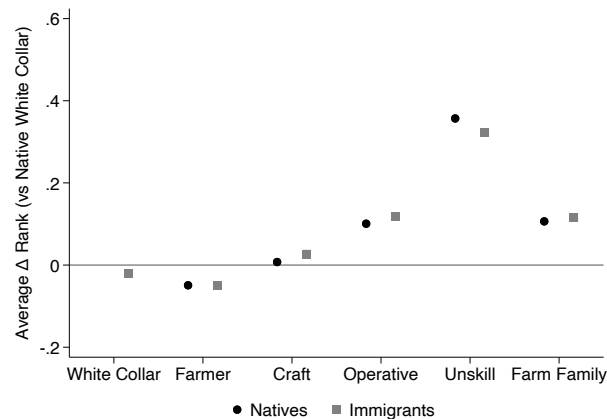
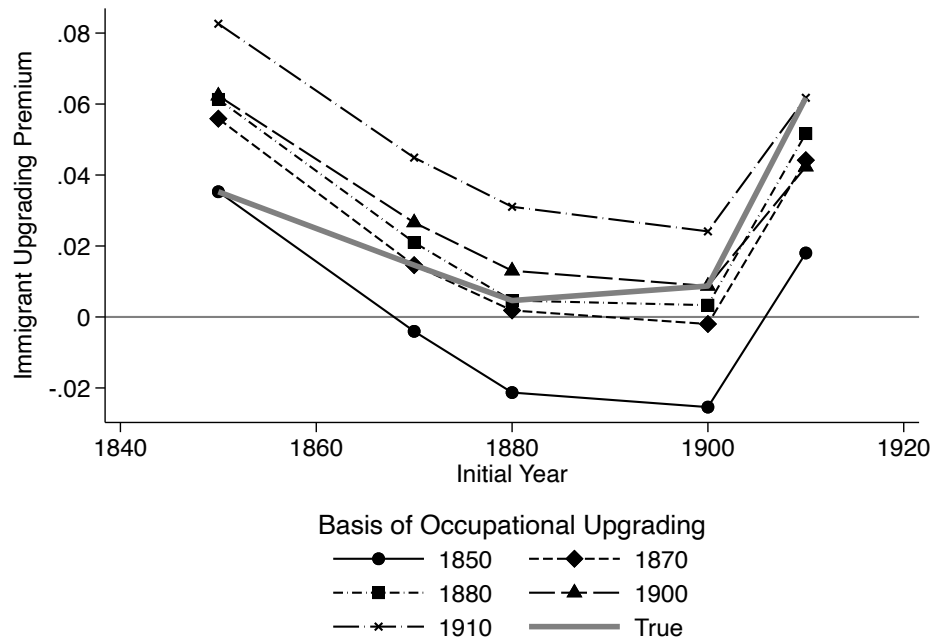


Figure C.6(e): 1910-1940



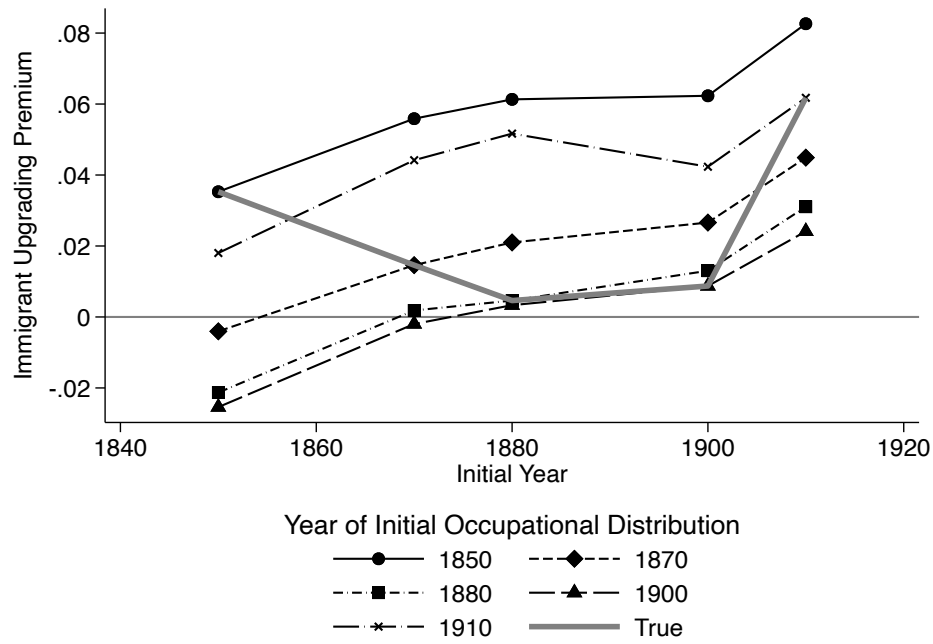
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.7: Unconditional assimilation, holding upgrading constant



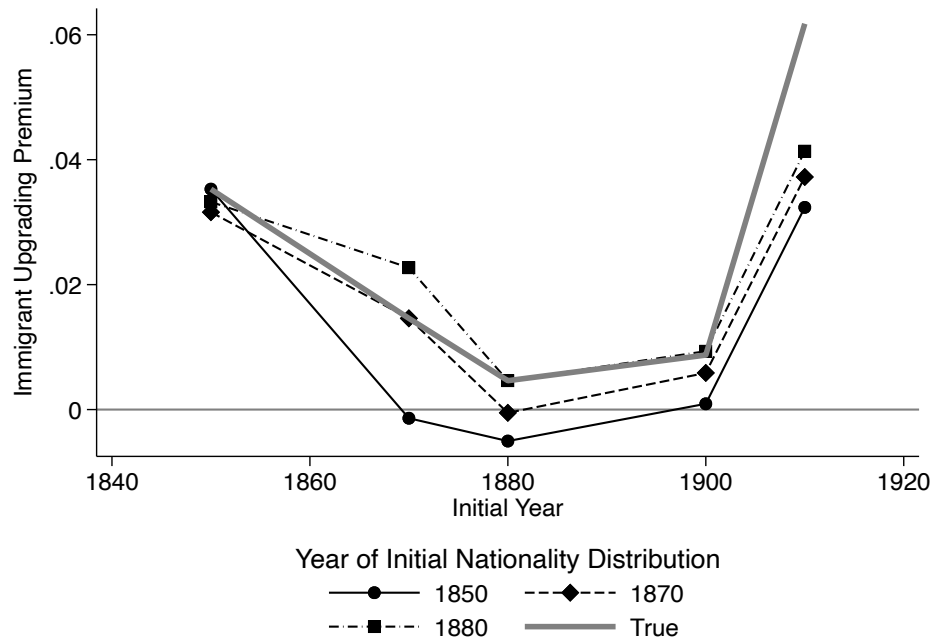
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.8: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.9: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.10: Occupational distributions

Figure C.10(a): Initial Year 1850

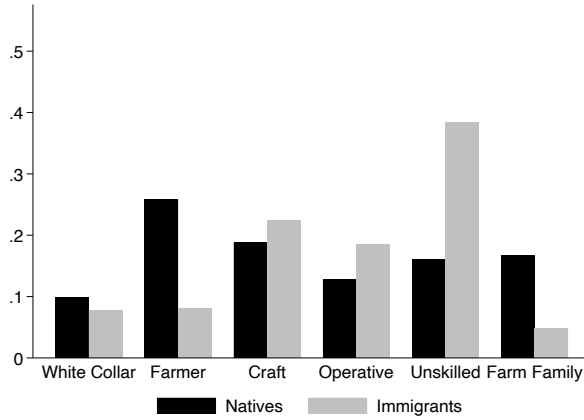


Figure C.10(b): Final Year 1880

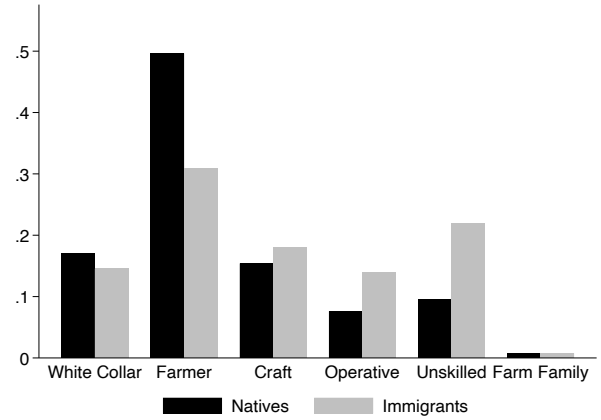


Figure C.10(c): Initial Year 1870

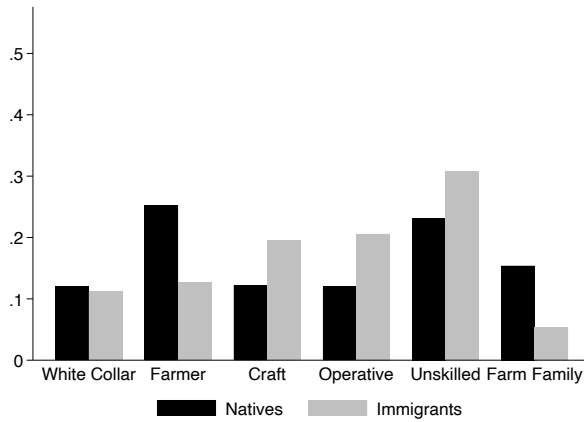


Figure C.10(d): Final Year 1900

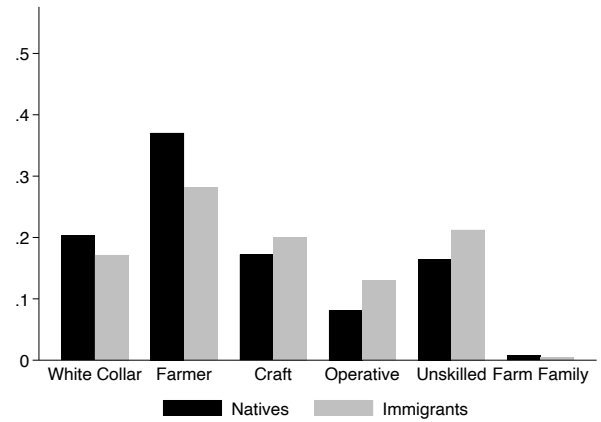


Figure C.10(e): Initial Year 1880

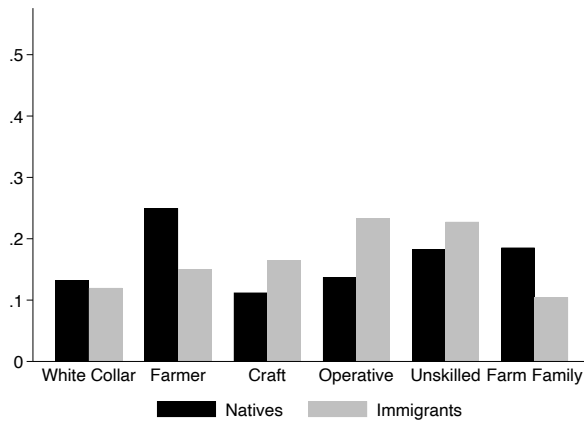


Figure C.10(f): Final Year 1910

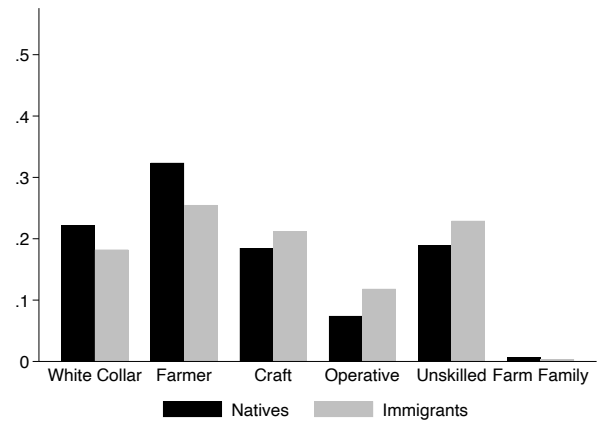


Figure C.10 (continued): Occupational distributions

Figure C.10(g): Initial Year 1900

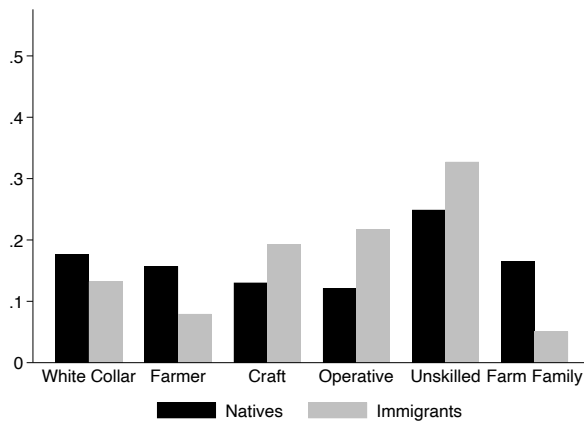


Figure C.10(h): Final Year 1930

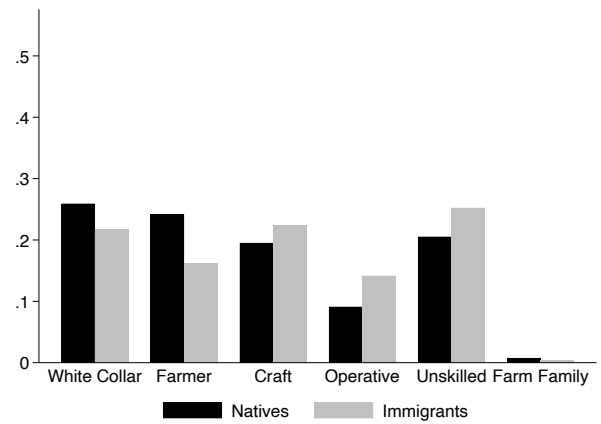


Figure C.10(i): Initial Year 1910

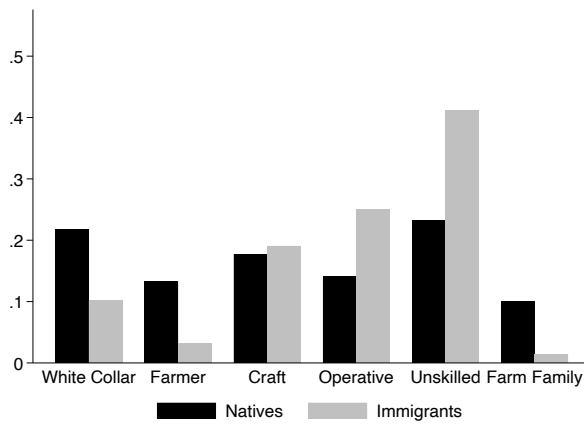
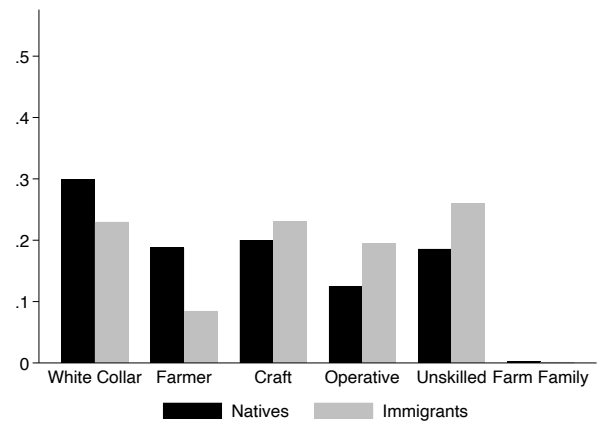
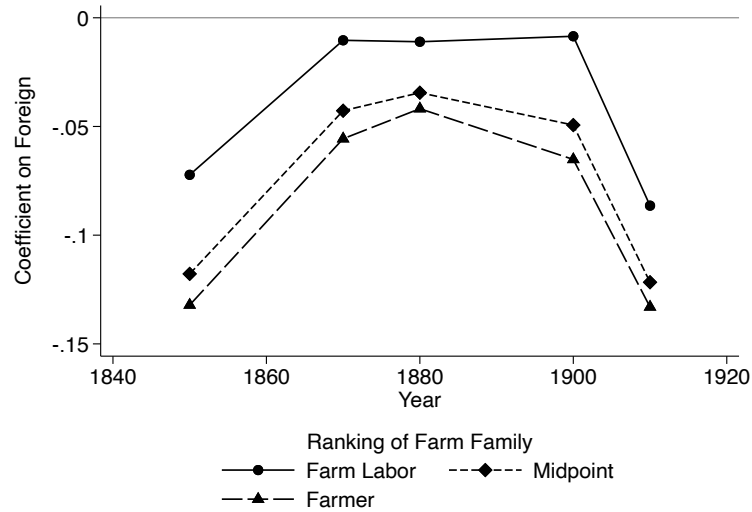


Figure C.10(j): Final Year 1940



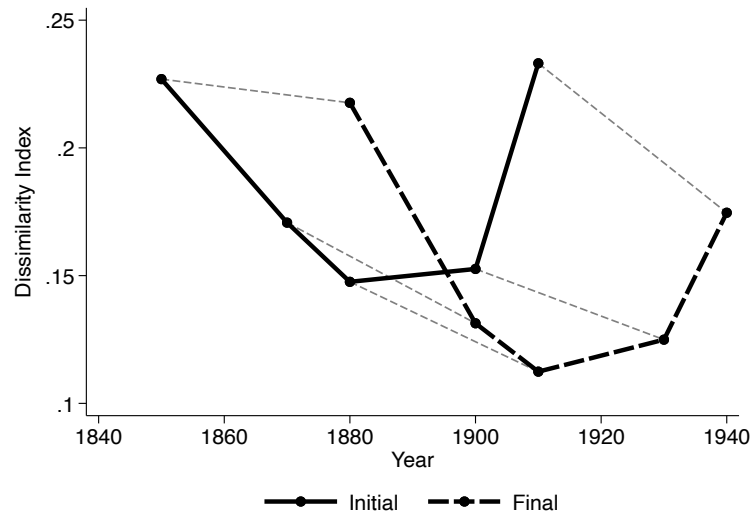
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure C.11: Initial gaps in occupational rank



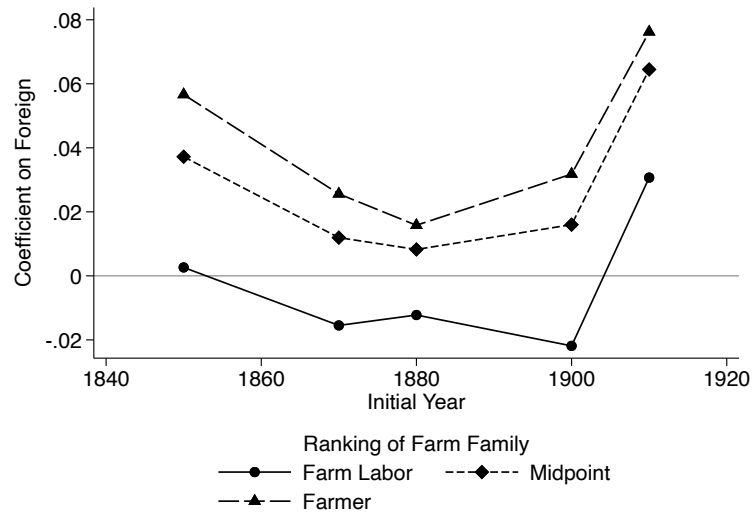
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure C.12: Dissimilarity indices



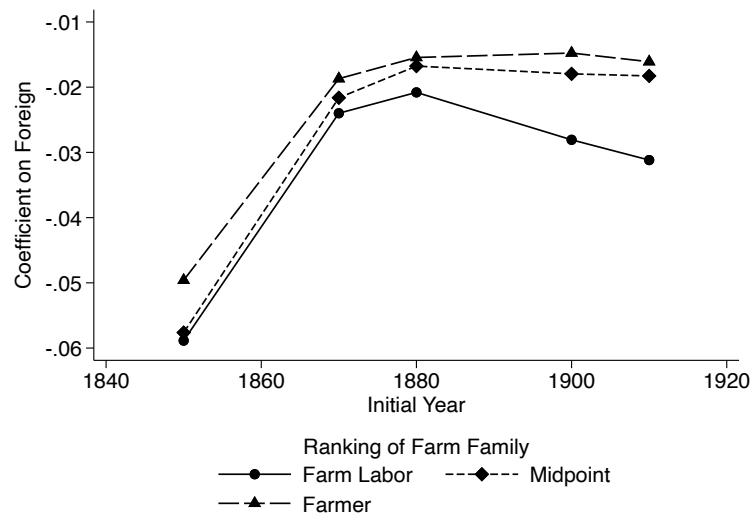
Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure C.13: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.14: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.15: Occupational upgrading by initial occupation

Figure C.15(a): 1850-1880

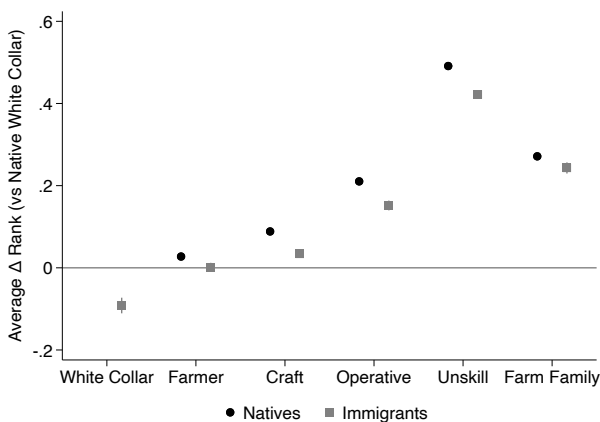


Figure C.15(b): 1870-1900

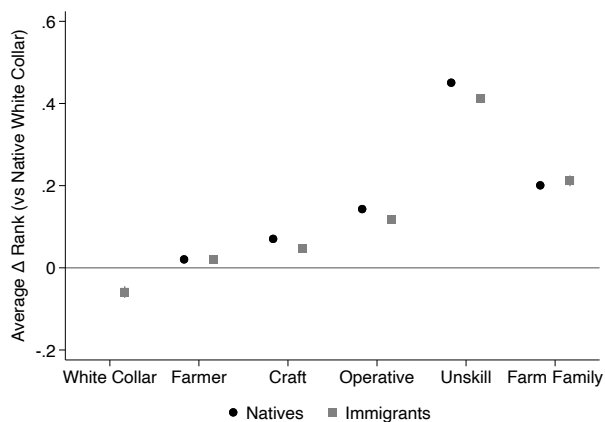


Figure C.15(c): 1880-1910

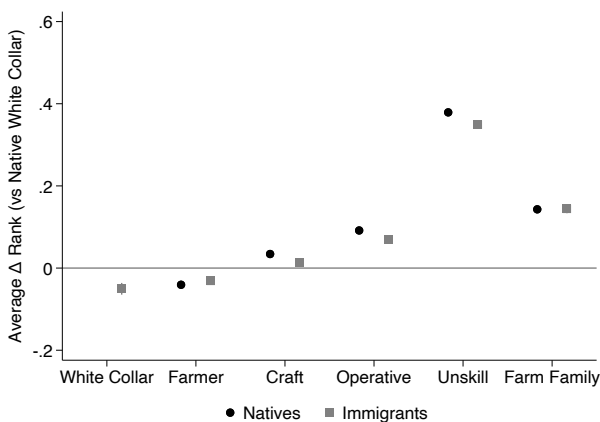


Figure C.15(d): 1900-1930

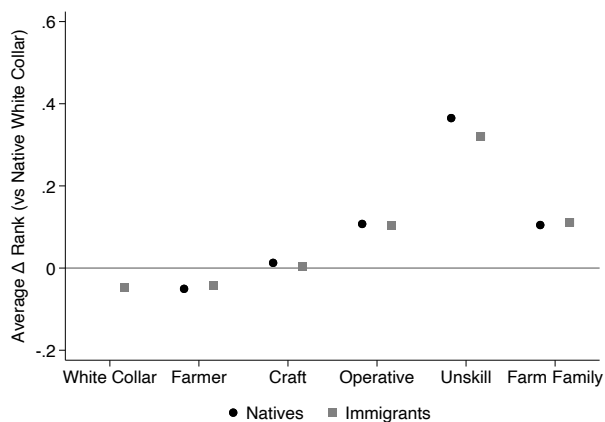
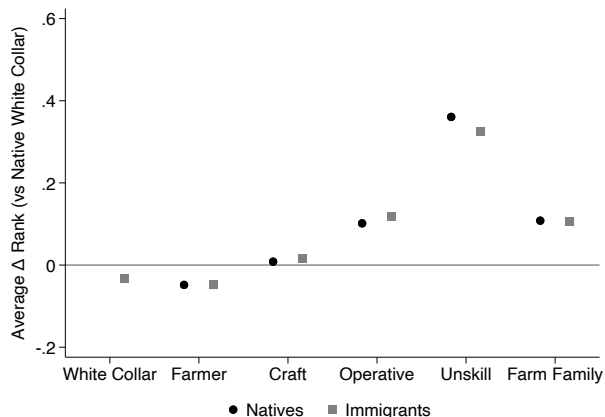
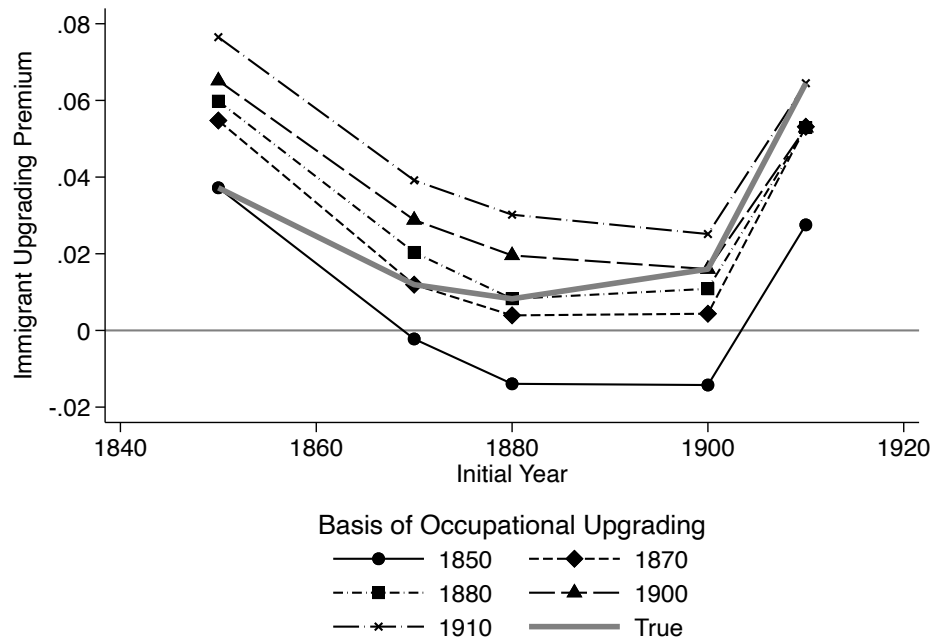


Figure C.15(e): 1910-1940



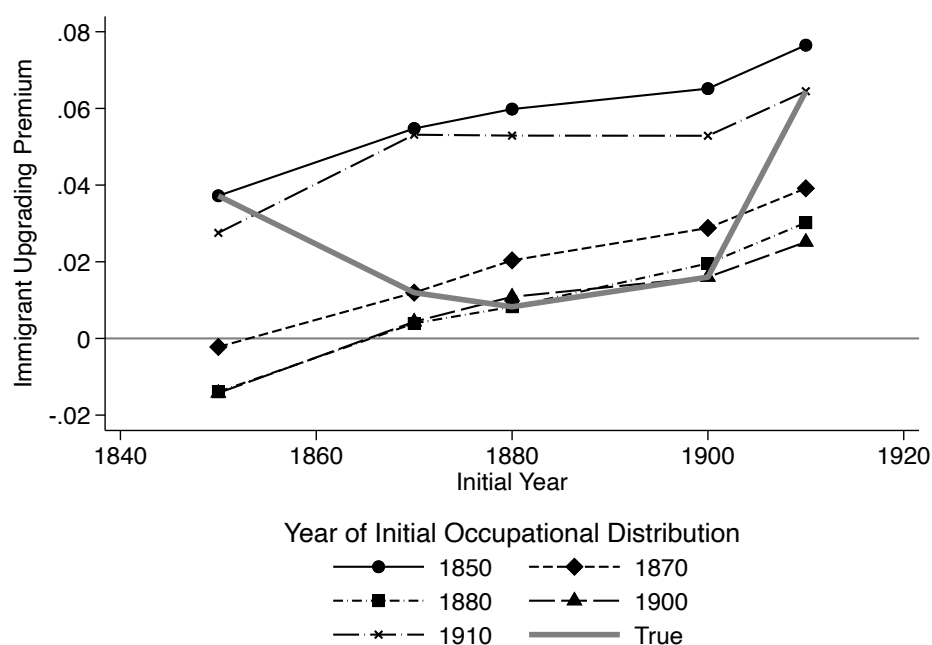
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.16: Unconditional assimilation, holding upgrading constant



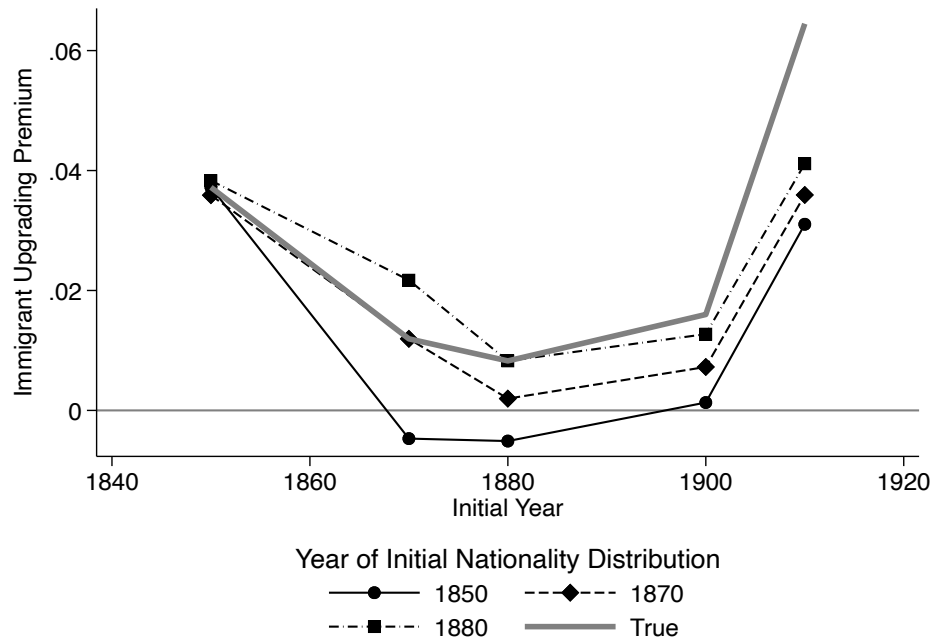
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.17: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.18: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.19: Occupational distributions

Figure C.19(a): Initial Year 1850

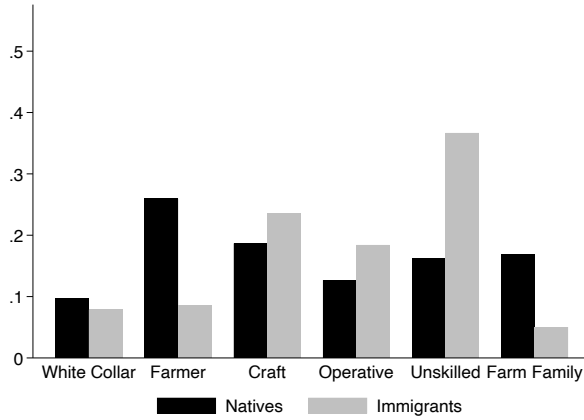


Figure C.19(b): Final Year 1880

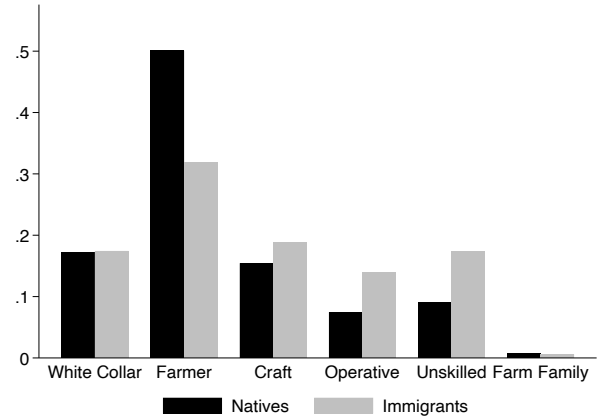


Figure C.19(c): Initial Year 1870

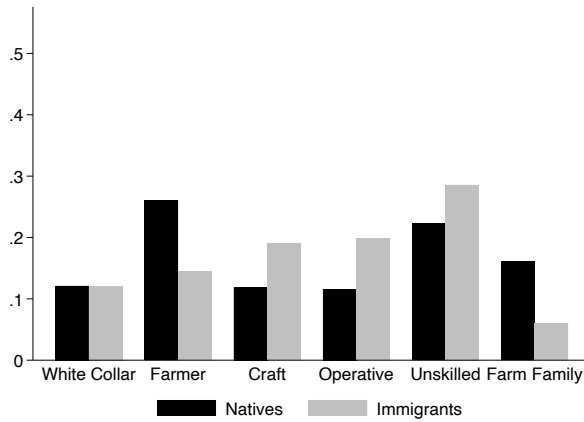


Figure C.19(d): Final Year 1900

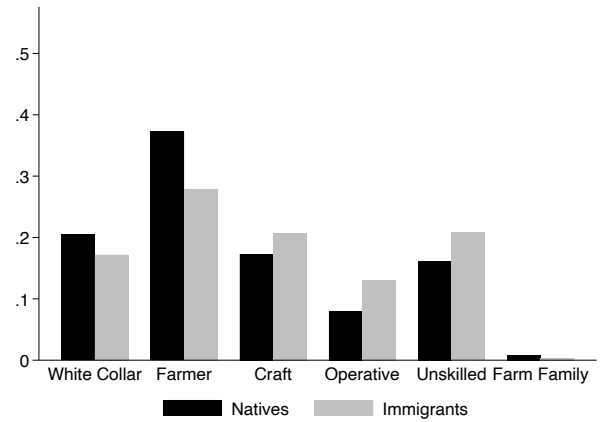


Figure C.19(e): Initial Year 1880

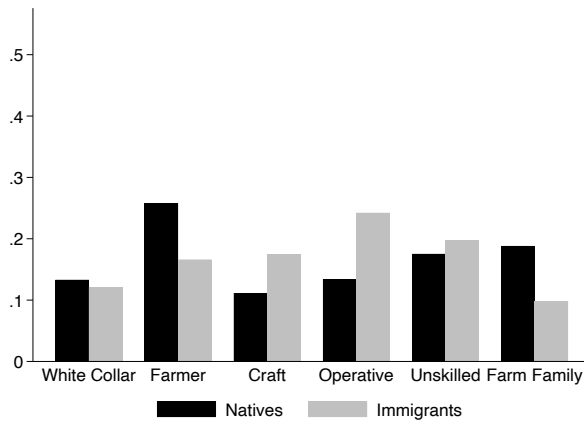


Figure C.19(f): Final Year 1910

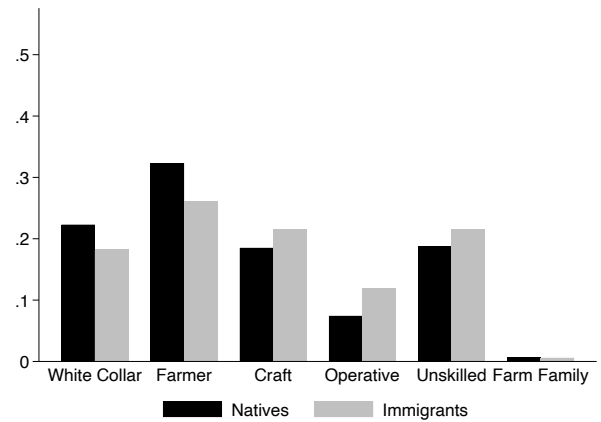


Figure C.19 (continued): Occupational distributions

Figure C.19(g): Initial Year 1900

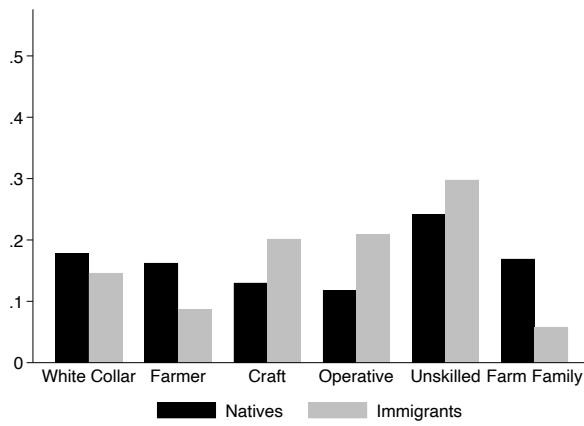


Figure C.19(h): Final Year 1930

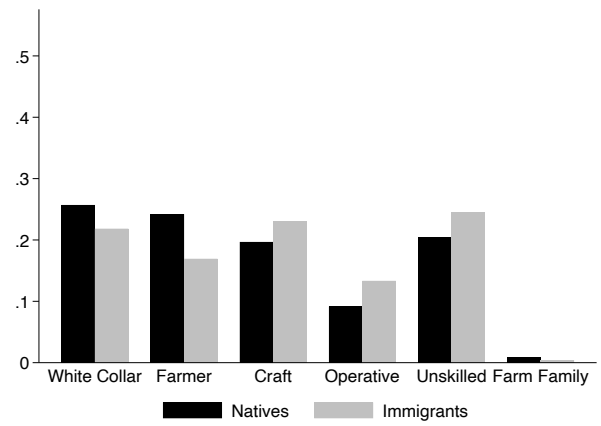


Figure C.19(i): Initial Year 1910

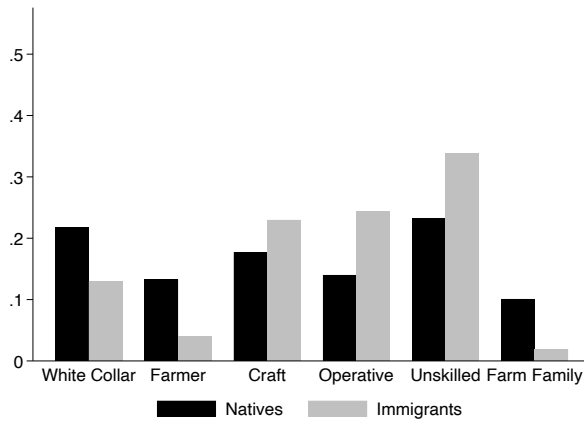
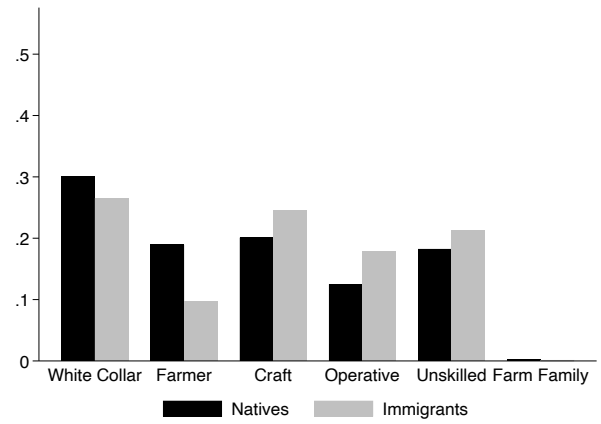


Figure C.19(j): Final Year 1940



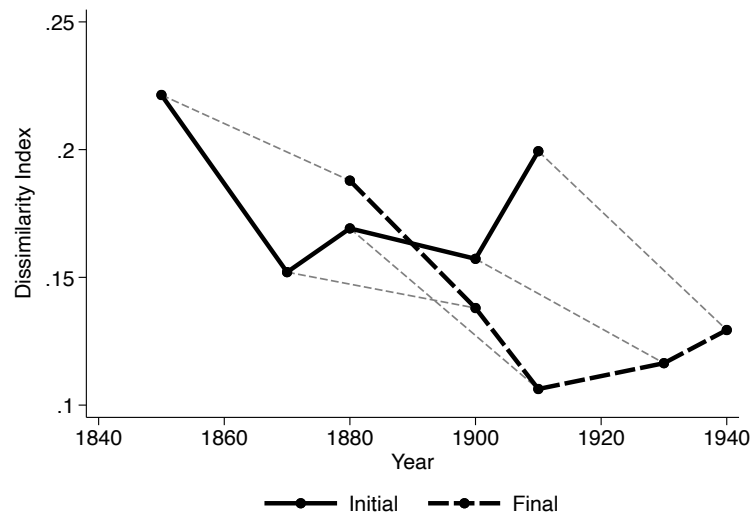
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure C.20: Initial gaps in occupational rank



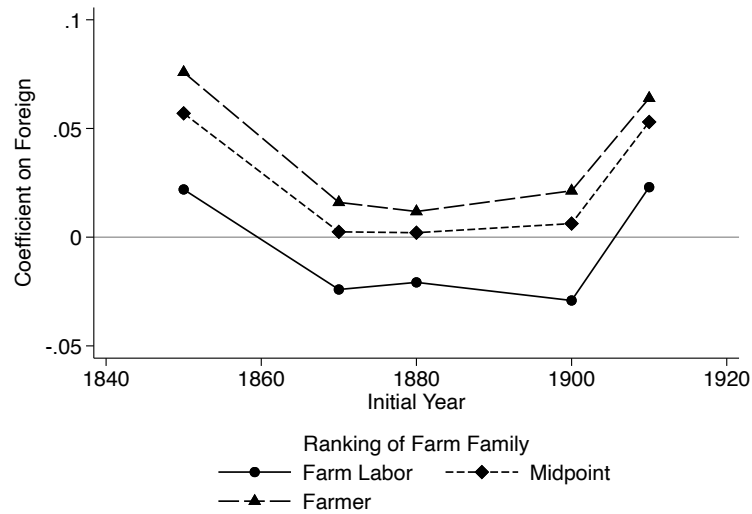
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure C.21: Dissimilarity indices



Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure C.22: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.23: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.24: Occupational upgrading by initial occupation

Figure C.24(a): 1850-1880

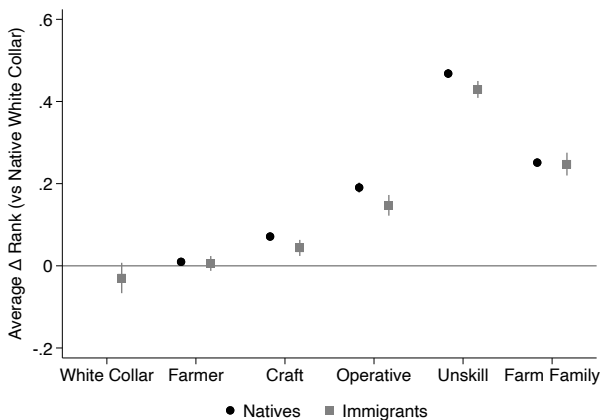


Figure C.24(b): 1870-1900

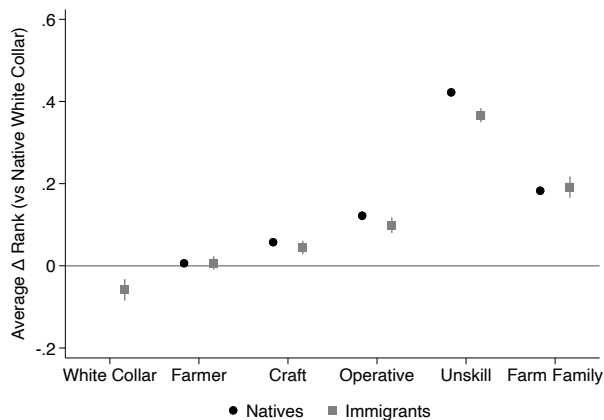


Figure C.24(c): 1880-1910

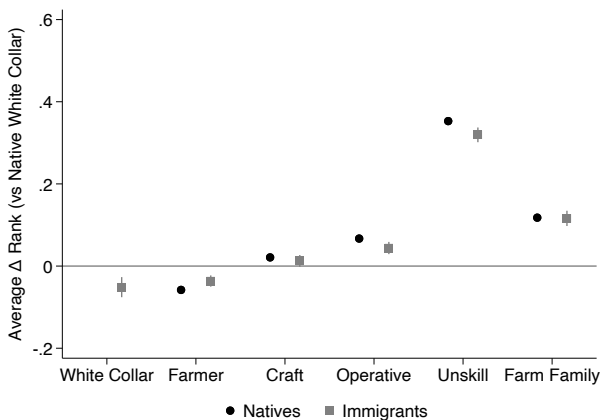


Figure C.24(d): 1900-1930

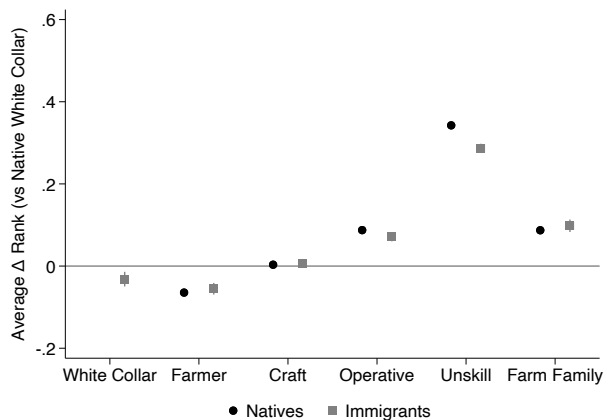
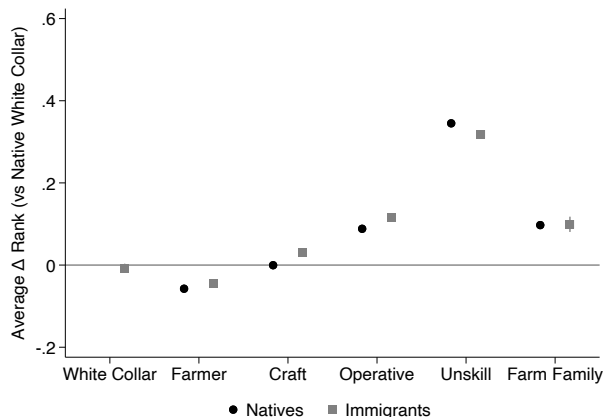
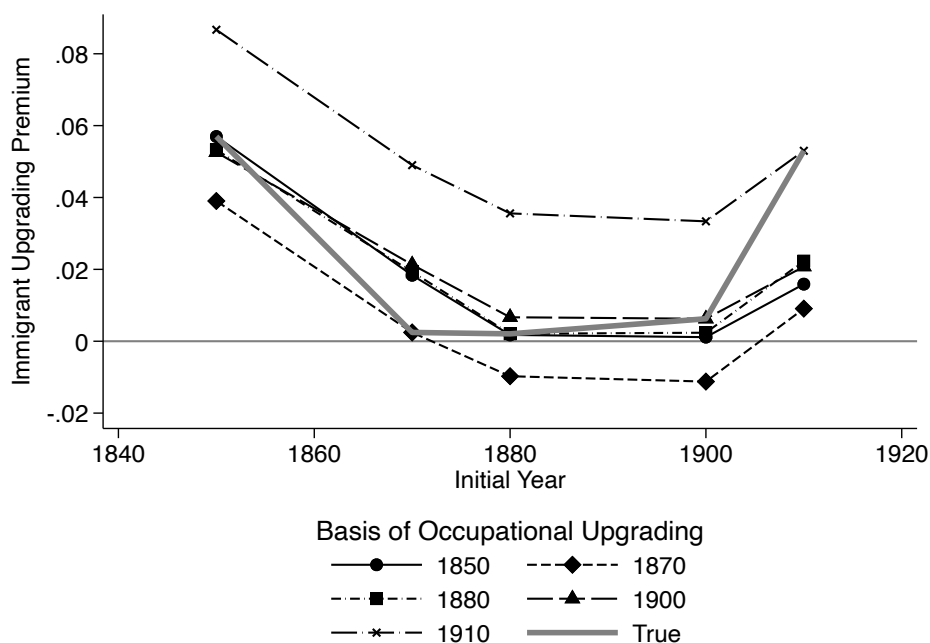


Figure C.24(e): 1910-1940



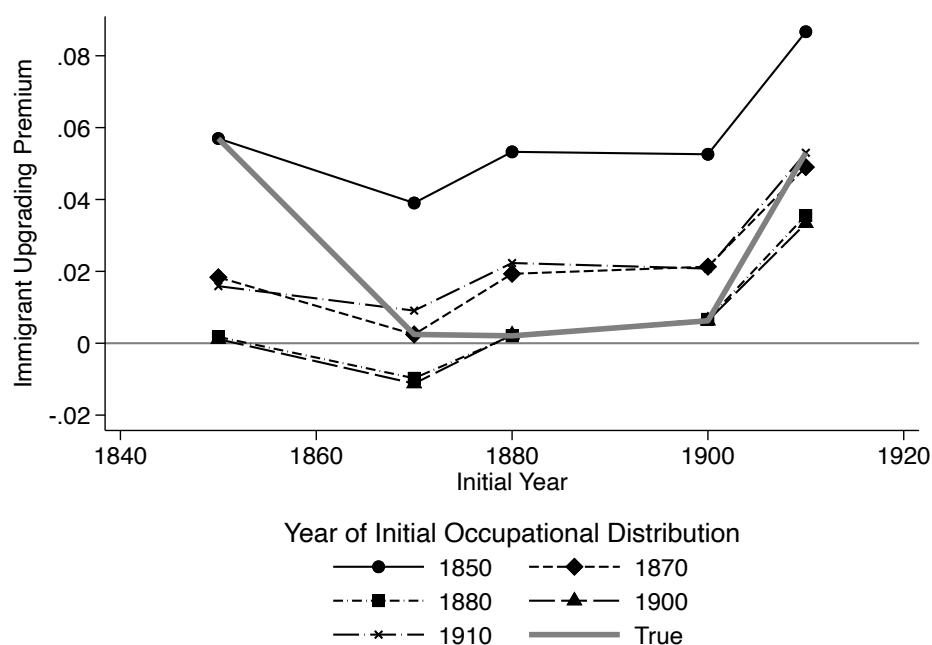
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.25: Unconditional assimilation, holding upgrading constant



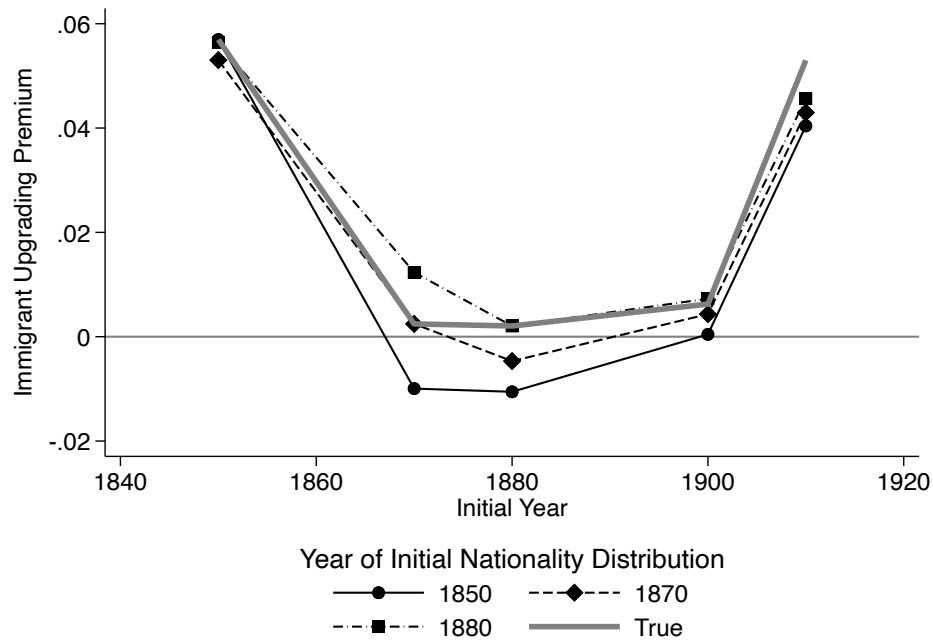
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.26: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.27: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.28: Occupational distributions

Figure C.28(a): Initial Year 1850

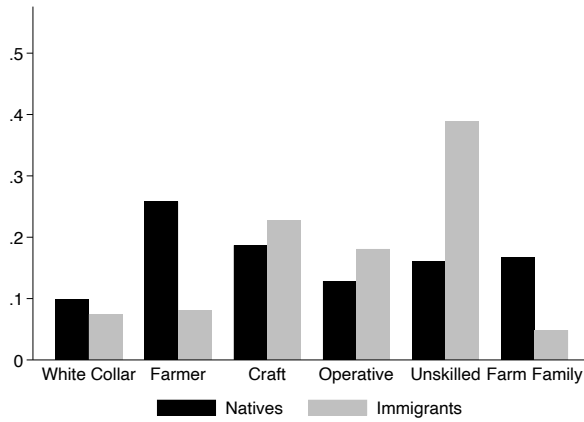


Figure C.28(b): Final Year 1880

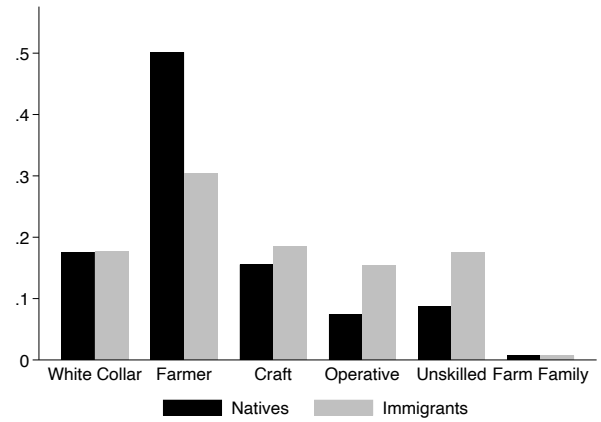


Figure C.28(c): Initial Year 1870

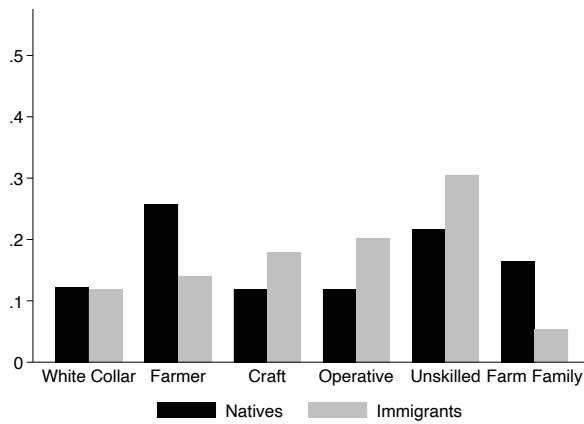


Figure C.28(d): Final Year 1900

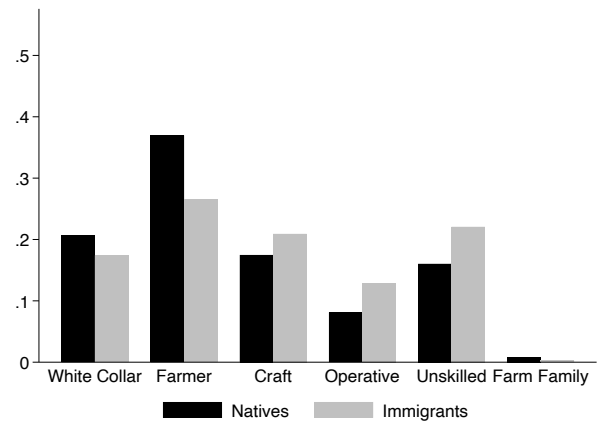


Figure C.28(e): Initial Year 1880

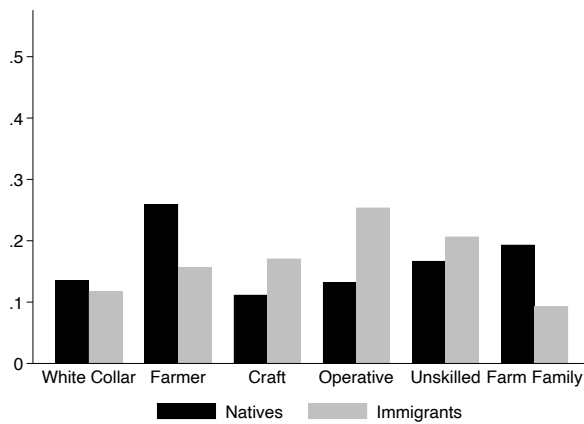


Figure C.28(f): Final Year 1910

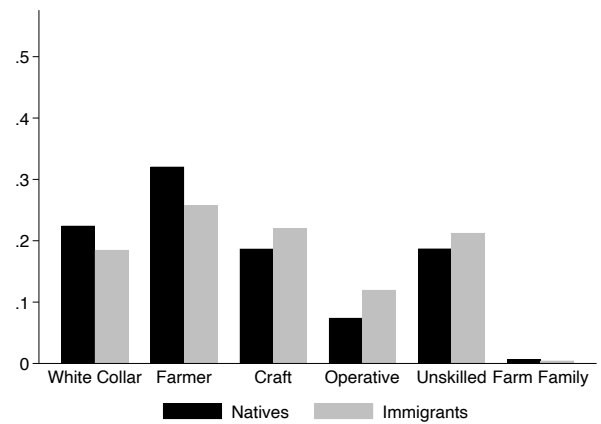


Figure C.28 (continued): Occupational distributions

Figure C.28(g): Initial Year 1900

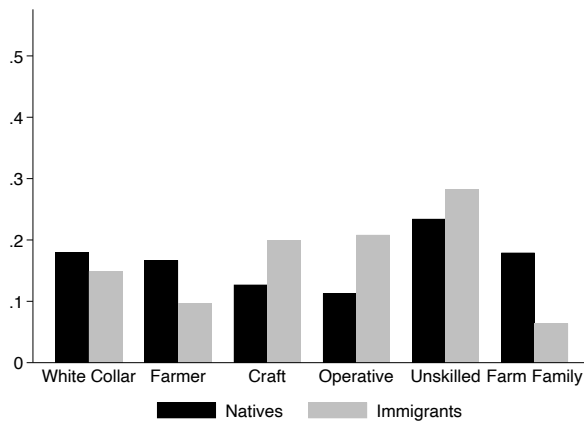


Figure C.28(h): Final Year 1930

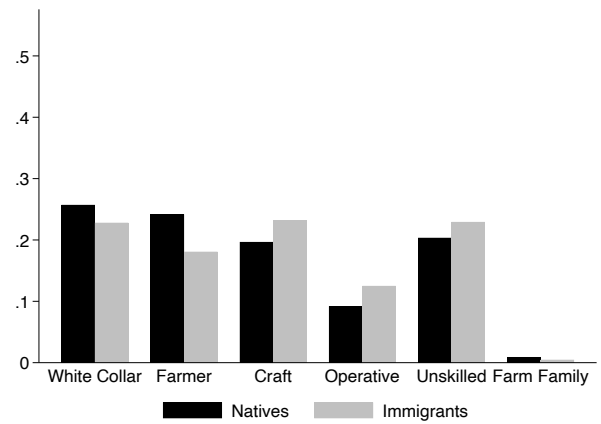


Figure C.28(i): Initial Year 1910

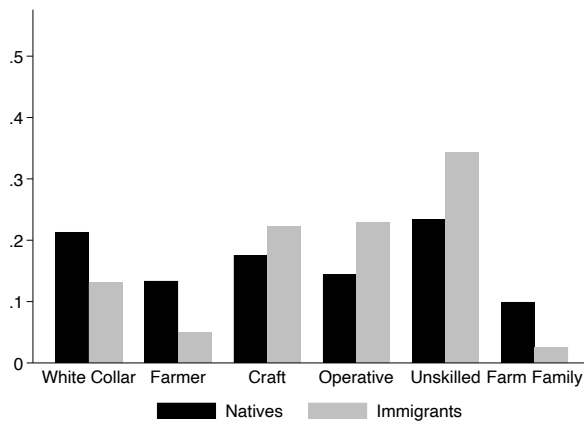
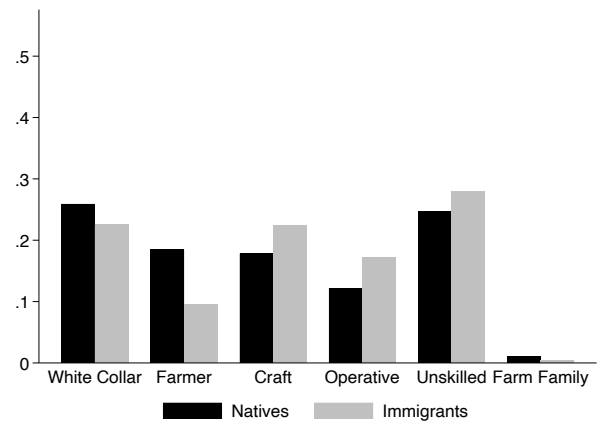
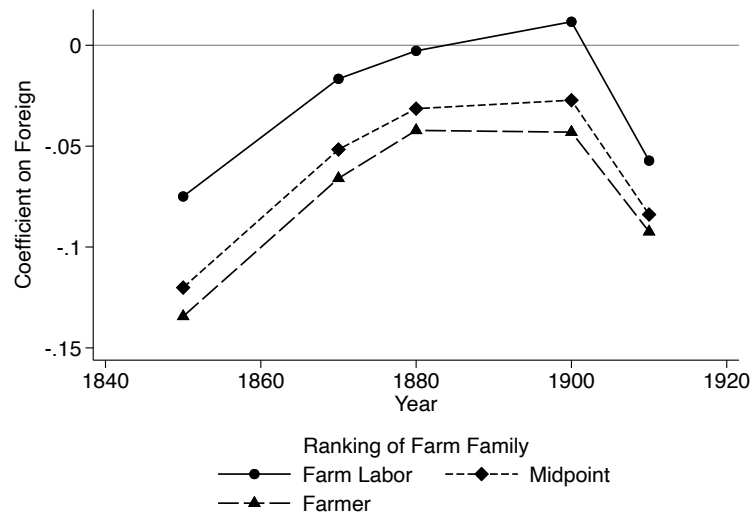


Figure C.28(j): Final Year 1940



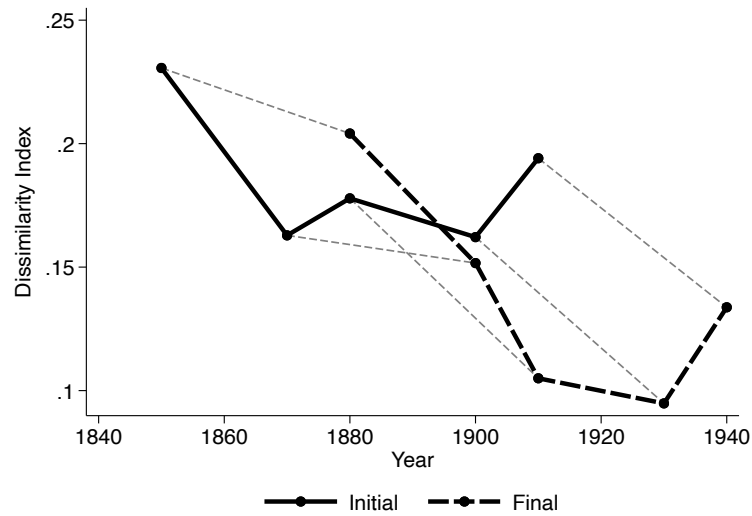
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure C.29: Initial gaps in occupational rank



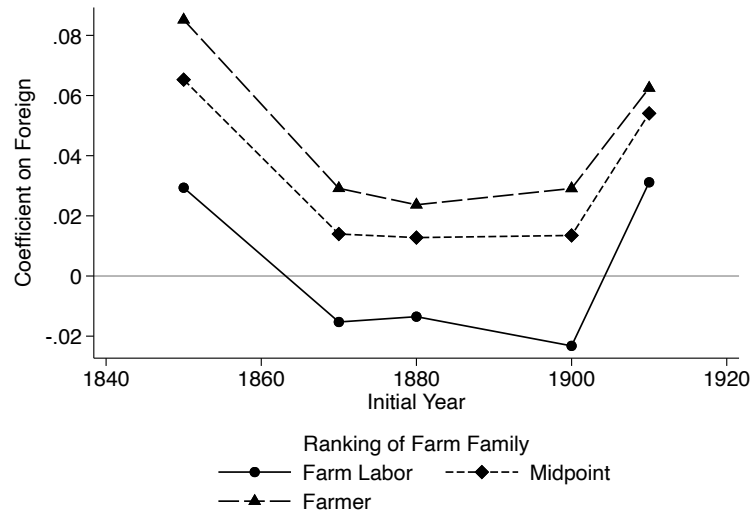
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure C.30: Dissimilarity indices



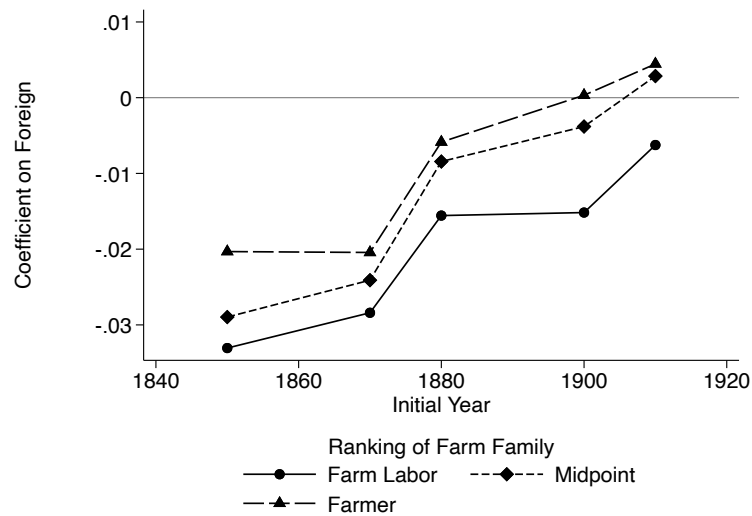
Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure C.31: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.32: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure C.33: Occupational upgrading by initial occupation

Figure C.33(a): 1850-1880

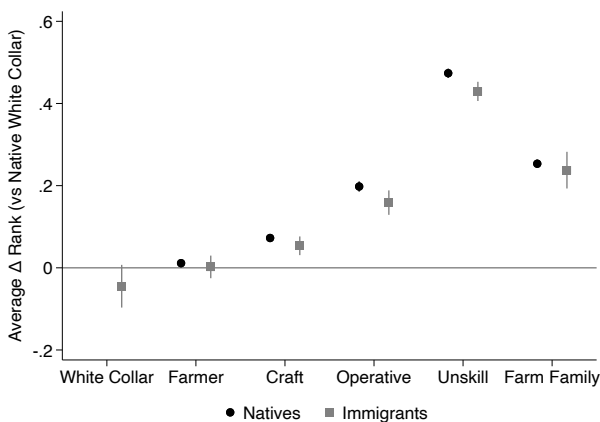


Figure C.33(b): 1870-1900

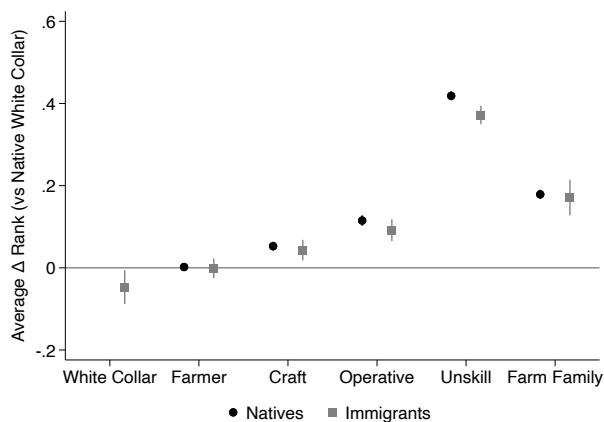


Figure C.33(c): 1880-1910

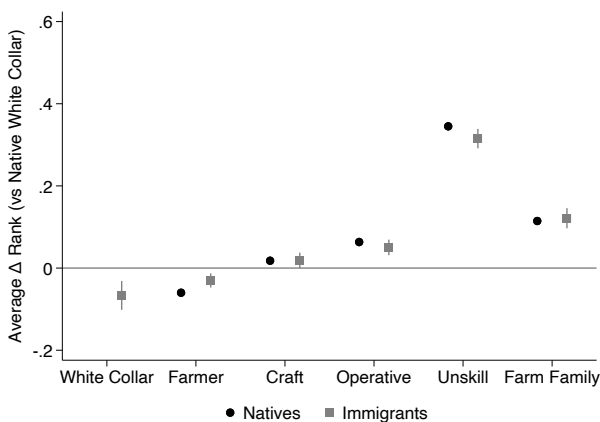


Figure C.33(d): 1900-1930

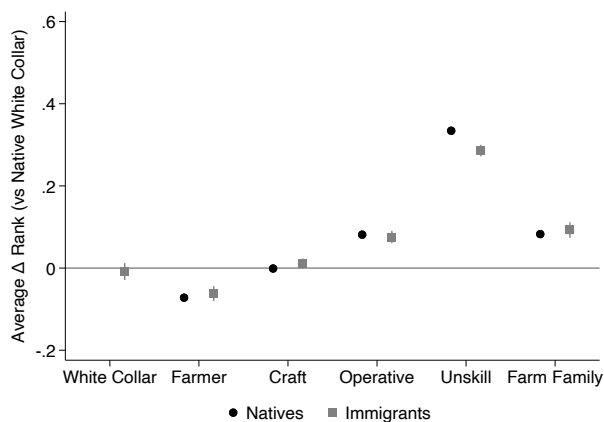
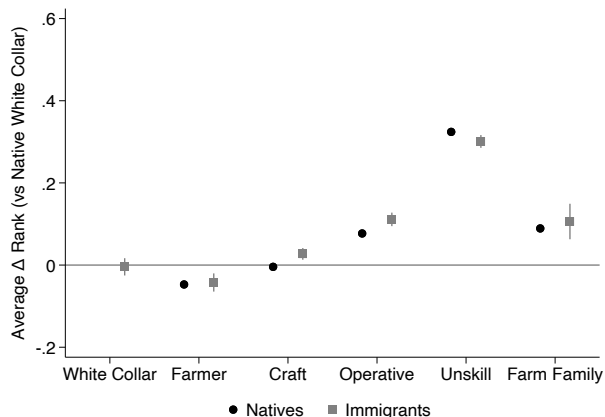
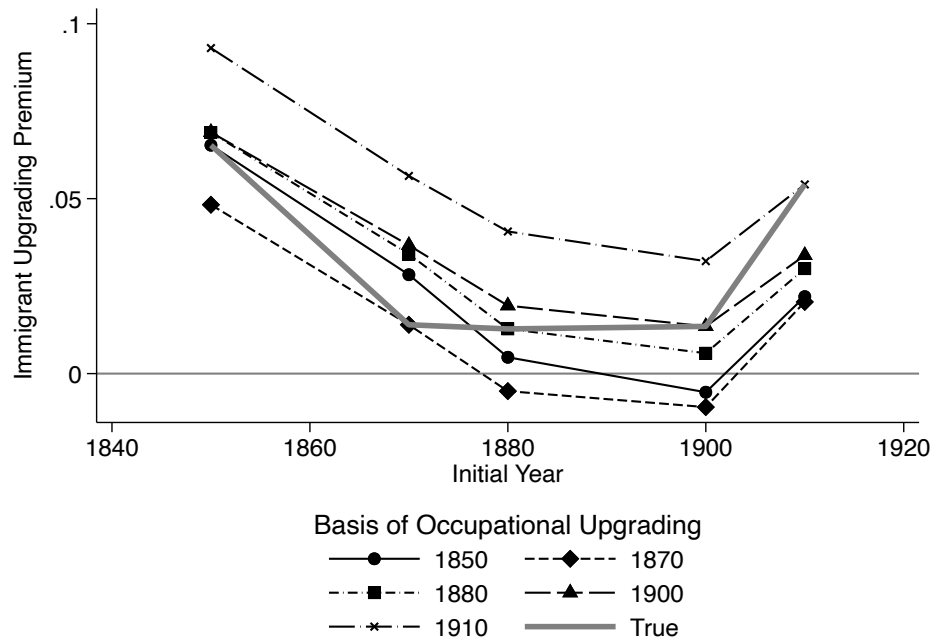


Figure C.33(e): 1910-1940



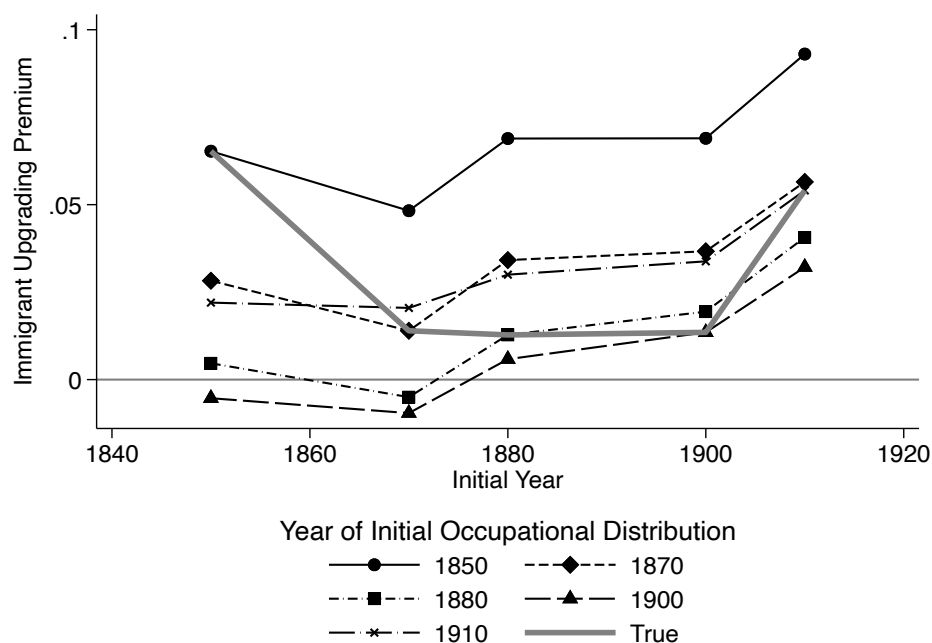
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.34: Unconditional assimilation, holding upgrading constant



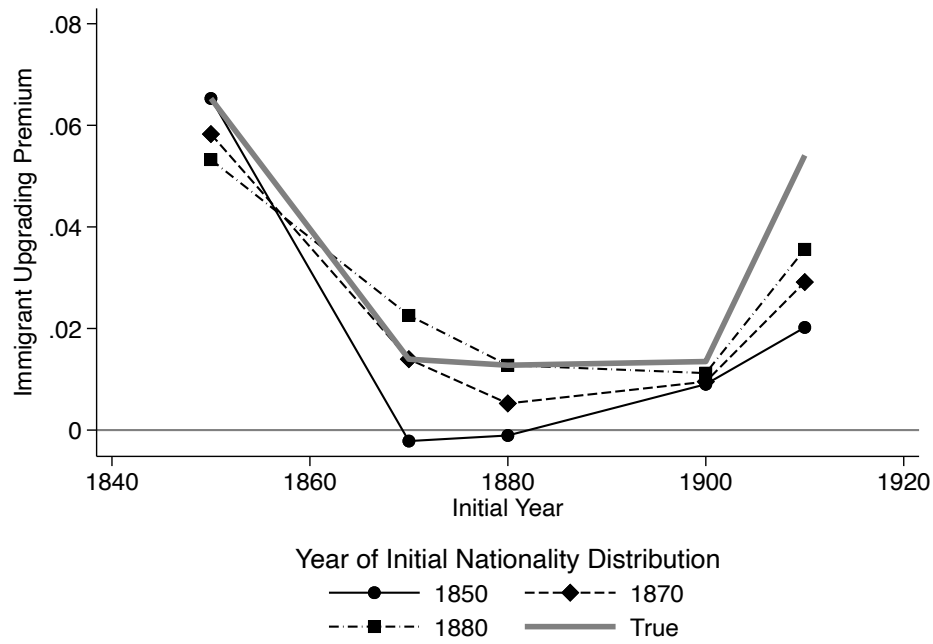
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.35: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure C.36: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Appendix D: Results with imputed occupation codes

The census microdata for the censuses of 1900, 1910, and 1930 (and 1920, though we do not use this census in our main results) provided by Ruggles et al. (2020) are preliminary. One consequence of this is that a considerable portion of these samples (9.8% of non-southern white men aged 18-30 in 1900, 16.0% in 1910, and 16.2% in 1930) have occupations that are listed as “Not Yet Classified” (*occ1950* code 979). Individuals classified in this way are omitted from the samples used in the main text. To ensure that the exclusion of a sizable number of observations from our analysis as a result of this missing code is not responsible for our findings, this Appendix repeats our main results with imputed occupational codes for these unclassified individuals. Specifically, we assign these uncategorized individuals the modal occupational code assigned to all occupational strings with a matching NYSIIS code of the listed occupational string, in a manner similar to Collins and Wanamaker (2021). We continue to use the listed occupational code for cases that were assigned a code by Ruggles et al. (2020). Our main results are qualitatively unaffected by including these individuals, implying that excluding these unclassified individuals does not drive our results.

Figure D.1: Occupational distributions

Figure D.1(a): Initial Year 1850

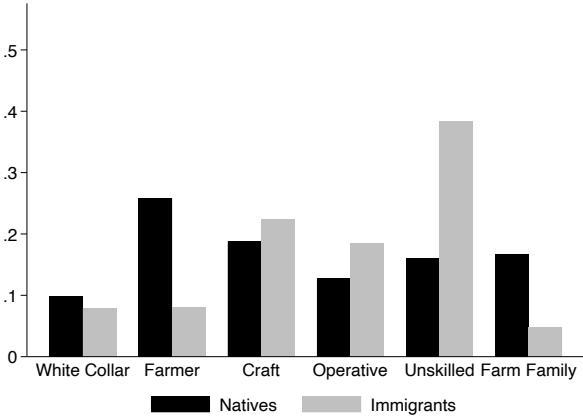


Figure D.1(b): Final Year 1880

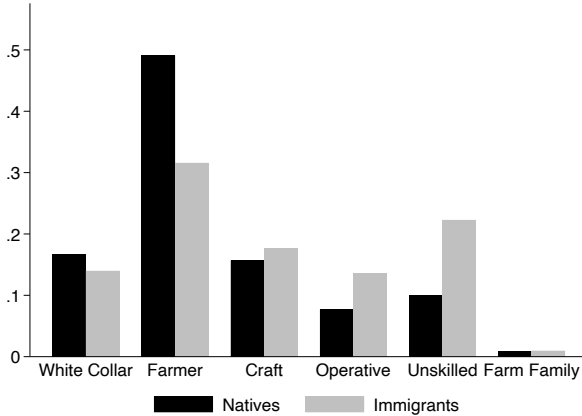


Figure D.1(c): Initial Year 1870

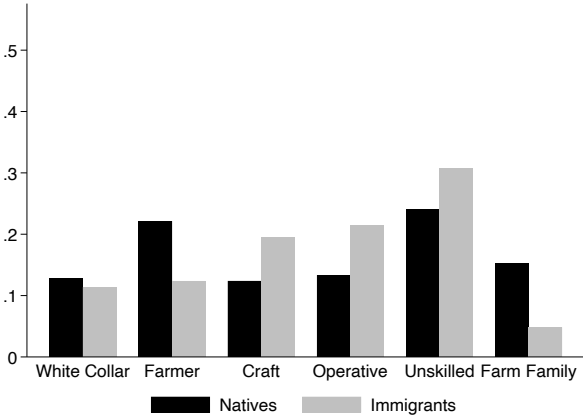


Figure D.1(d): Final Year 1900

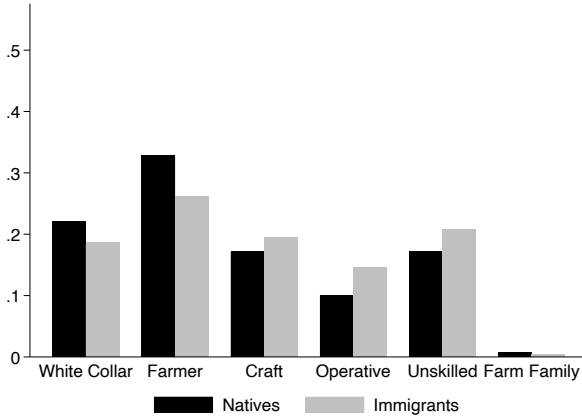


Figure D.1(e): Initial Year 1880

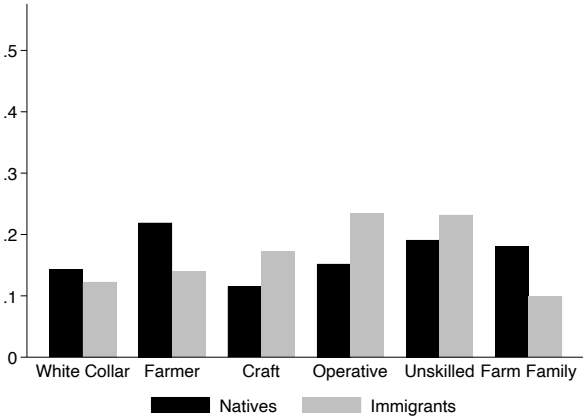


Figure D.1(f): Final Year 1910

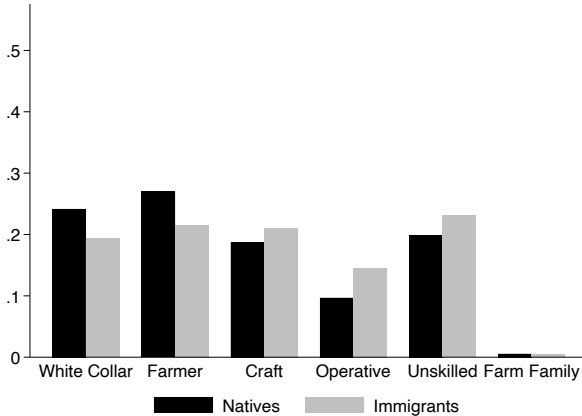


Figure D.1 (continued): Occupational distributions

Figure D.1(g): Initial Year 1900

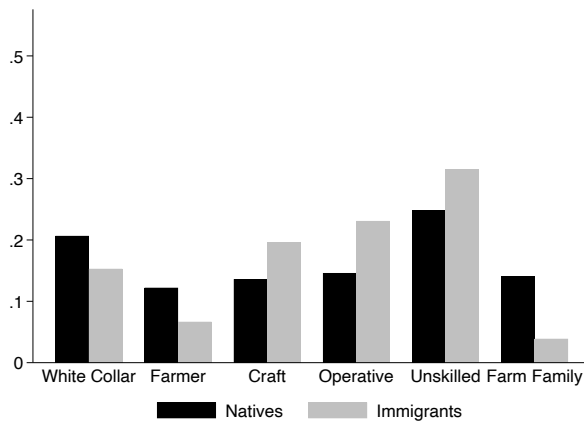


Figure D.1(h): Final Year 1930

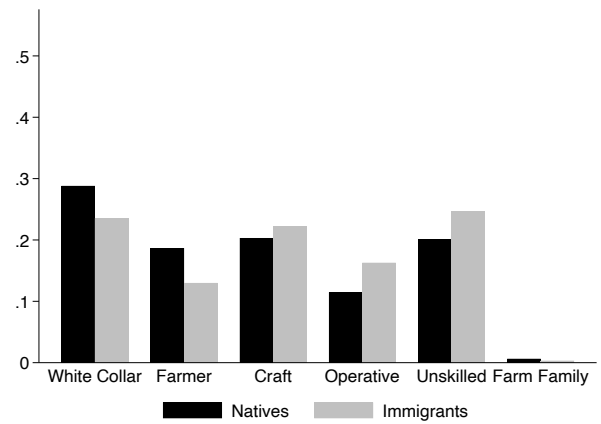


Figure D.1(i): Initial Year 1910

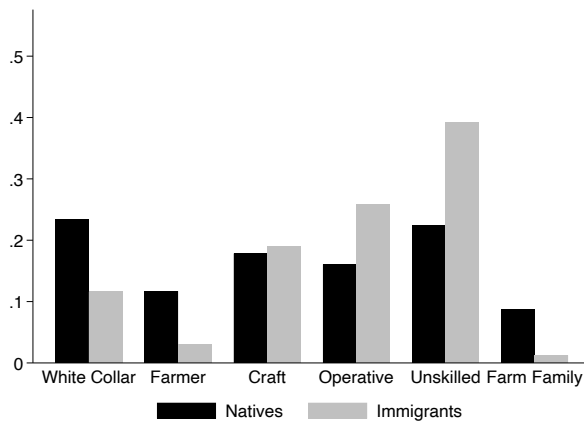
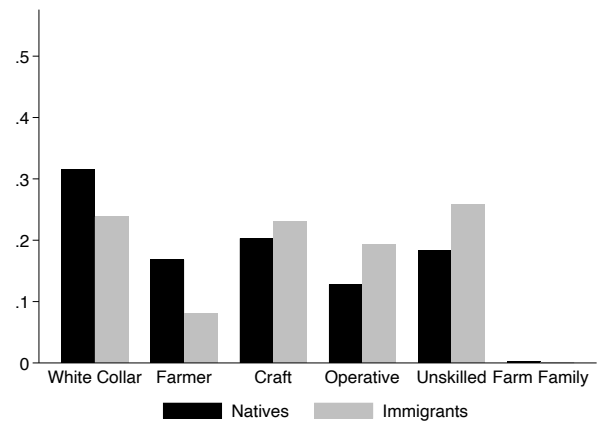
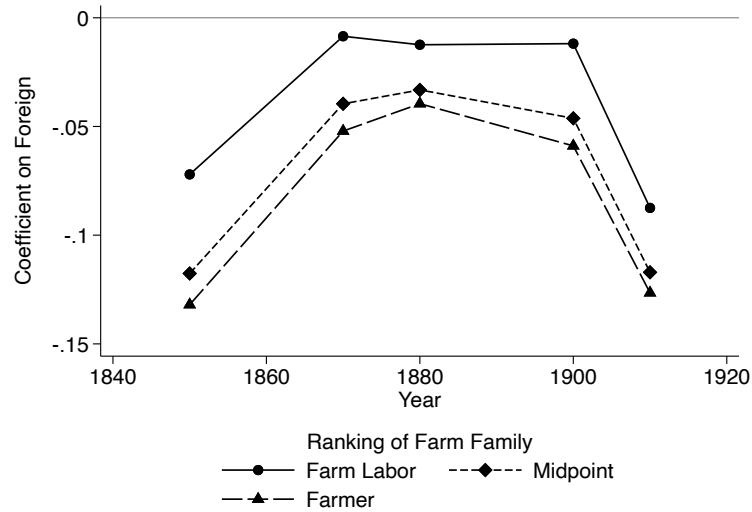


Figure D.1(j): Final Year 1940



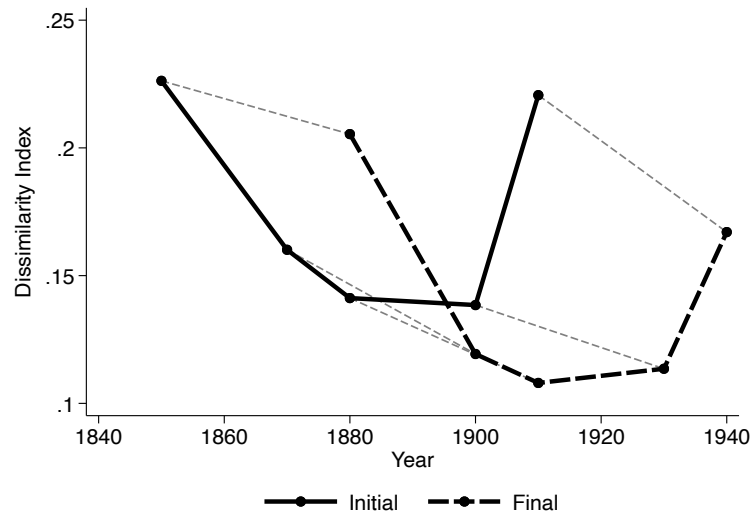
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure D.2: Initial gaps in occupational rank



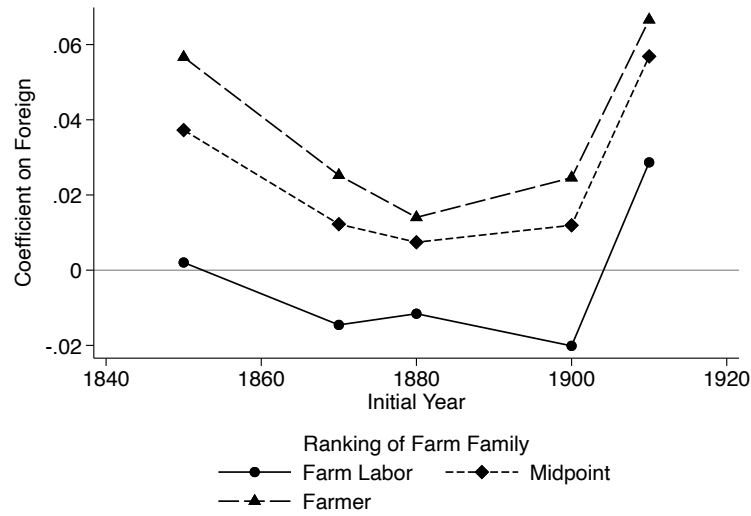
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure D.3: Dissimilarity indices



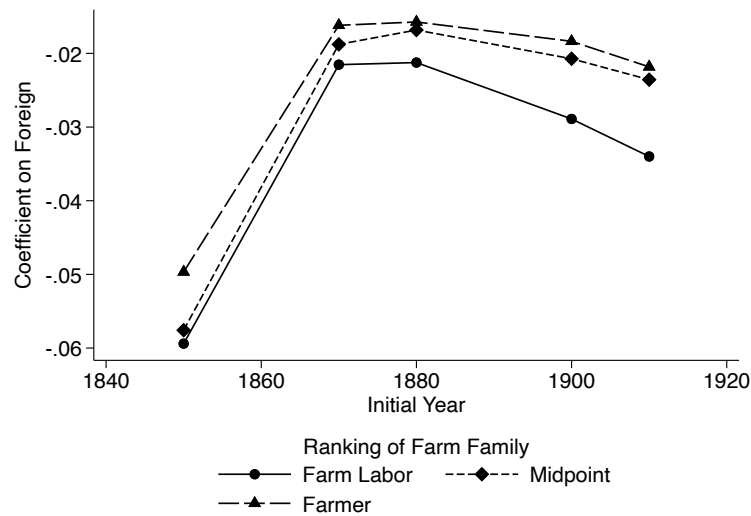
Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure D.4: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure D.5: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure D.6: Occupational upgrading by initial occupation

Figure D.6(a): 1850-1880

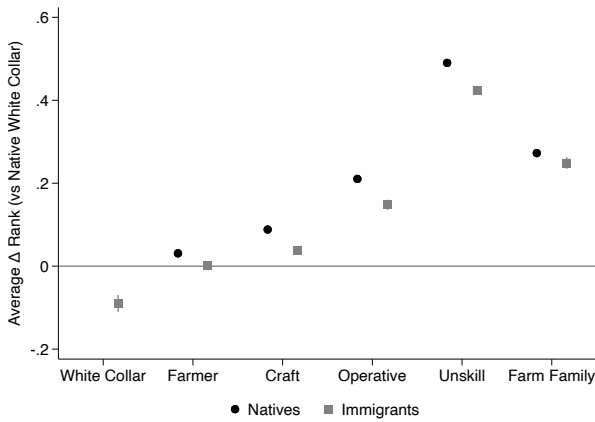


Figure D.6(b): 1870-1900

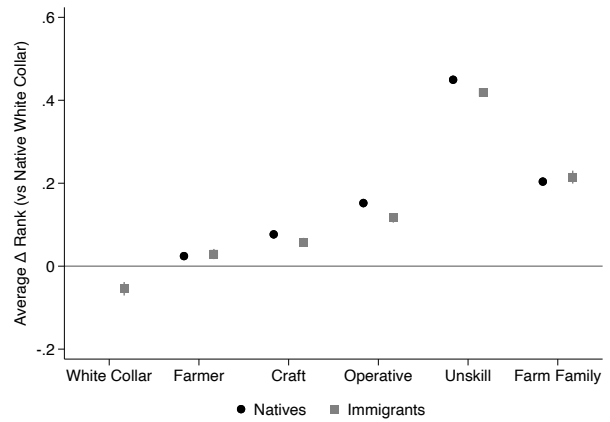


Figure D.6(c): 1880-1910

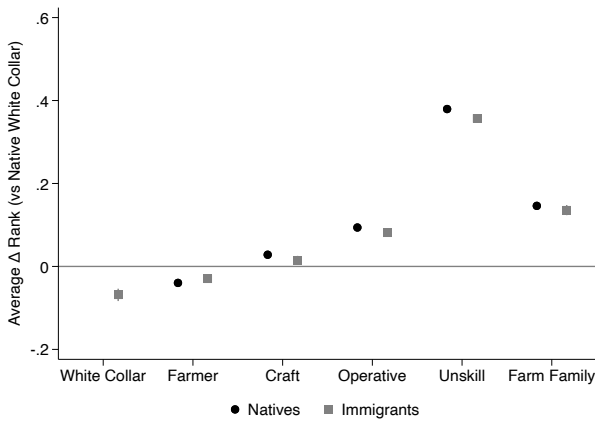


Figure D.6(d): 1900-1930

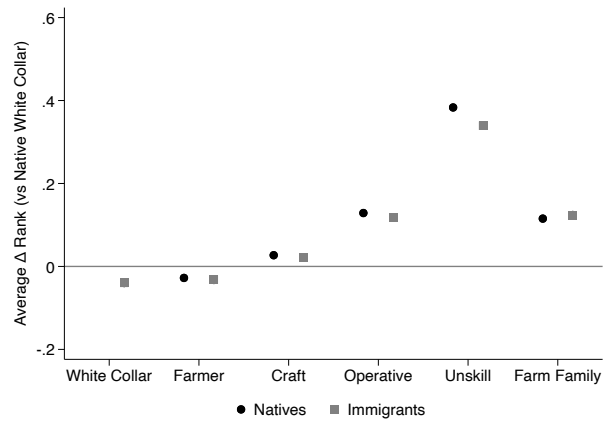
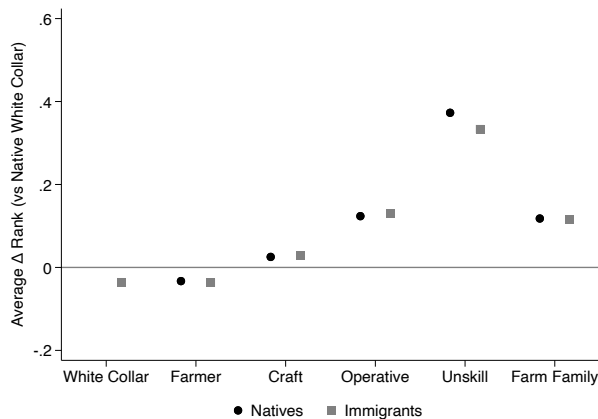
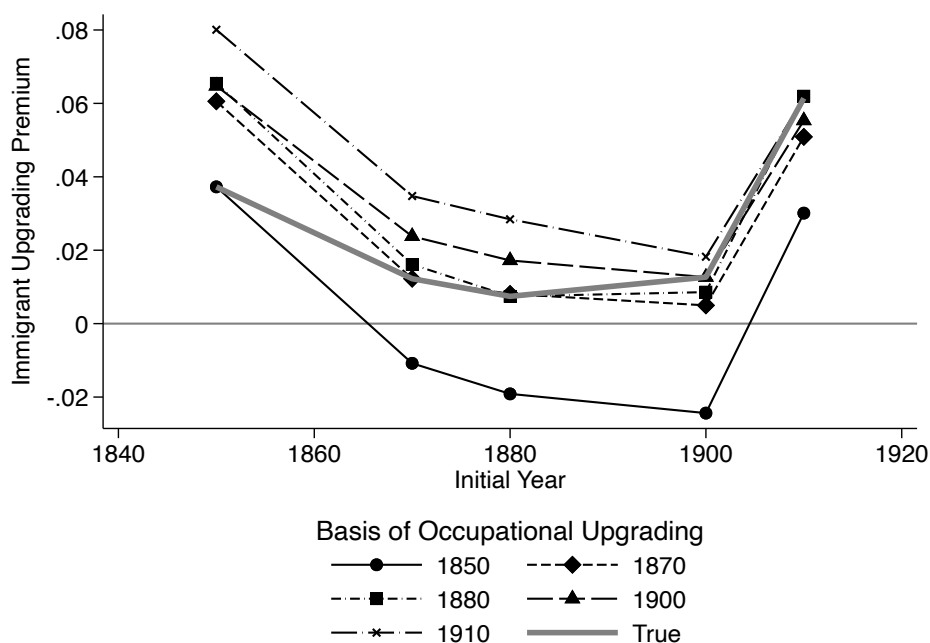


Figure D.6(e): 1910-1940



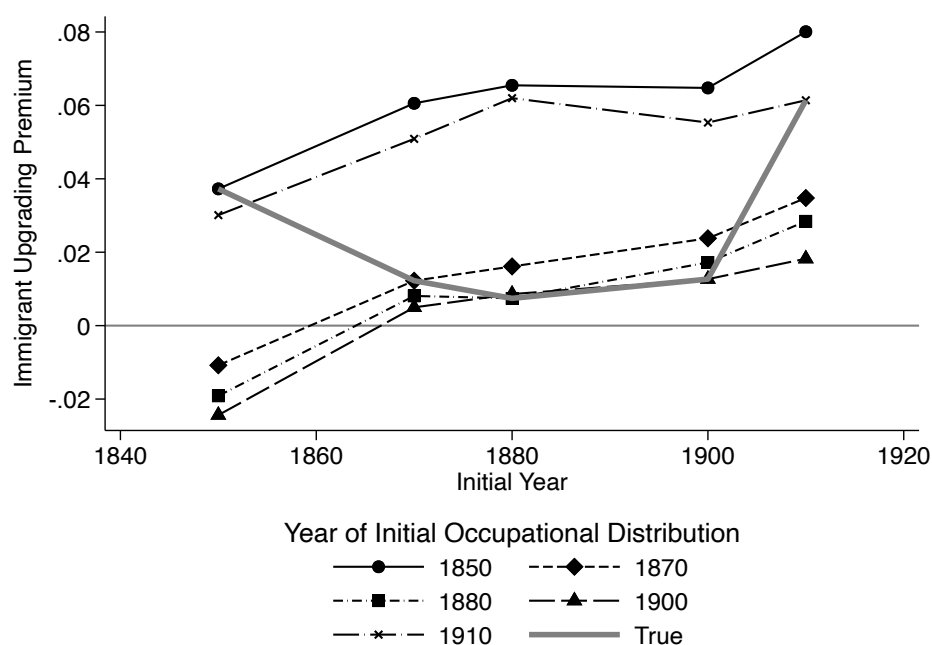
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure D.7: Unconditional assimilation, holding upgrading constant



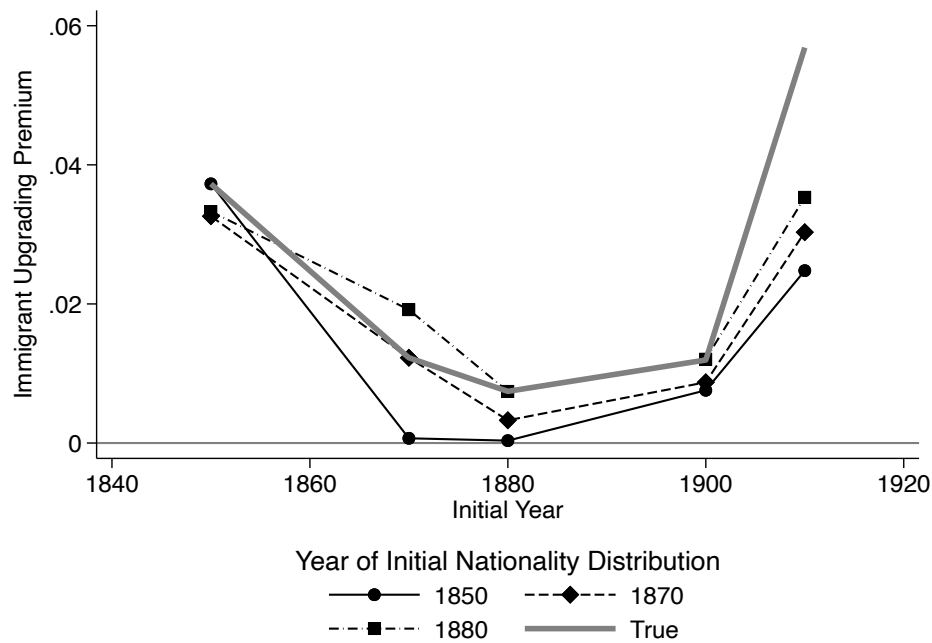
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure D.8: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure D.9: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Appendix E: Procedure to recategorize farm family members in 1850

In a sample of all non-southern white men aged 18-30 who were not household heads, lived in a household headed by a farmer, and reported an agricultural occupation (occ1950 codes 100-199 and 800-899), we estimate a probit model of the form

$$P(y_i = 1) = \Phi(\mathbf{x}_i' \beta + \varepsilon_i) \quad (1)$$

where y_i is an indicator taking a value of 1 if an individual was classified as a farmer and 0 if they were classified as a farm laborer, and \mathbf{x}_i is a vector containing a variety of individual characteristics that can be observed in both 1850 and 1870. The most important of these is indicators for relation to the head of household. We then take the estimates of this model to the 1850 census, limiting attention to the same class of individuals as in 1870. For each such individual in the 1850 census, we compute $\hat{p}_i = \Phi(\mathbf{x}_i' \hat{\beta})$, where the $\hat{\beta}$ are from the estimates of equation (1) using the 1870 data and the \mathbf{x}_i are from 1850. Finally, we randomly assign these 1850 individuals to be either farmers (with probability \hat{p}_i) or farm laborers. That is, we drew a uniform random variable η_i and assigned individuals to be farmers if $\eta_i < \hat{p}_i$ and farm laborers otherwise.

Appendix F: Results with 20-year linkage spans

This appendix repeats the main results with linkage spans of 20 years instead of the 30-year linkage spans of the main text. The specific linkage spans included are 1850-1870, 1880-1900, 1900-1920, 1910-1930, and 1920-1940. The 1860-1880 linkage span is not used due to issues mentioned in the main text regarding the coding of occupations in the 1860 census. The main benefit of the 20-year linkage spans is that it enables us to study the cohort of immigrants in the United States by 1910 while linking them to a census in which immigrants' year of arrival is reported. This enables us to construct weights that do not include eventual returners in the sample at risk for linkage. Although the assimilation measures for this 1910 cohort are lower than those in the main text, consistent with negatively selected return migration being an issue, there is still a rise in the assimilation measure from the 1900 cohort to the 1910 cohort. Moreover, the assimilation for the 1910 cohort can effectively be bounded below at zero and is substantially positive for our preferred midpoint ranking of farm family members.

Figure F.1: Occupational distributions

Figure F.1(a): Initial Year 1850

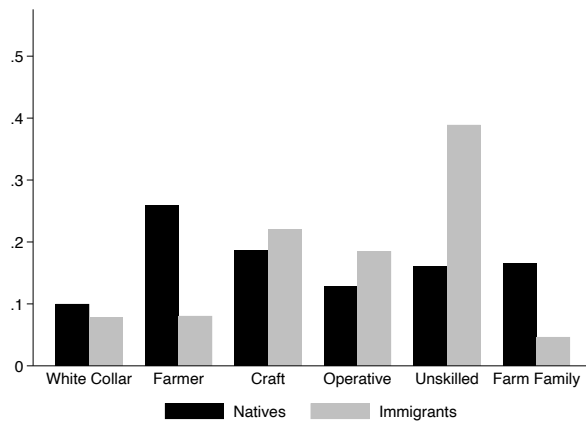


Figure F.1(b): Final Year 1870

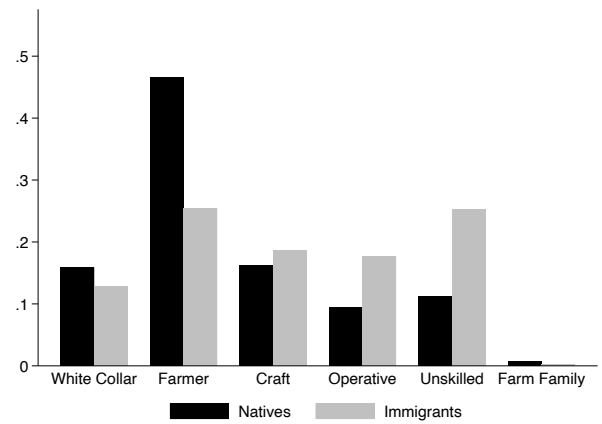


Figure F.1(c): Initial Year 1880

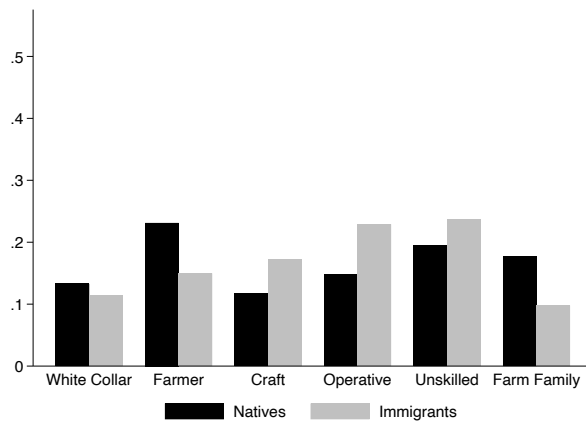


Figure F.1(d): Final Year 1900

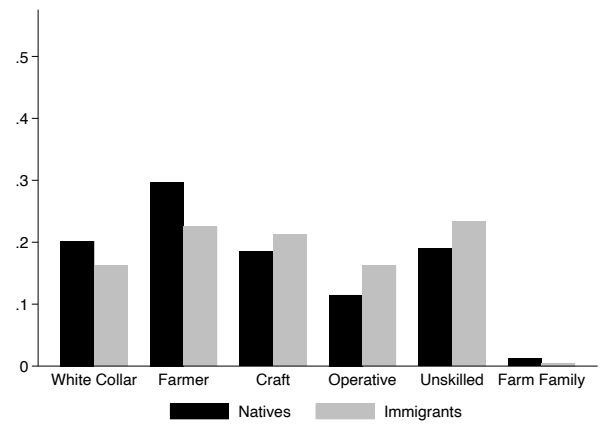


Figure F.1(e): Initial Year 1900

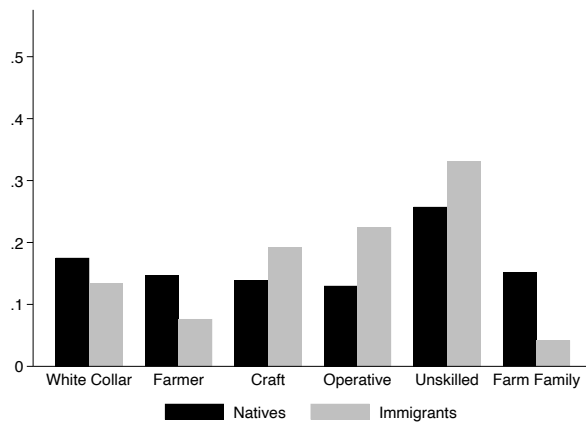


Figure F.1(f): Final Year 1920

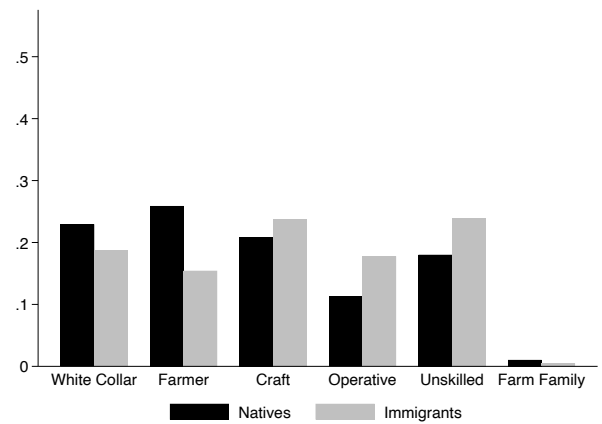


Figure F.1 (continued): Occupational distributions

Figure F.1(g): Initial Year 1910

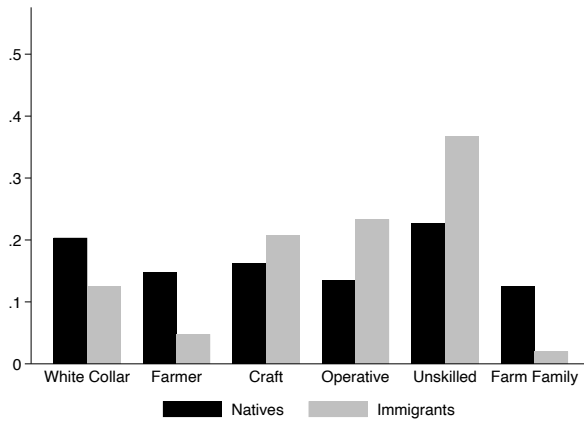


Figure F.1(h): Final Year 1930

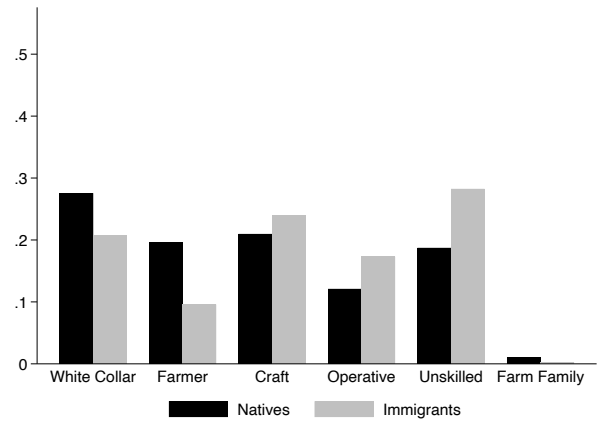


Figure F.1(i): Initial Year 1920

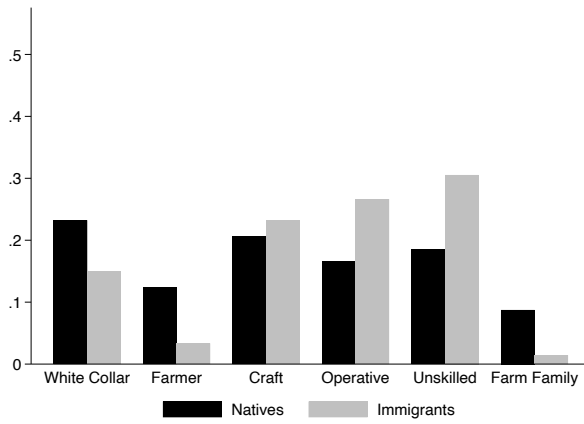
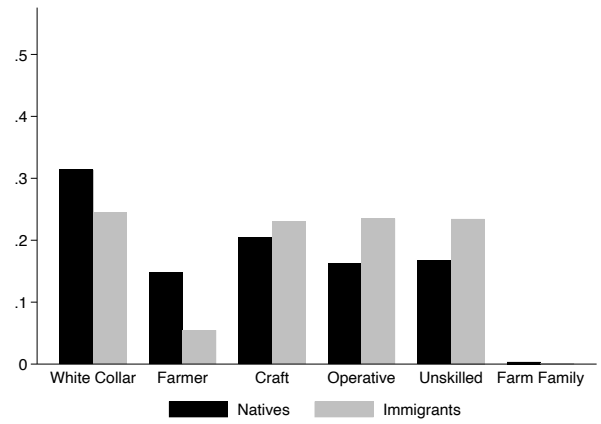
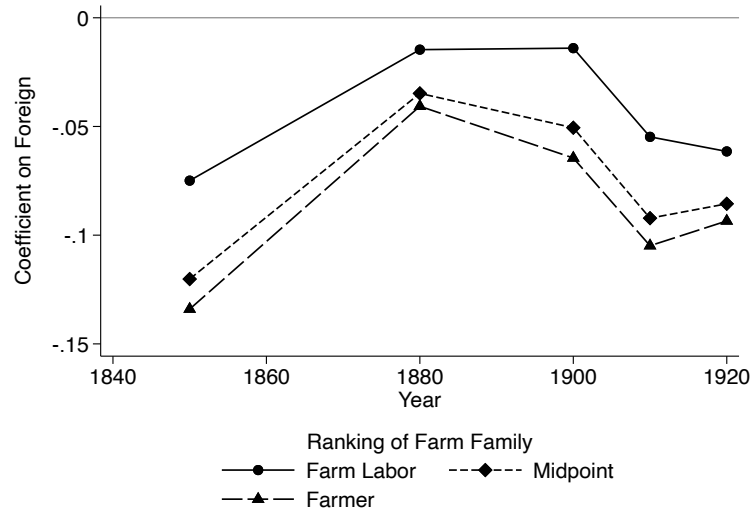


Figure F.1(j): Final Year 1940



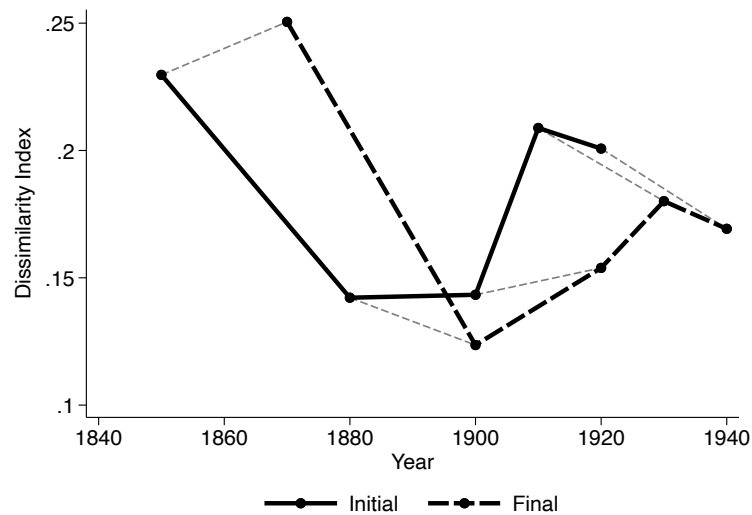
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure F.2: Initial gaps in occupational rank



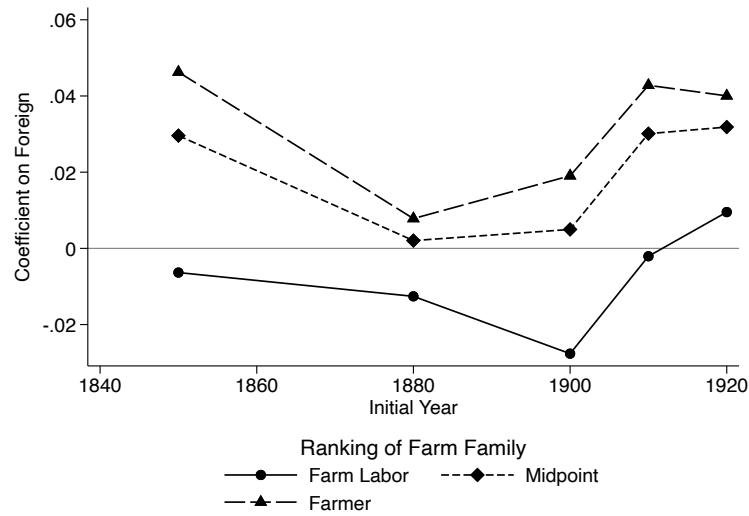
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure F.3: Dissimilarity indices



Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure F.4: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure F.5: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure F.6: Occupational upgrading by initial occupation

Figure F.6(a): 1850-1870

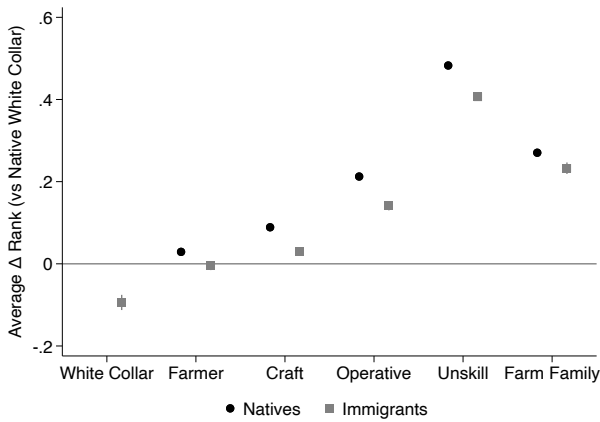


Figure F.6(b): 1880-1900

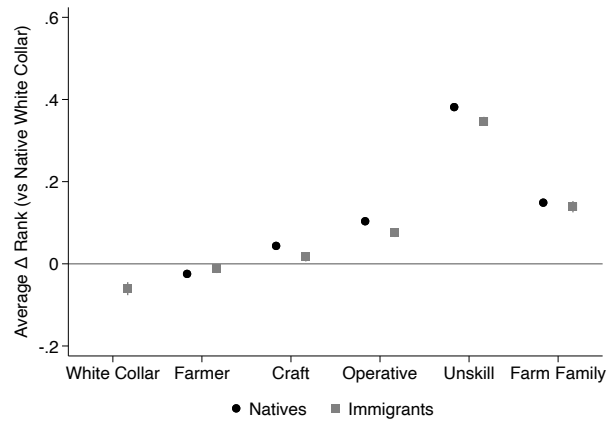


Figure F.6(c): 1900-1920

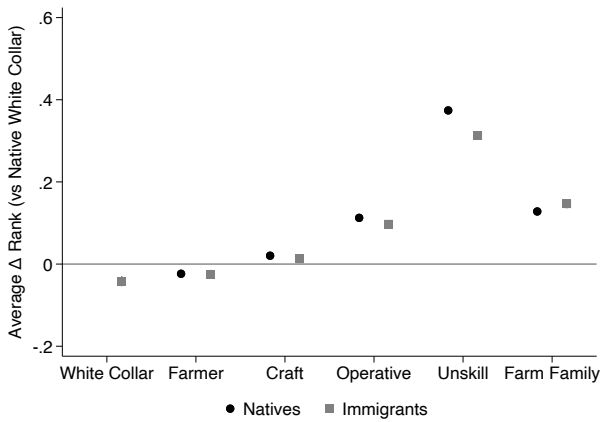


Figure F.6(d): 1910-1930

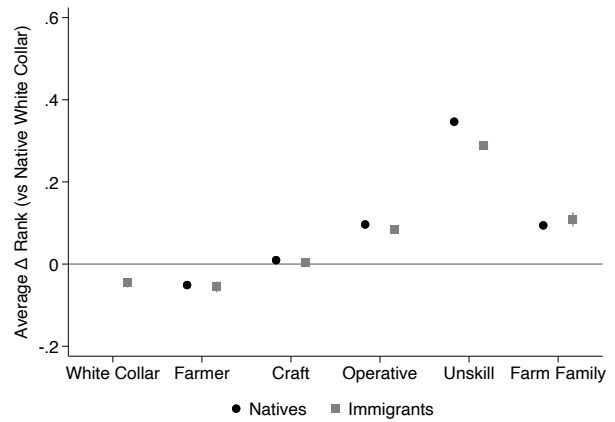
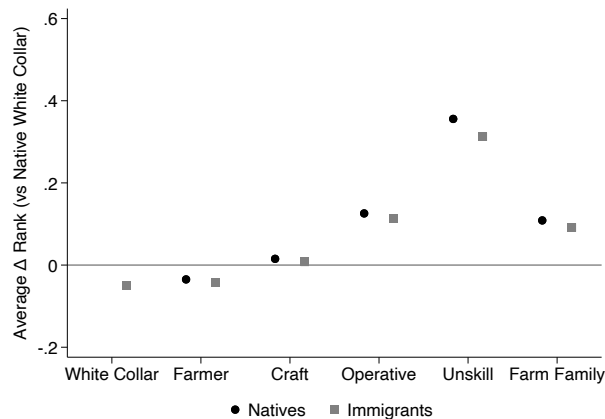
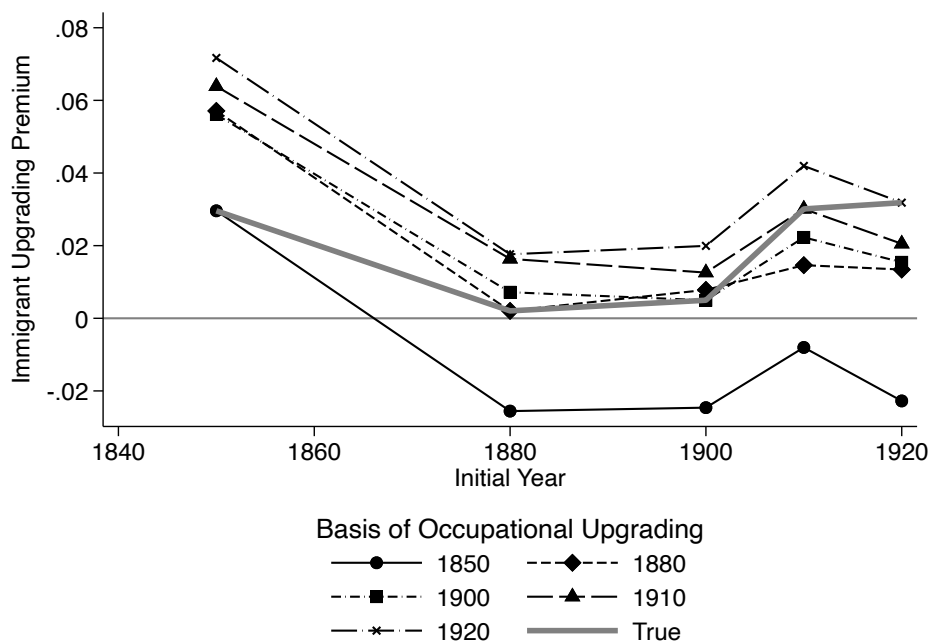


Figure F.6(e): 1920-1940



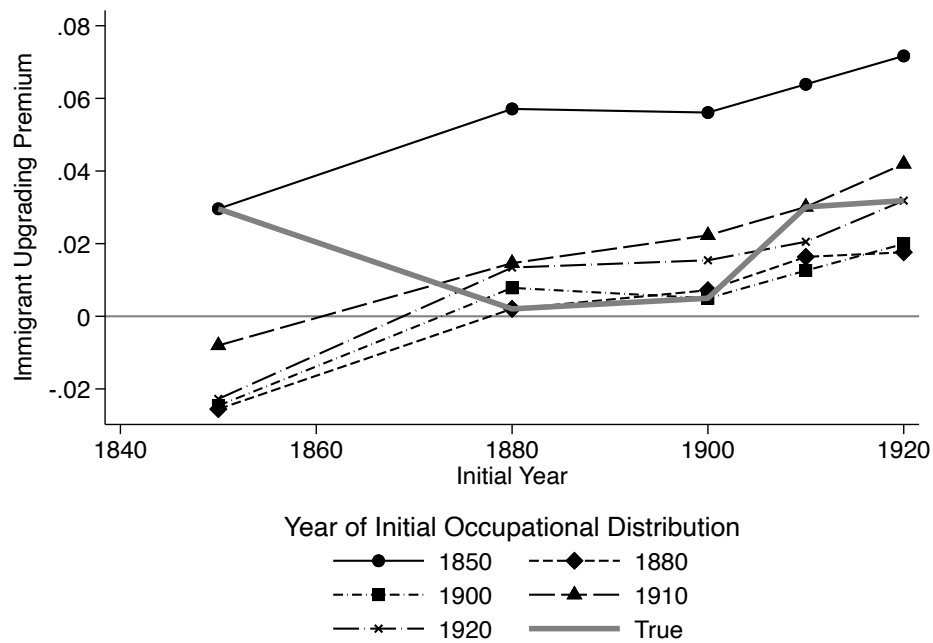
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure F.7: Unconditional assimilation, holding upgrading constant



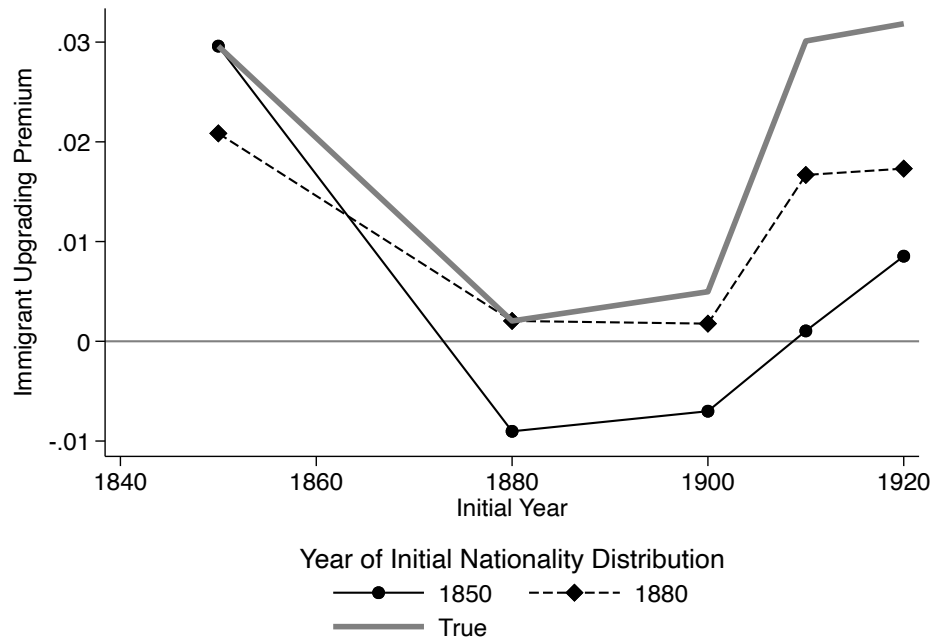
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure F.8: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure F.9: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Appendix G: Results with alternate weighting methods

This Appendix adopts an alternative approach to reweighting the linked samples to address selection into linkage on the basis of observable characteristics. In the main text, the 1870-1900, 1880-1910, and 1900-1930 samples are weighted to match the population at risk for linkage in the latter year of the sample. This is done because the population at risk for linkage in the initial year includes those who would eventually return to their home countries and are thus not of interest in our analysis. This creates a concern, however, that the decline in immigrant assimilation from the 1850-1880 cohort to the intermediate ones may have come from an increase in return migration over this period, which might have caused individuals who experienced the most upgrading to return to their home countries and not be observed, thus reducing observed assimilation. Weighting the sample to match the population at risk for linkage in the initial year in effect increases the weight assigned to the lower (and most upwardly mobile) part of the immigrant distribution, but does not qualitatively change our results.

Figure G.1: Occupational distributions

Figure G.1(a): Initial Year 1850

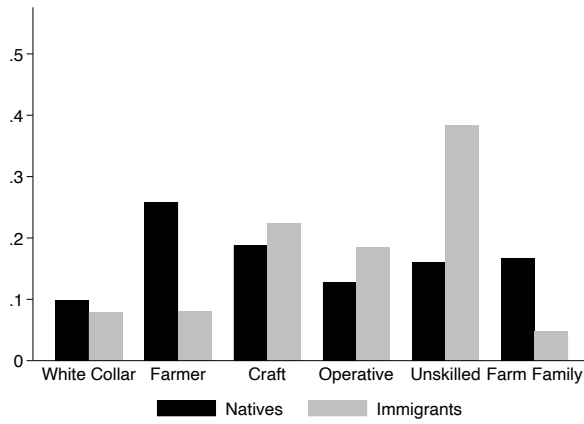


Figure G.1(b): Final Year 1880

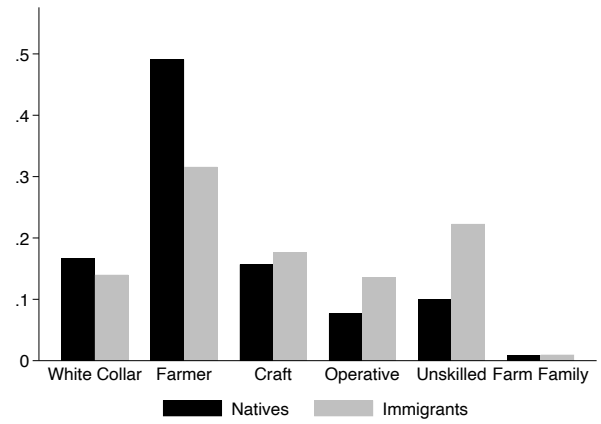


Figure G.1(c): Initial Year 1870

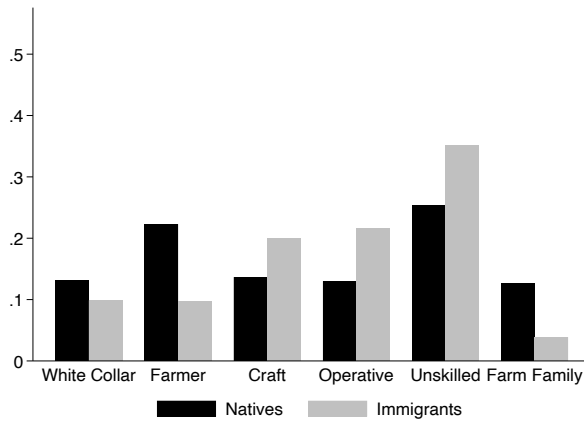


Figure G.1(d): Final Year 1900

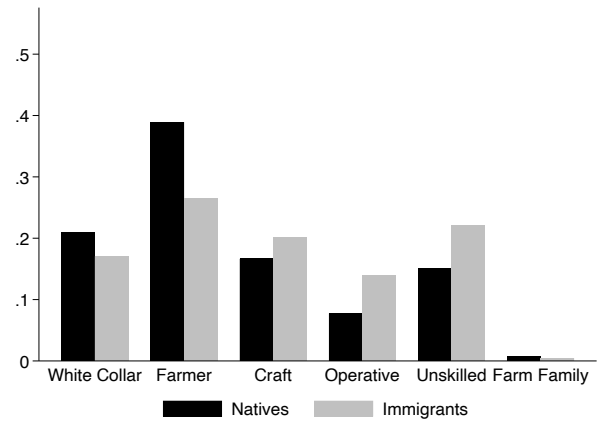


Figure G.1(e): Initial Year 1880

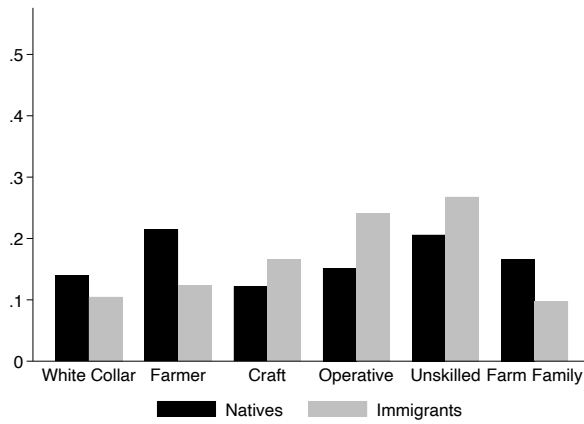


Figure G.1(f): Final Year 1910

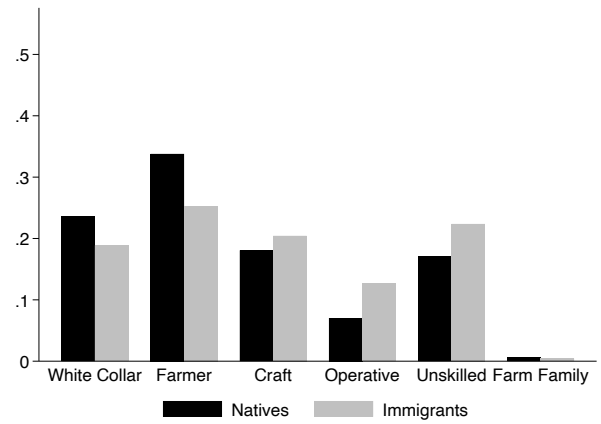


Figure G.1 (continued): Occupational distributions

Figure G.1(g): Initial Year 1900

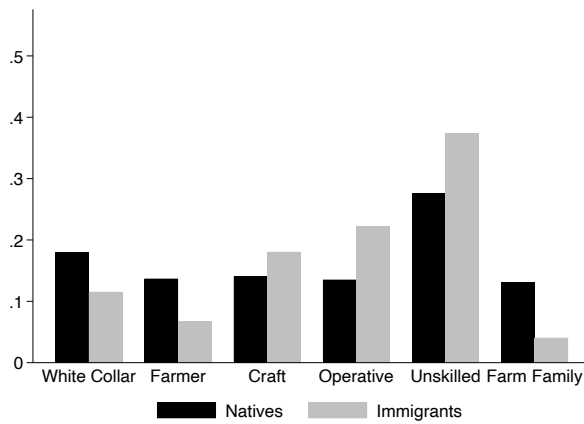


Figure G.1(h): Final Year 1930

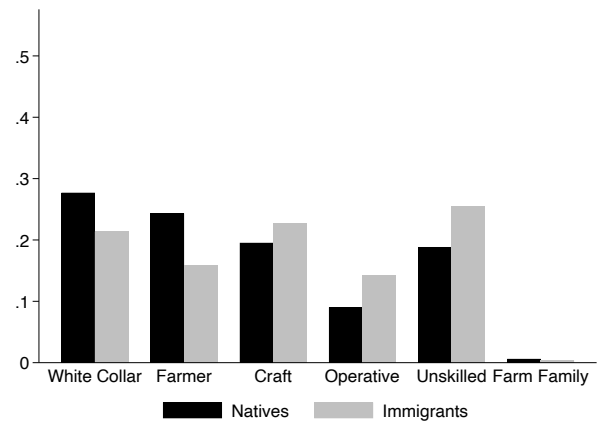


Figure G.1(i): Initial Year 1910

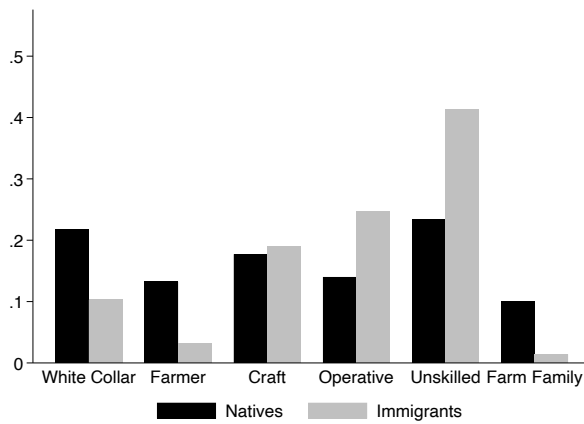
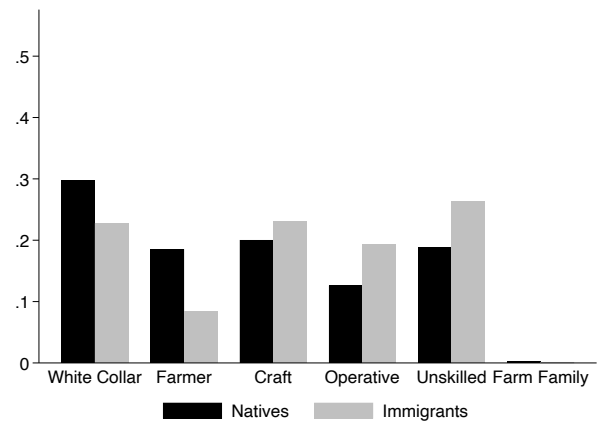


Figure G.1(j): Final Year 1940



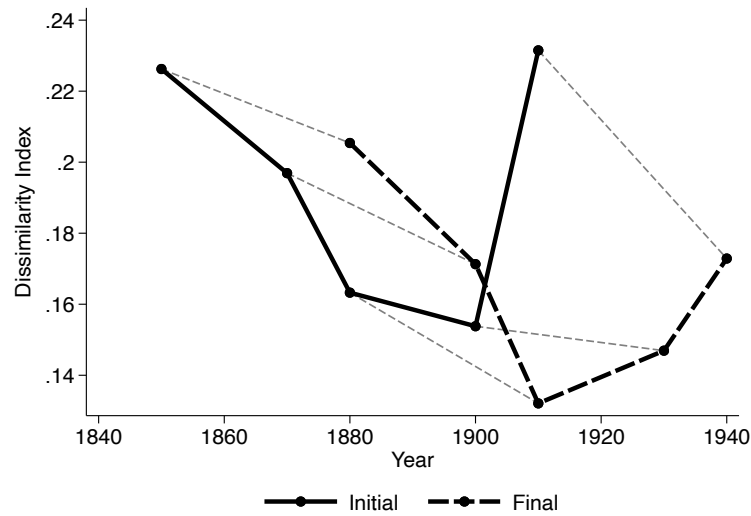
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure G.2: Initial gaps in occupational rank



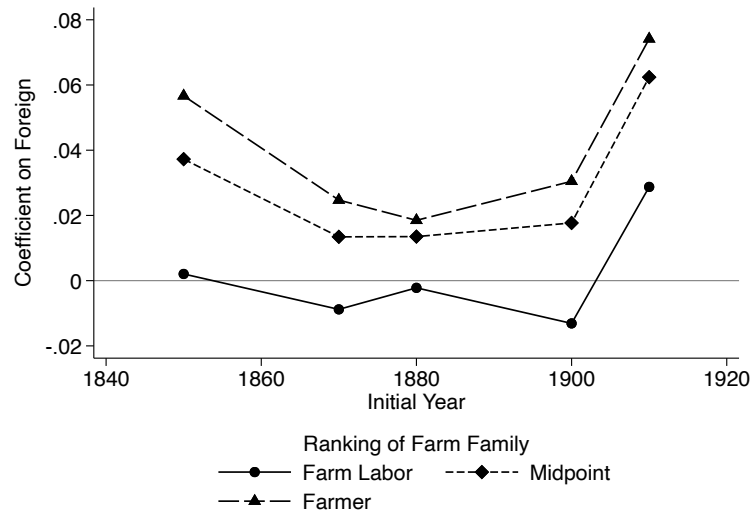
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure G.3: Dissimilarity indices



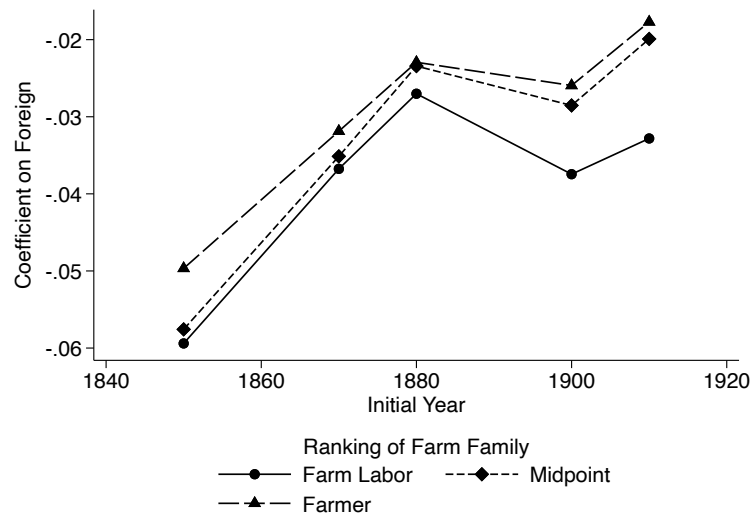
Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure G.4: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure G.5: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure G.6: Occupational upgrading by initial occupation

Figure G.6(a): 1850-1880

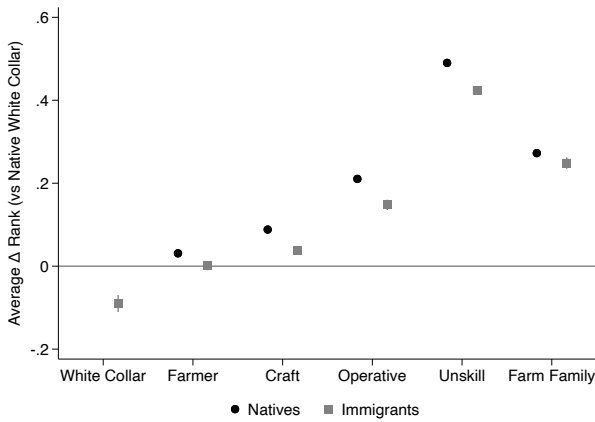


Figure G.6(b): 1870-1900

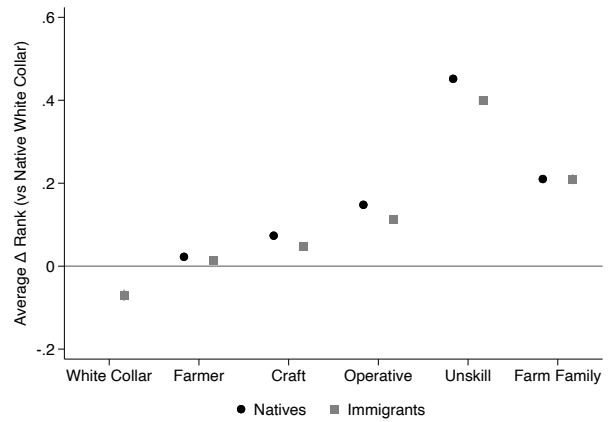


Figure G.6(c): 1880-1910

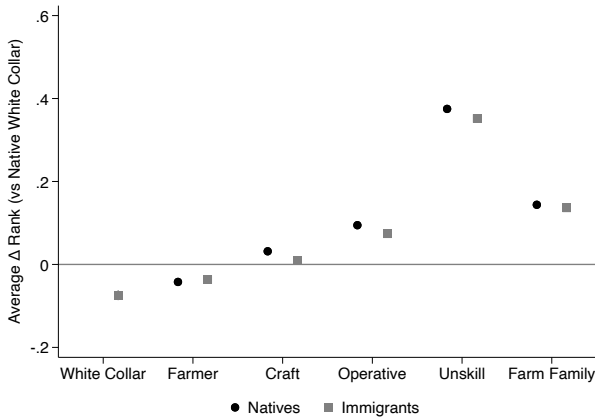


Figure G.6(d): 1900-1930

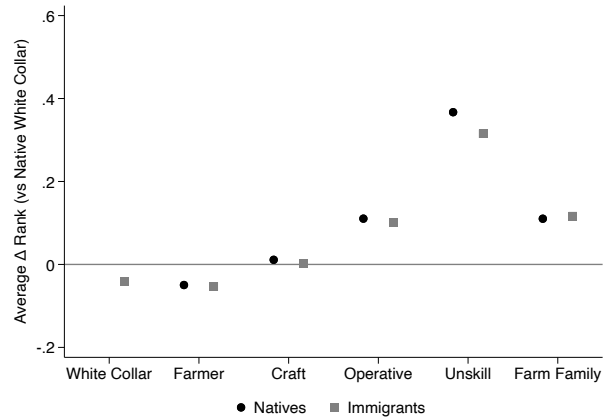
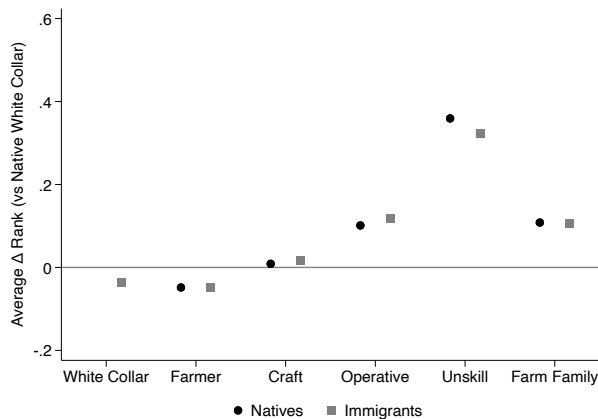
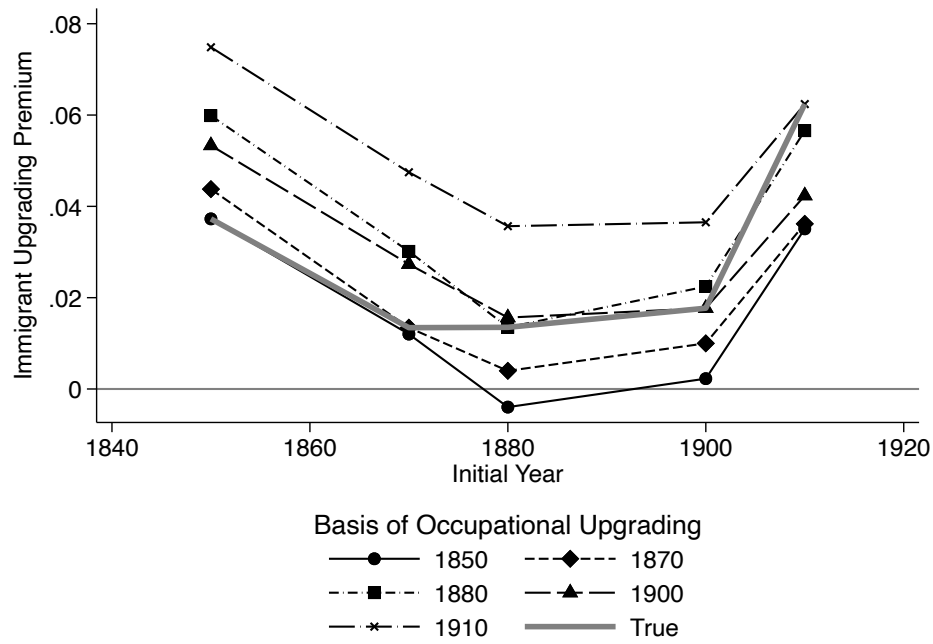


Figure G.6(e): 1910-1940



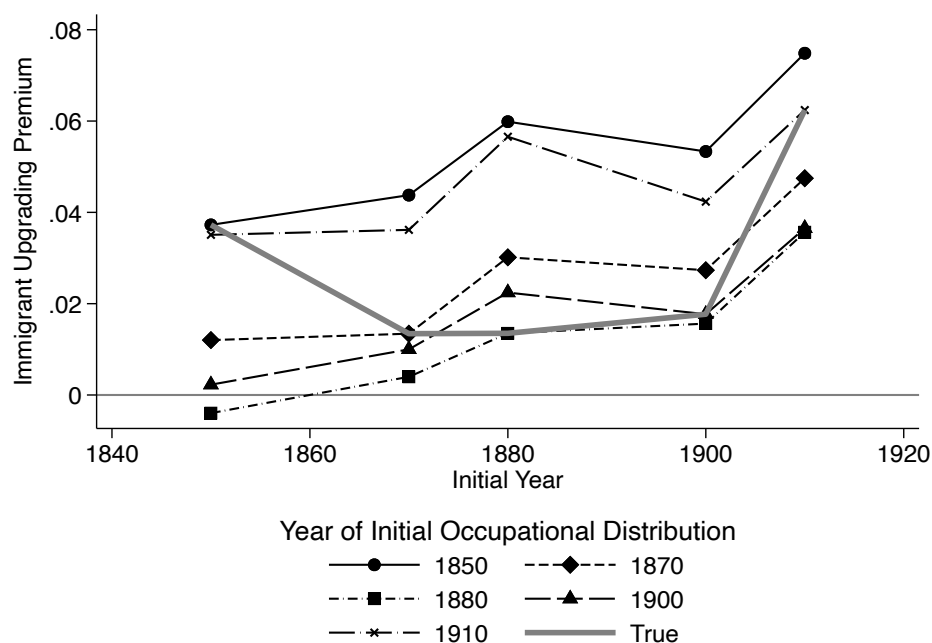
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure G.7: Unconditional assimilation, holding upgrading constant



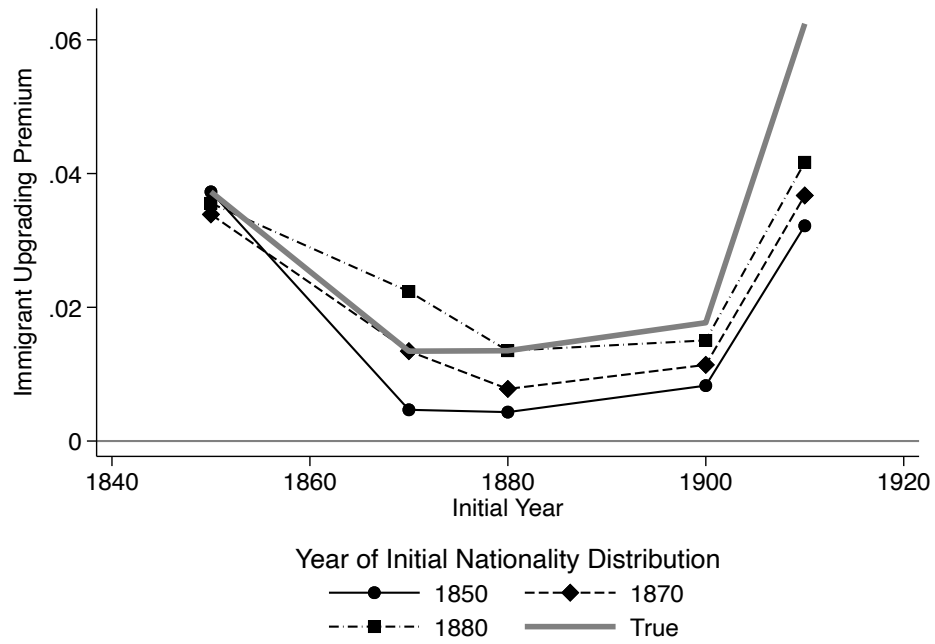
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure G.8: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure G.9: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Appendix H: Results omitting natives with foreign-born fathers

This Appendix omits from consideration native-born men who report that their fathers were foreign-born. The goal of this exercise is to address the concern that second-generation immigrants might be undergoing their own process of assimilation relative to the children of natives (Abramitzky et al. 2021b) and that comparing first-generation immigrants to second-generation immigrants might not capture the assimilation of interest. Our results are qualitatively unchanged relative to the main text.

Figure H.1: Occupational distributions

Figure H.1(a): Initial Year 1850

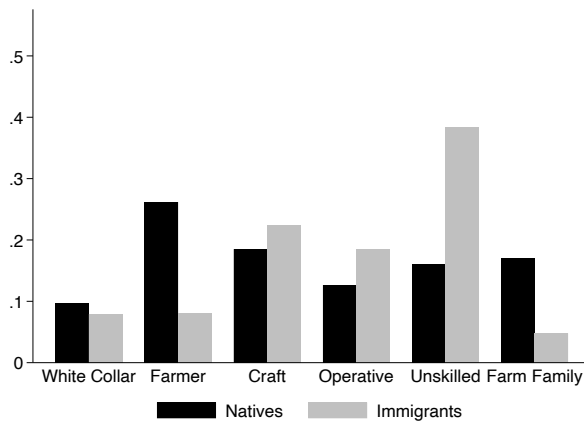


Figure H.1(b): Final Year 1880

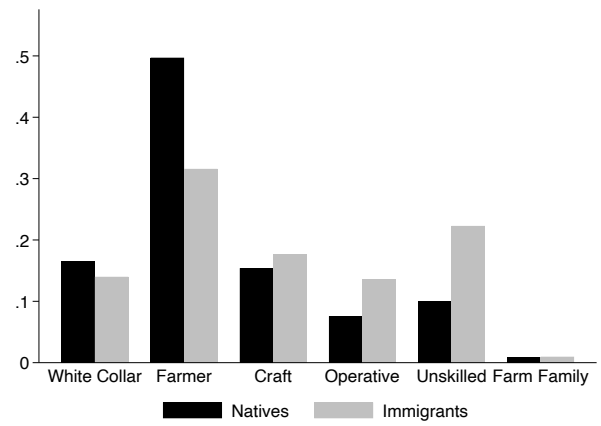


Figure H.1(c): Initial Year 1870

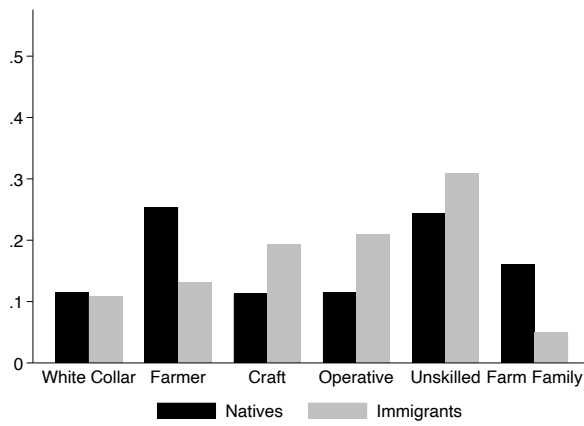


Figure H.1(d): Final Year 1900

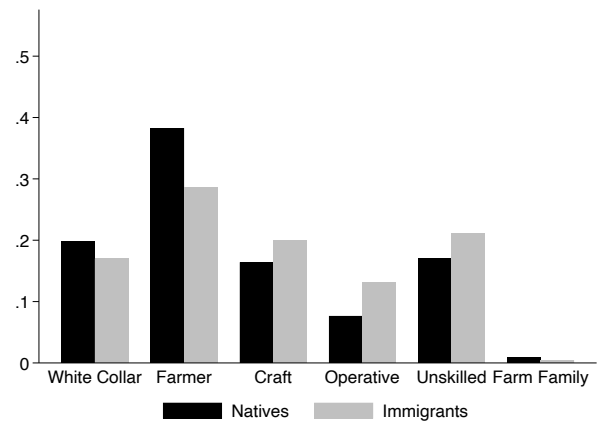


Figure H.1(e): Initial Year 1880

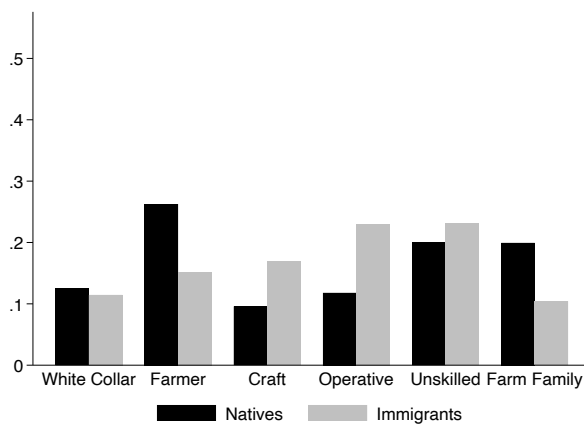


Figure H.1(f): Final Year 1910

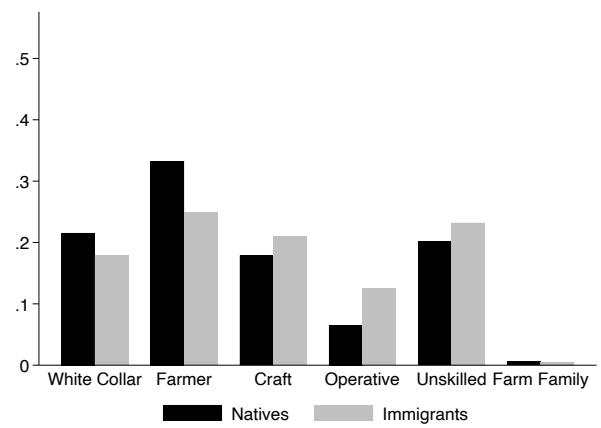


Figure H.1 (continued): Occupational distributions

Figure H.1(g): Initial Year 1900

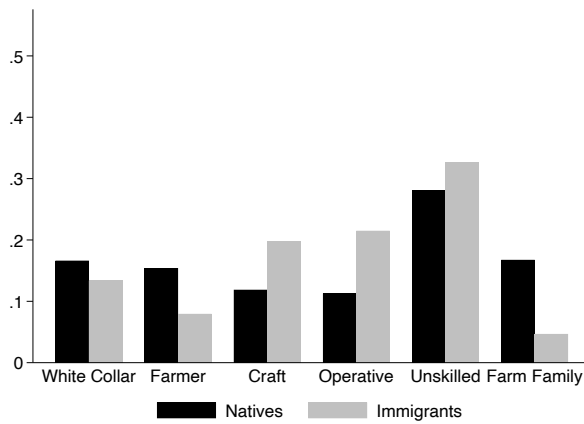


Figure H.1(h): Final Year 1930

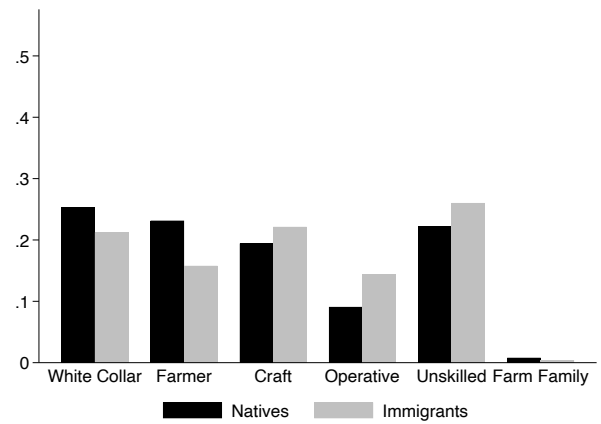


Figure H.1(i): Initial Year 1910

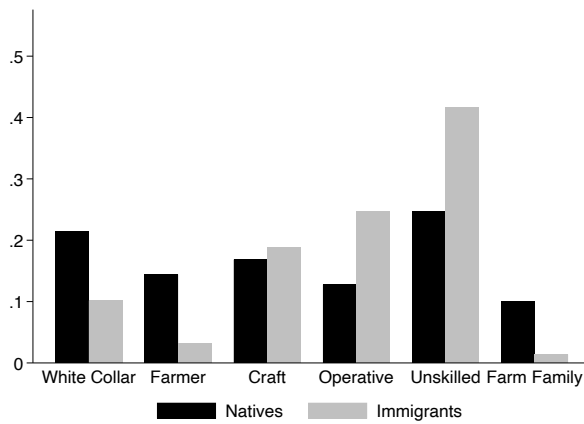
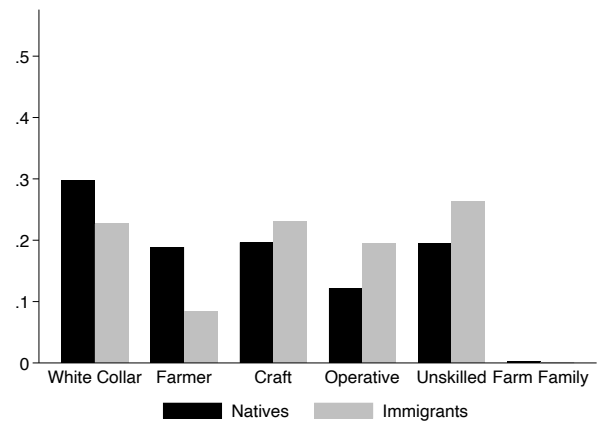
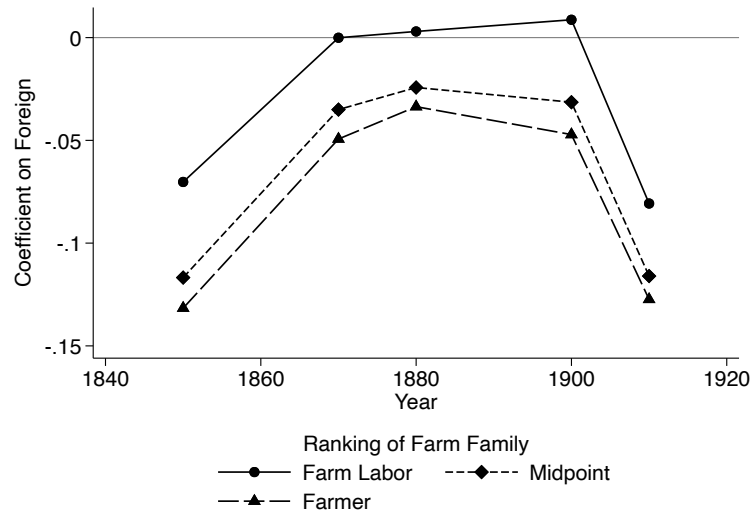


Figure H.1(j): Final Year 1940



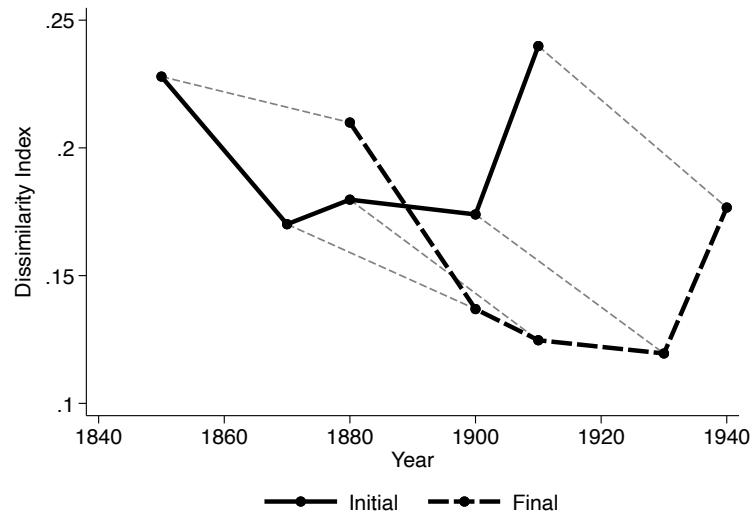
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure H.2: Initial gaps in occupational rank



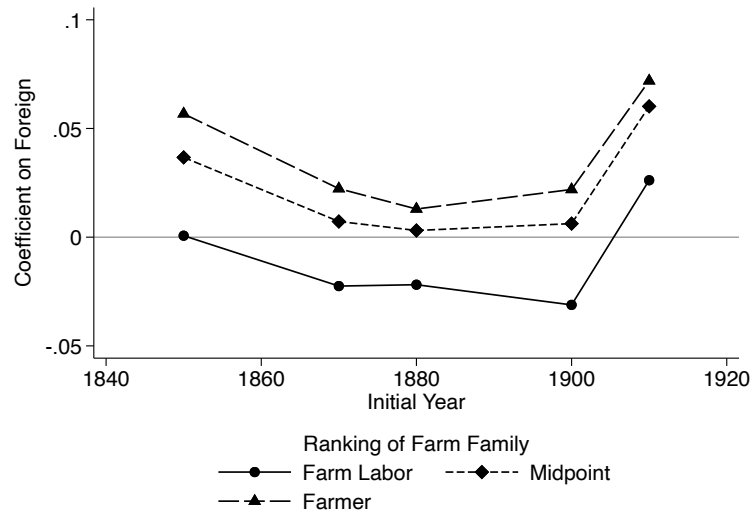
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure H.3: Dissimilarity indices



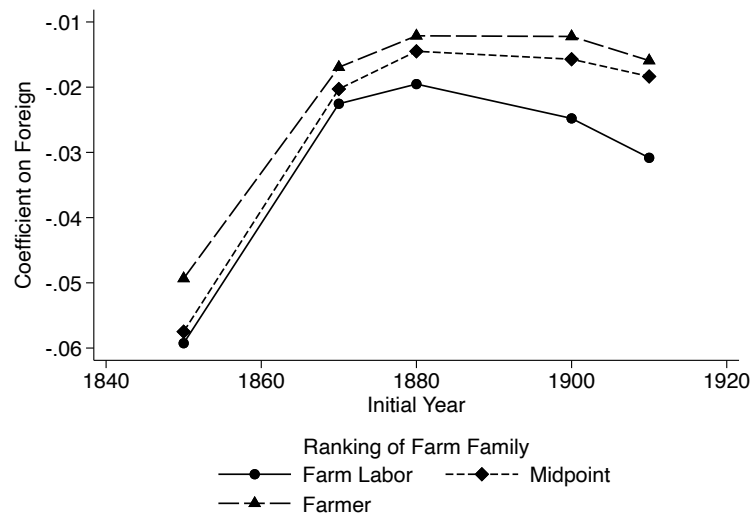
Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure H.4: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure H.5: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure H.6: Occupational upgrading by initial occupation

Figure H.6(a): 1850-1880

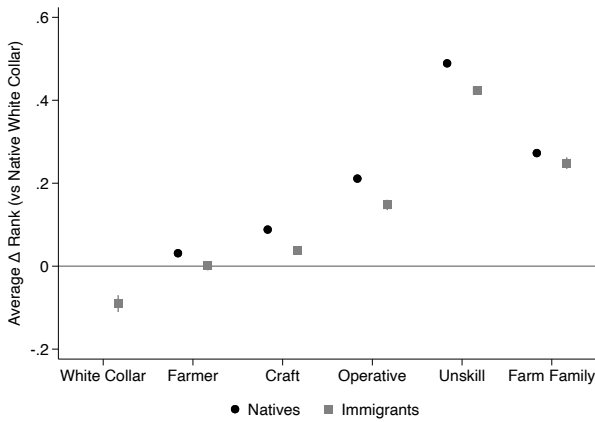


Figure H.6(b): 1870-1900

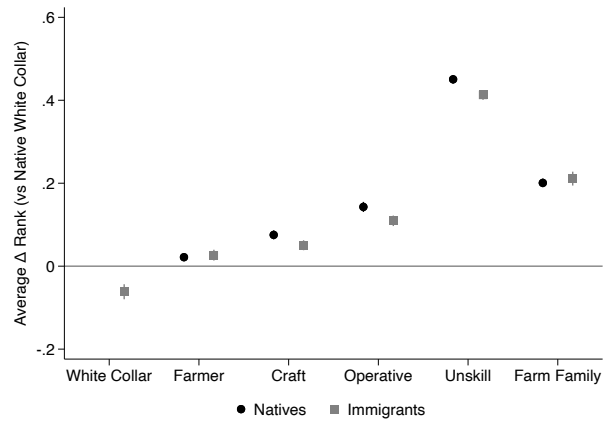


Figure H.6(c): 1880-1910

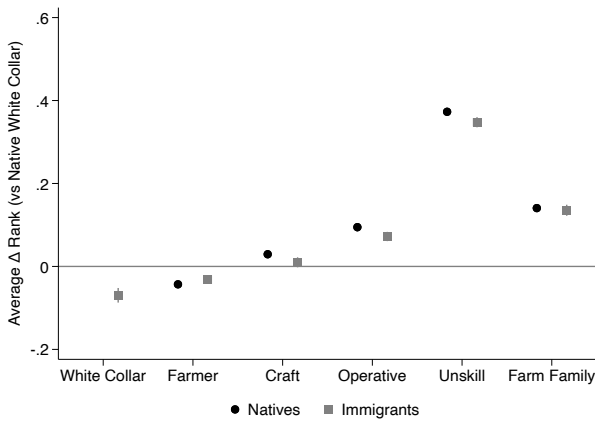


Figure H.6(d): 1900-1930

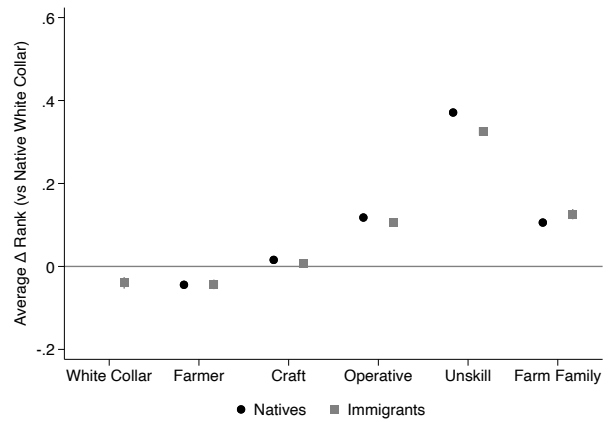
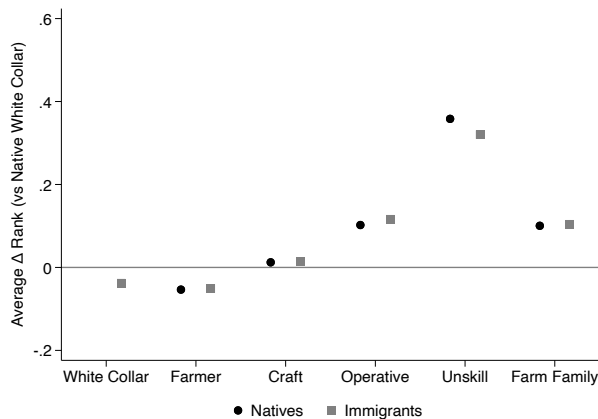
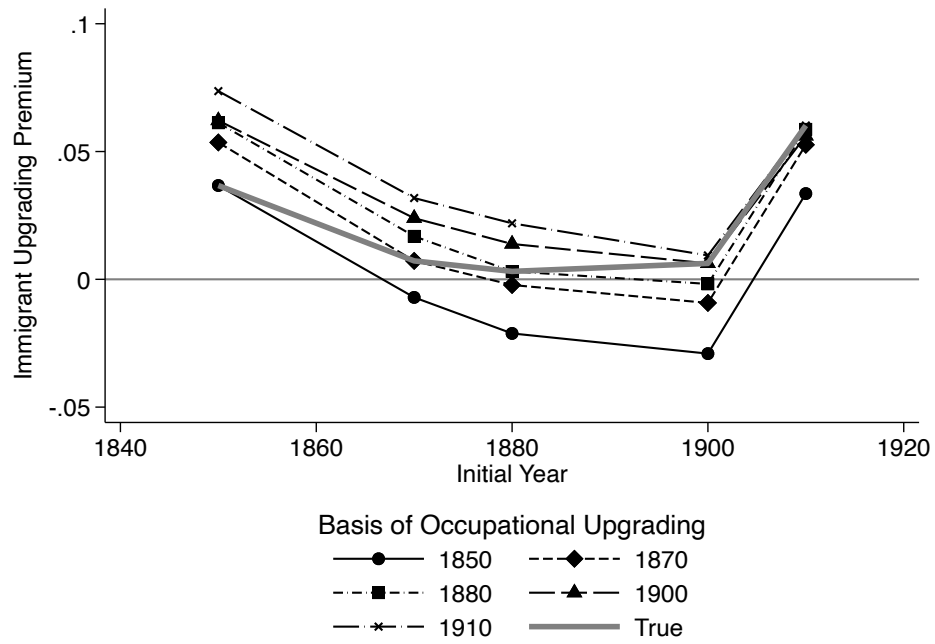


Figure H.6(e): 1910-1940



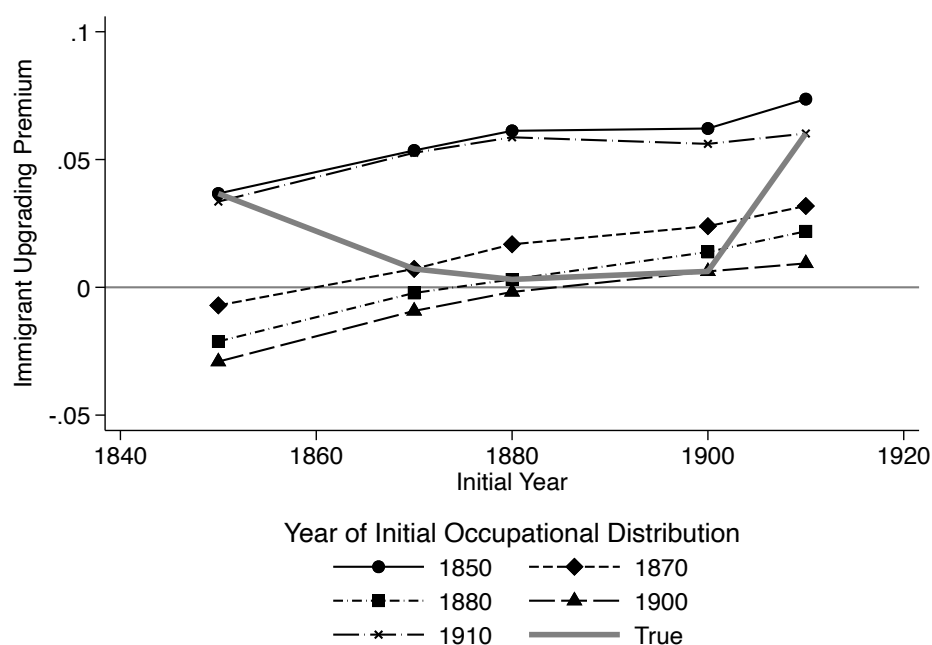
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure H.7: Unconditional assimilation, holding upgrading constant



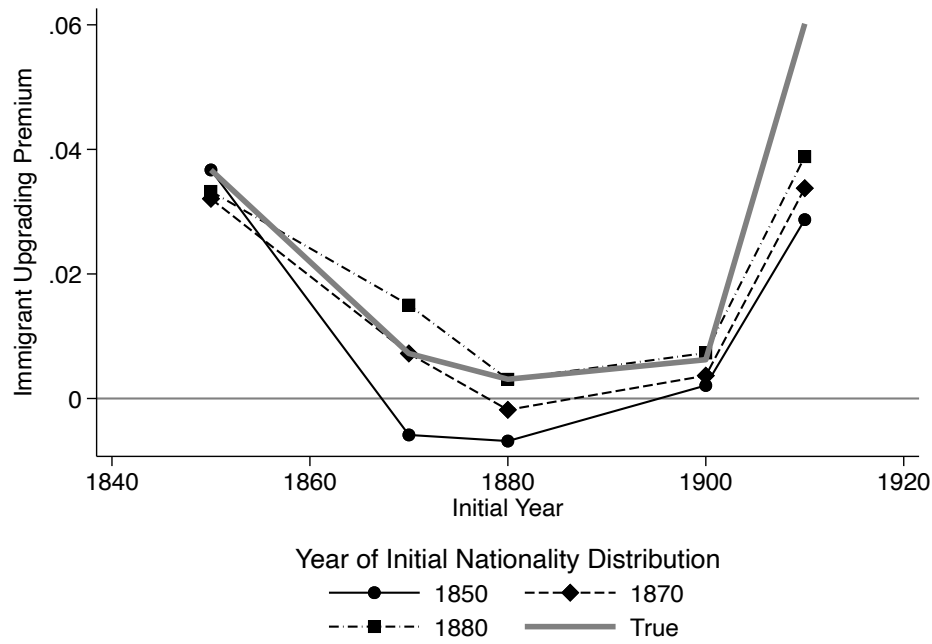
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure H.8: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure H.9: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Appendix I: Results omitting immigrants more than 10 years from arrival

This Appendix removes from consideration any immigrant who had been in the United States for more than 10 years when first observed—a restriction that can be imposed for samples beginning 1870 and later, but not for the 1850-1880 linked sample. This restriction is intended to address the possibility that, whereas immigrants observed in 1850 had likely arrived no earlier than 1840 due to the low rates of immigration prior to that year (which is why we are not concerned that we cannot implement this restriction for 1850-1880), those observed in later years might have been in the United States for some time and might have already experienced some assimilation before being observed. This concern would cause the *observed* assimilation of cohorts after 1850-1880 to be lower even if it was not truly lower. The results in this Appendix are qualitatively unchanged, however, indicating that this mechanism is not responsible for our results.

Figure I.1: Occupational distributions

Figure I.1(a): Initial Year 1850

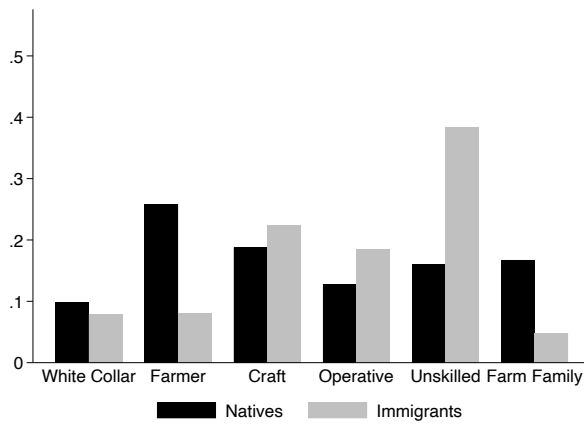


Figure I.1(b): Final Year 1880

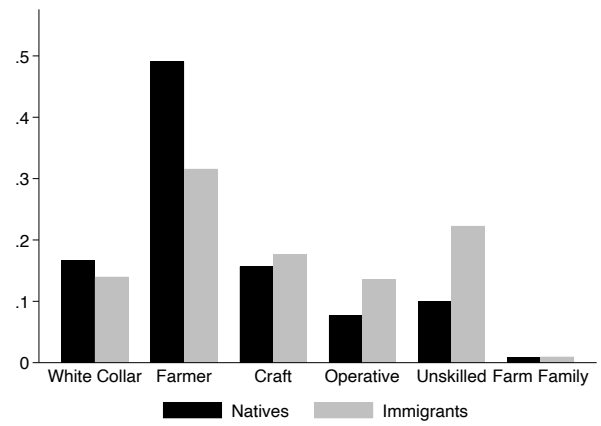


Figure I.1(c): Initial Year 1870

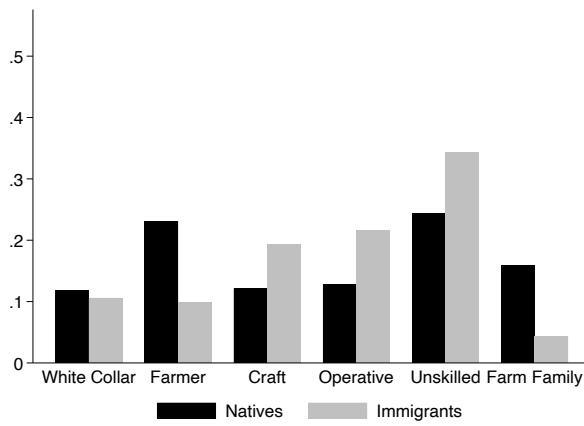


Figure I.1(d): Final Year 1900

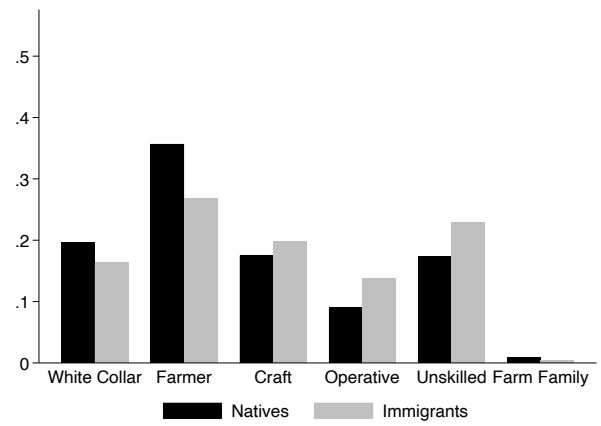


Figure I.1(e): Initial Year 1880

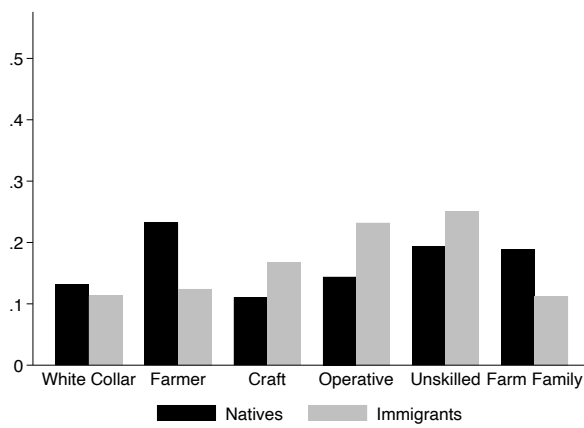


Figure I.1(f): Final Year 1910

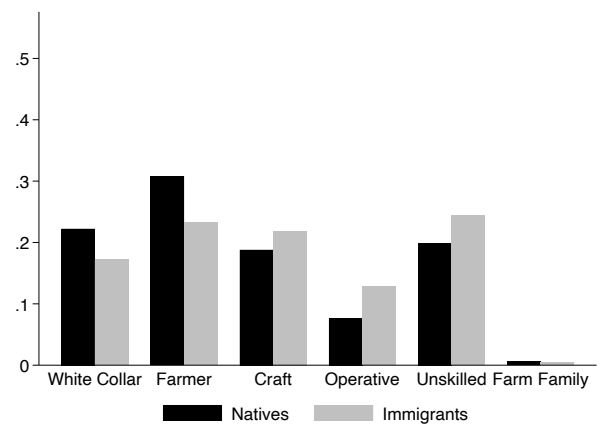


Figure I.1 (continued): Occupational distributions

Figure I.1(g): Initial Year 1900

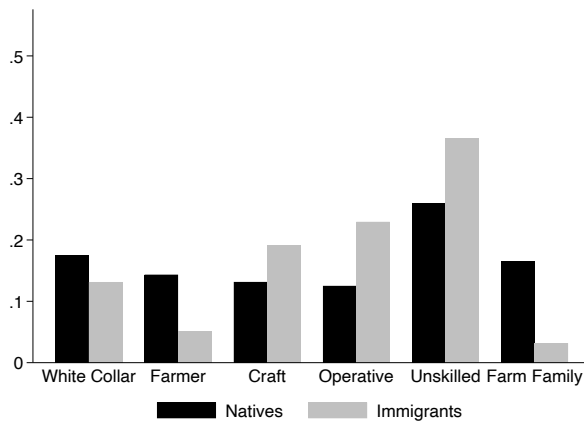


Figure I.1(h): Final Year 1930

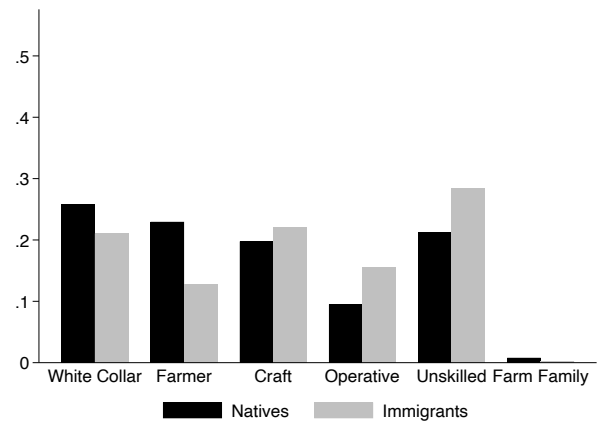


Figure I.1(i): Initial Year 1910

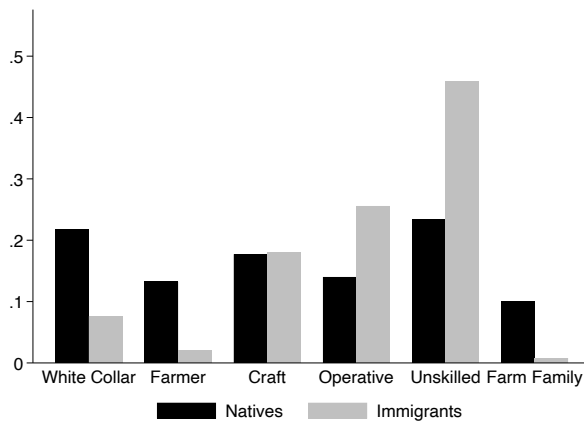
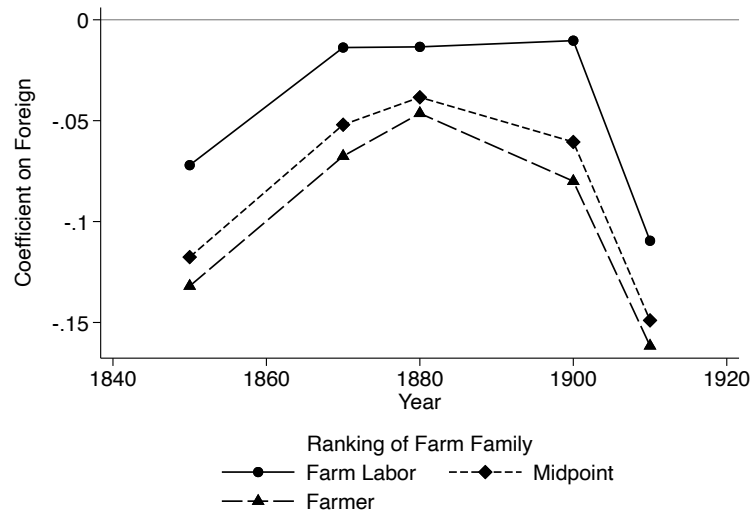


Figure I.1(j): Final Year 1940

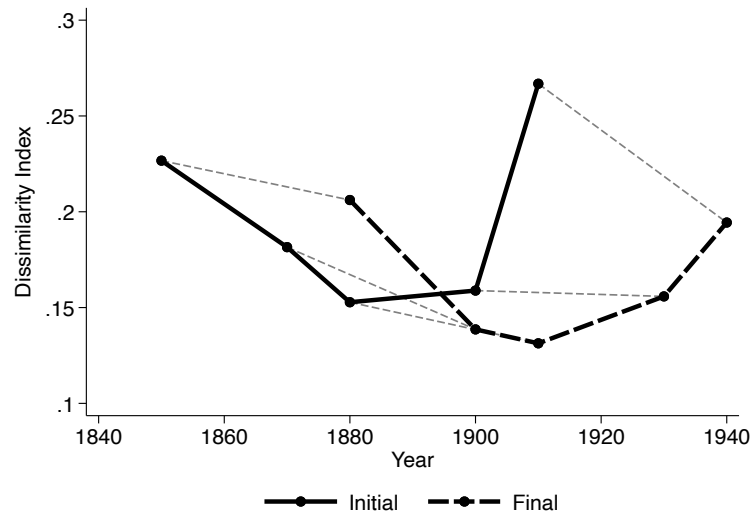
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure I.2: Initial gaps in occupational rank



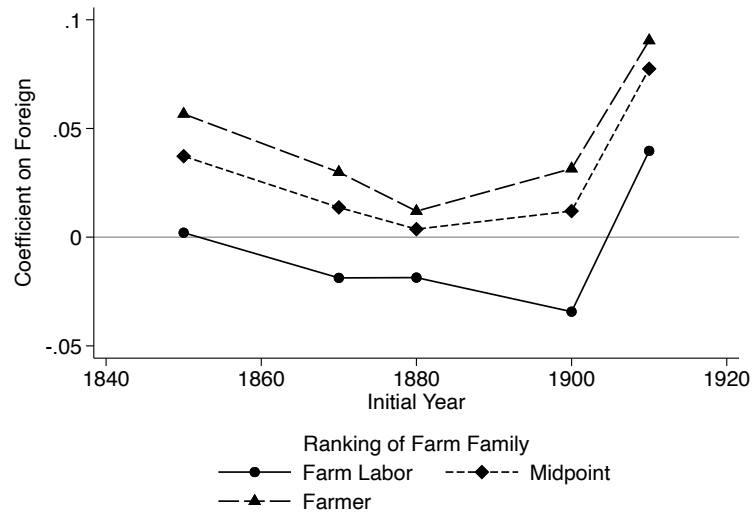
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure I.3: Dissimilarity indices



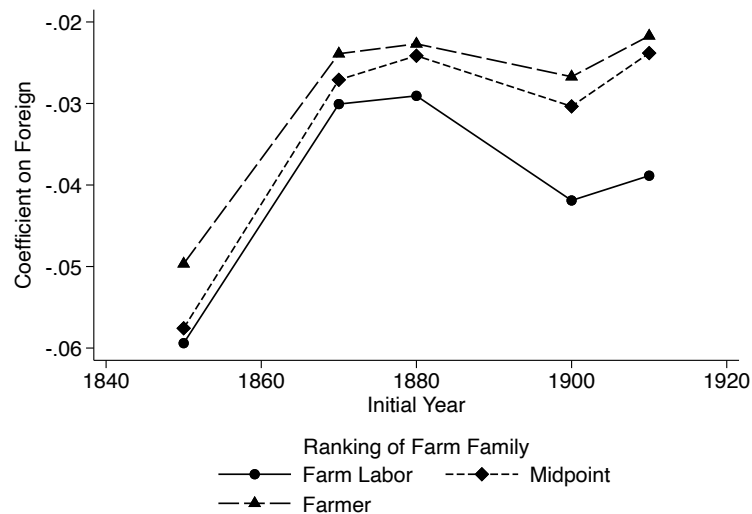
Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure I.4: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure I.5: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure I.6: Occupational upgrading by initial occupation

Figure I.6(a): 1850-1880

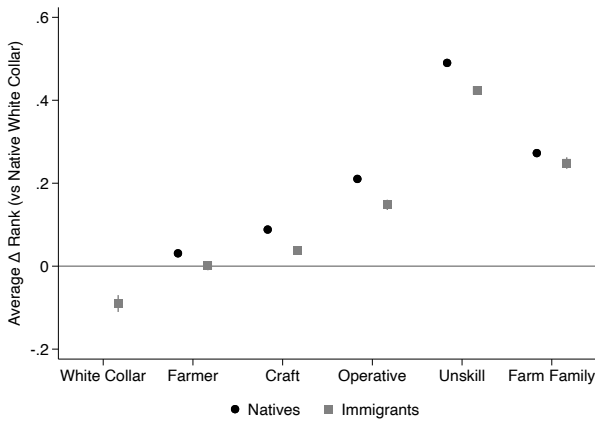


Figure I.6(b): 1870-1900

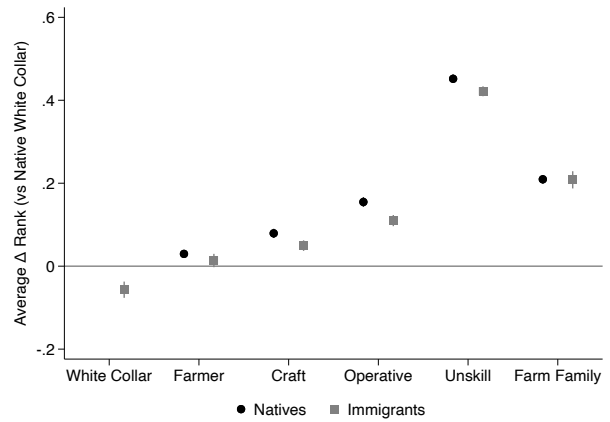


Figure I.6(c): 1880-1910

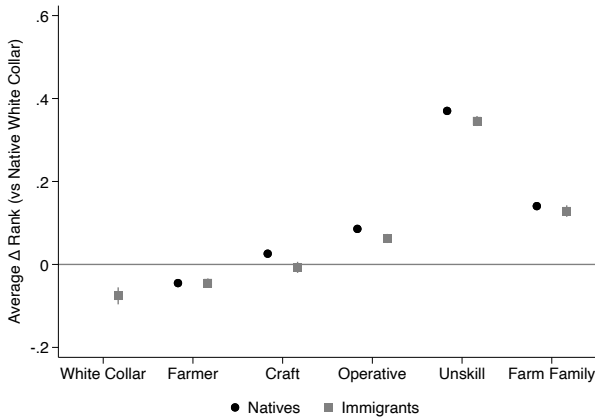


Figure I.6(d): 1900-1930

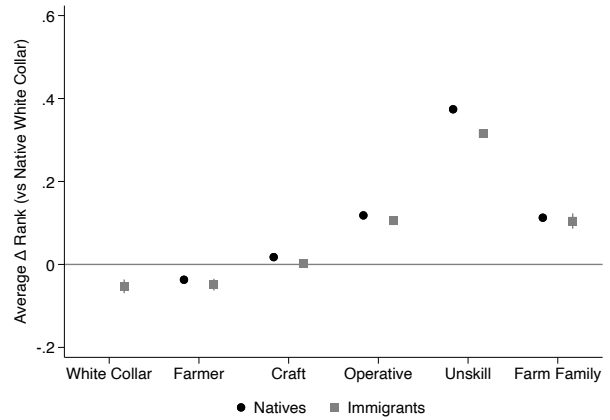
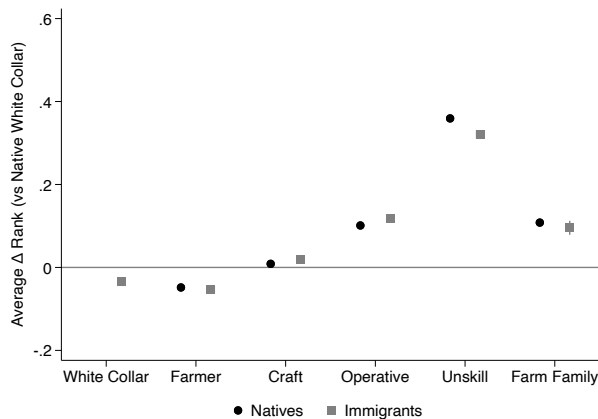
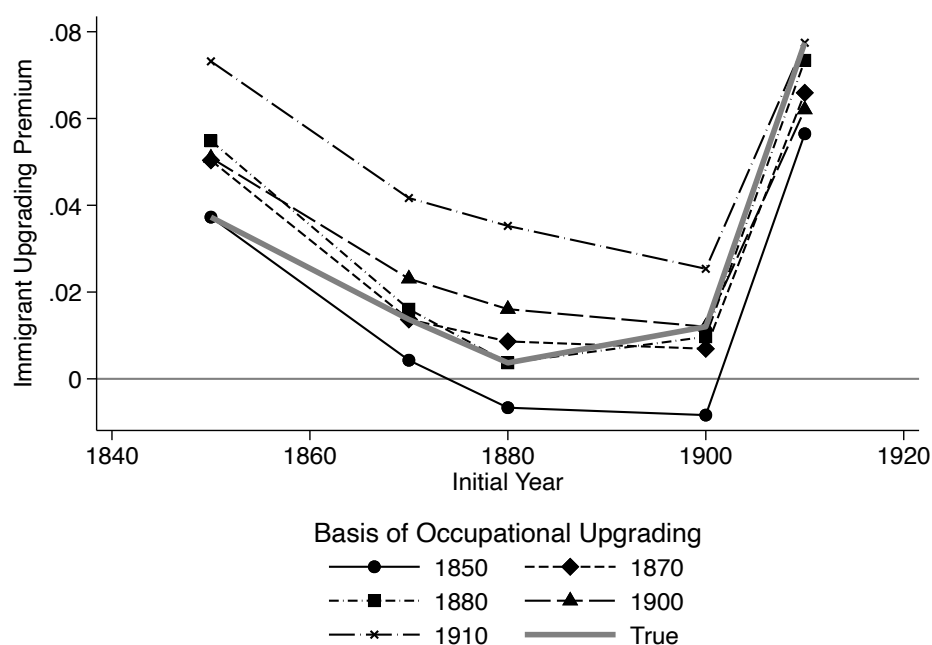


Figure I.6(e): 1910-1940



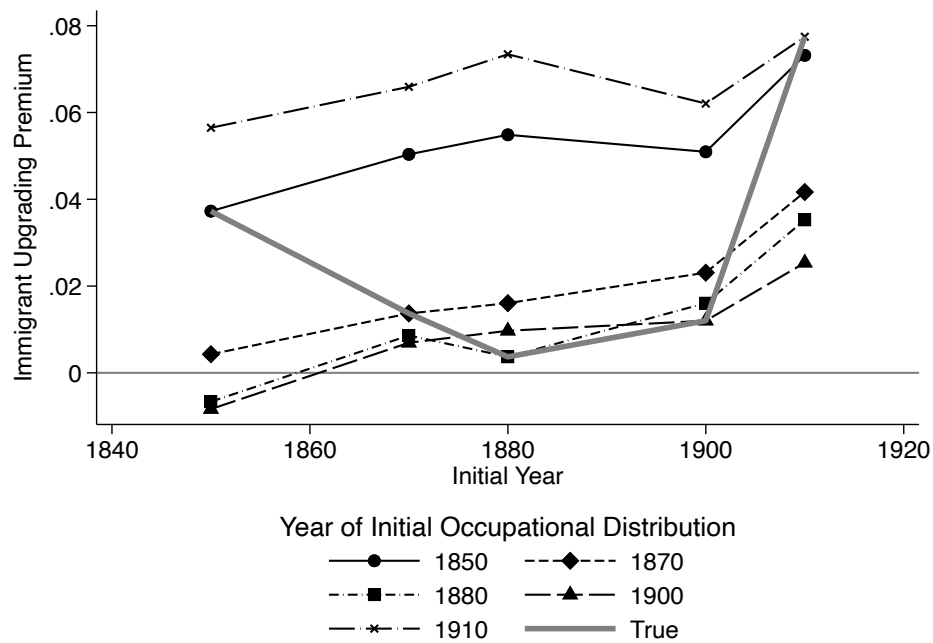
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure I.7: Unconditional assimilation, holding upgrading constant



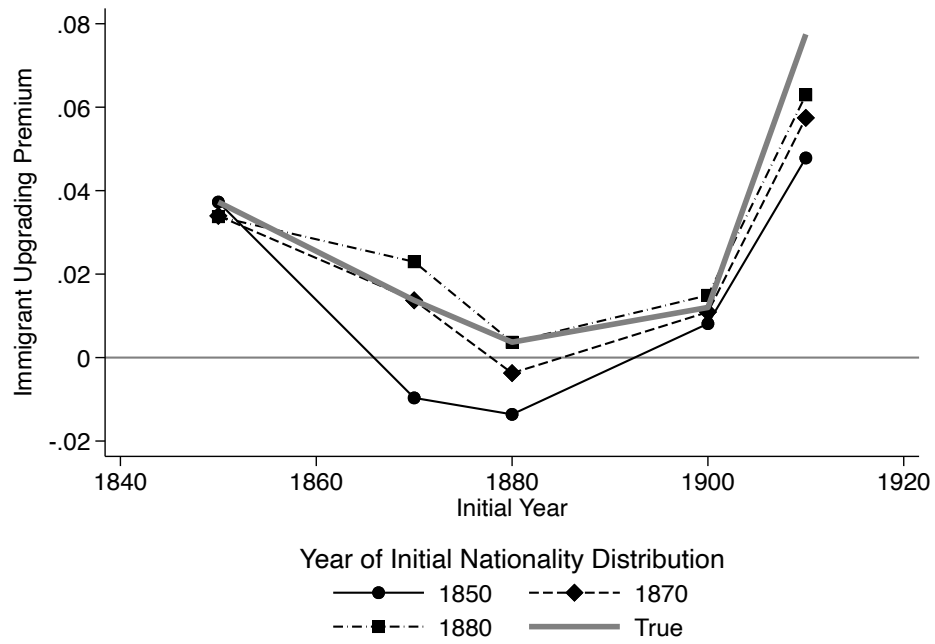
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure I.8: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

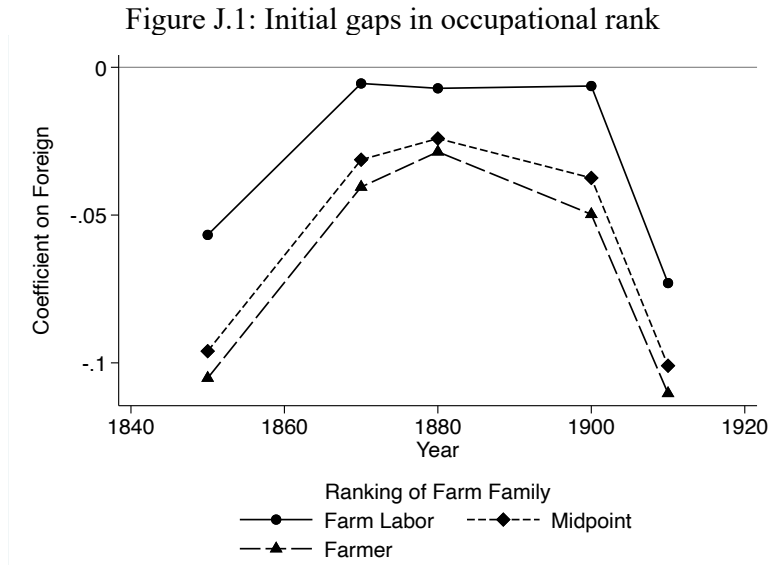
Figure I.9: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

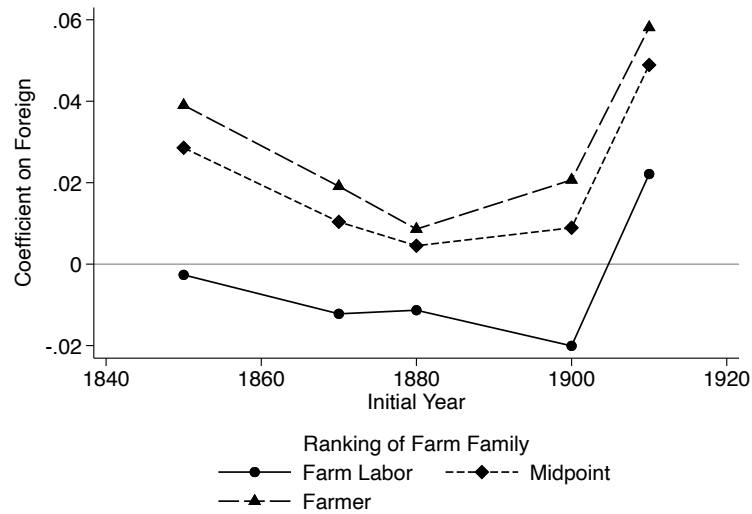
Appendix J: Results with alternate occupational wealth scores

This Appendix replaces the 1870 occupational wealth score used in the main text with two alternatives. Whereas the score in the main text is based on the average wealth holdings for all non-southern individuals aged 30-65 with a given occupation, this Appendix reports results with two alternatives. The first is a score that varies by age, with separate scores for each occupation for ages 18-30 and 44-64. This score enables us to capture, for instance, concerns that farmers might upgrade their status over the lifecycle without changing occupations. The second score varies by state, capturing geographic variation in status within occupations. In both cases, the results are qualitatively unchanged relative to the results in the main text.



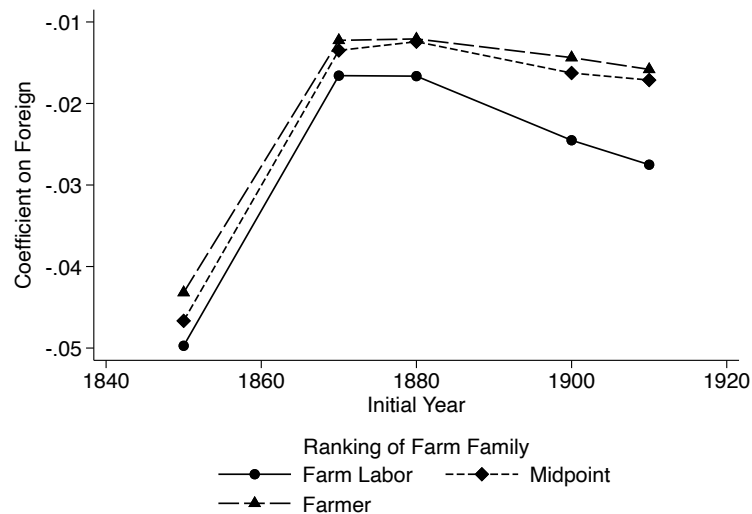
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure J.2: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure J.3: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure J.4: Occupational upgrading by initial occupation

Figure J.4(a): 1850-1880

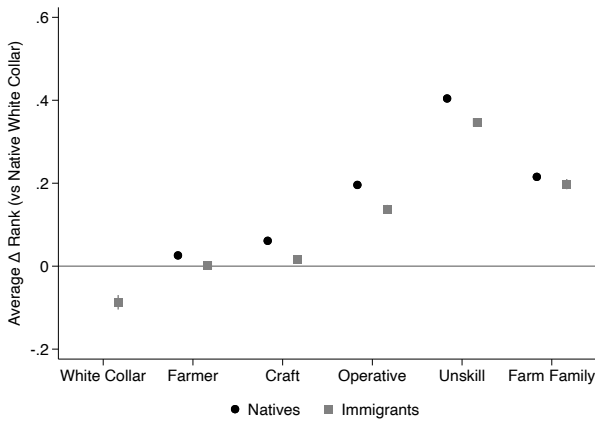


Figure J.4(b): 1870-1900

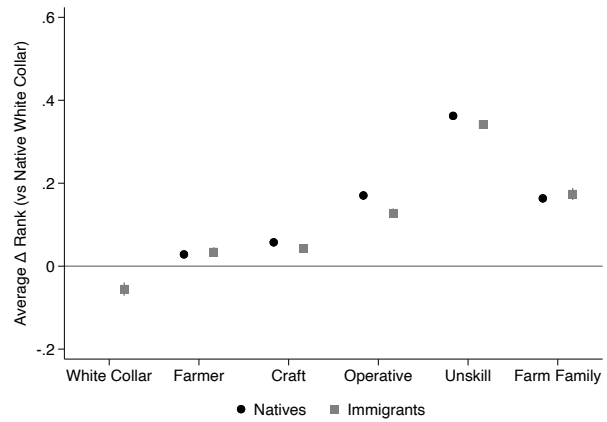


Figure J.4(c): 1880-1910

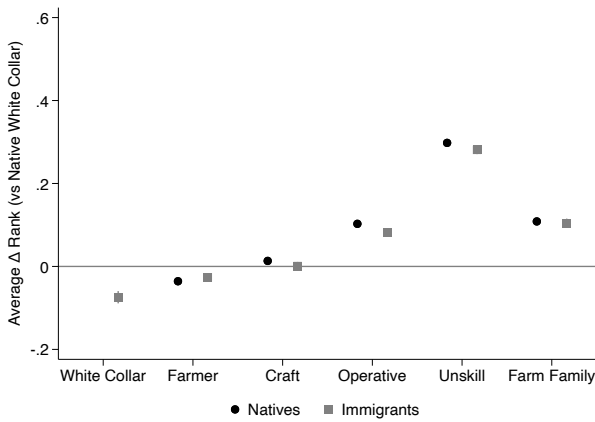


Figure J.4(d): 1900-1930

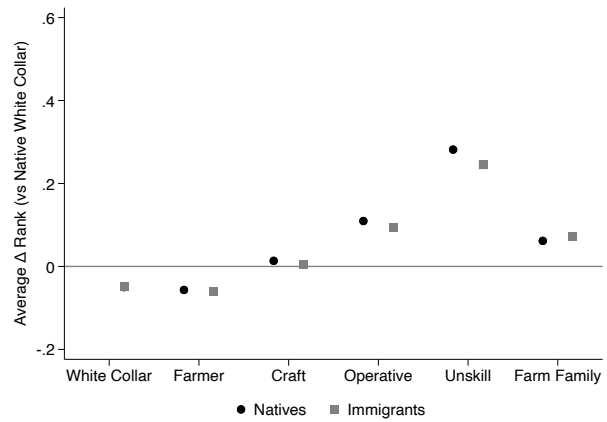
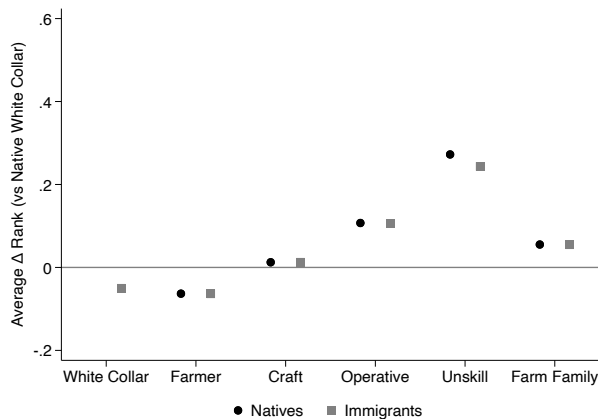
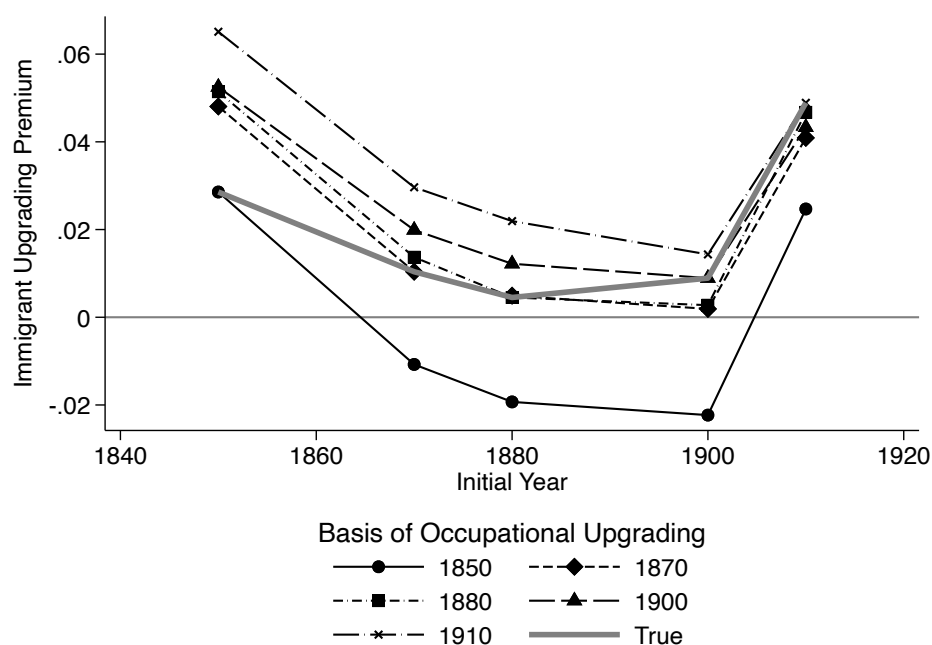


Figure J.4(e): 1910-1940



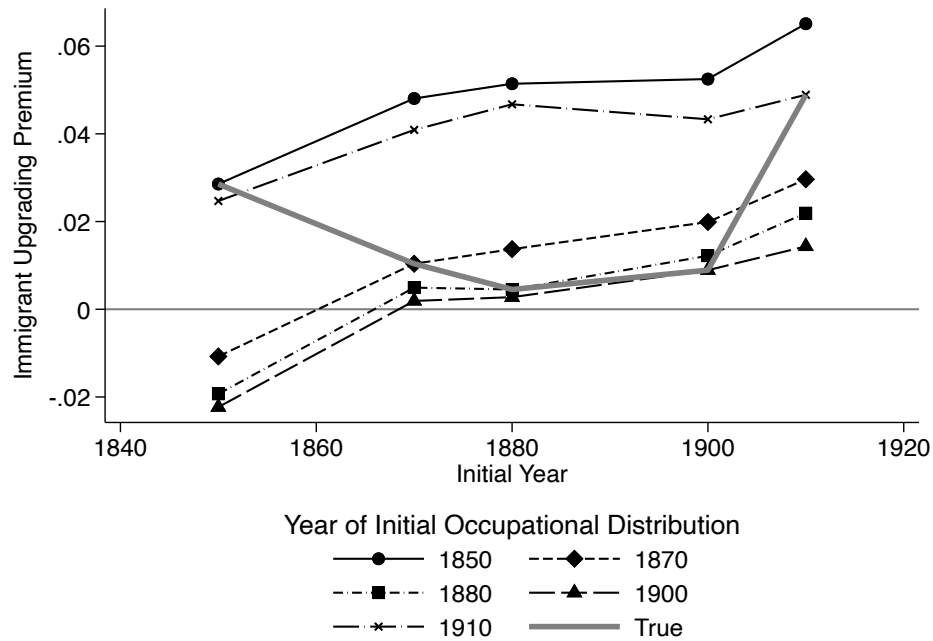
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure J.5: Unconditional assimilation, holding upgrading constant



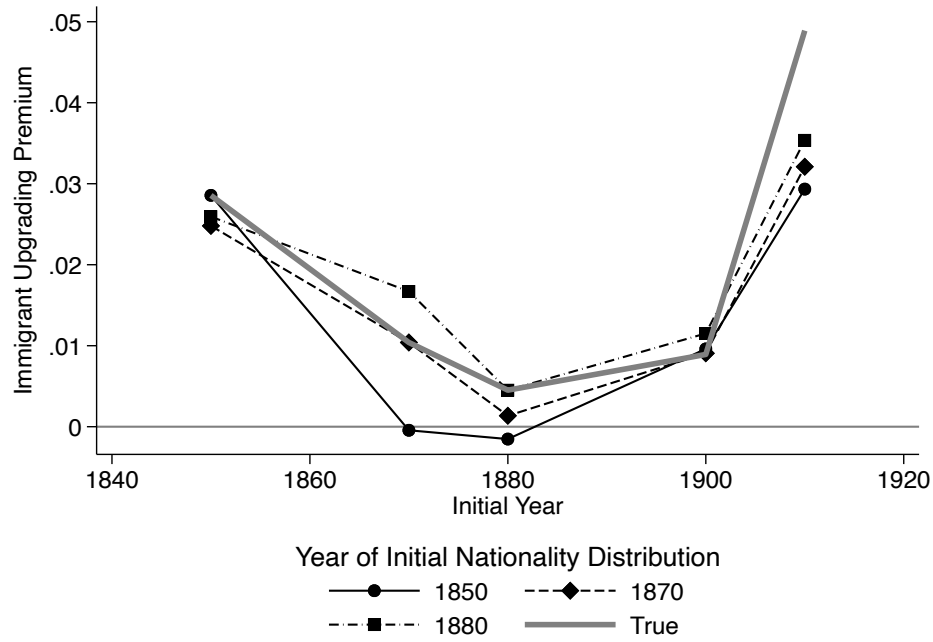
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure J.6: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure J.7: Unconditional assimilation, holding nationality distributions constant



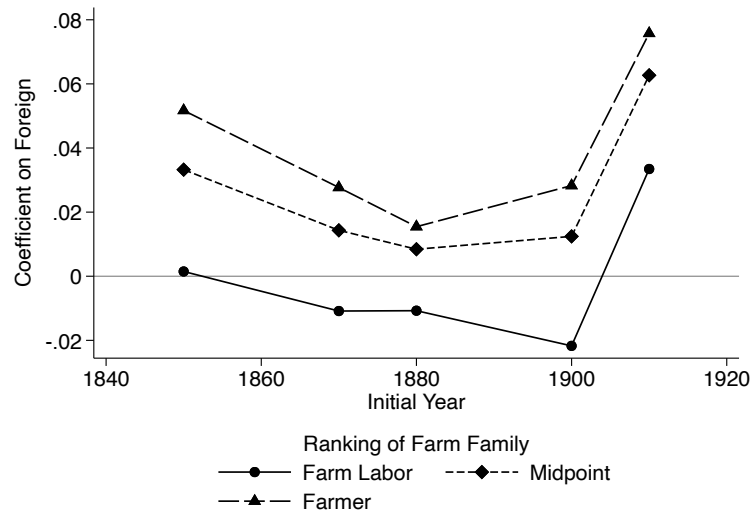
Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure J.8: Initial gaps in occupational rank



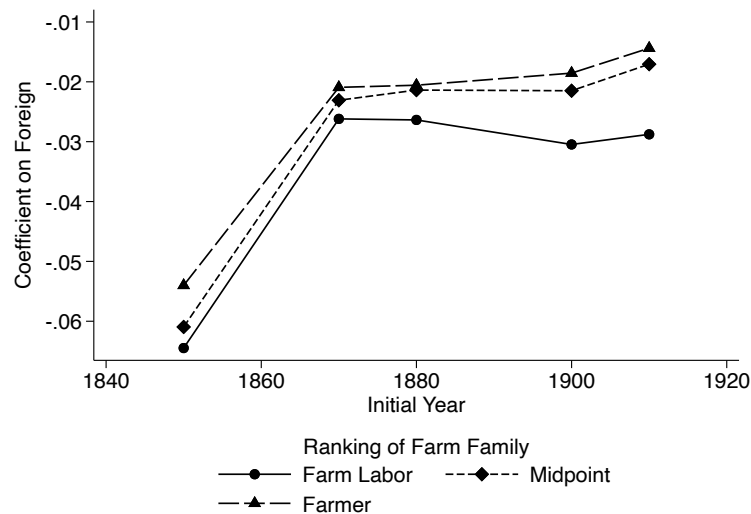
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure J.9: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure J.10: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure J.11: Occupational upgrading by initial occupation

Figure J.11(a): 1850-1880

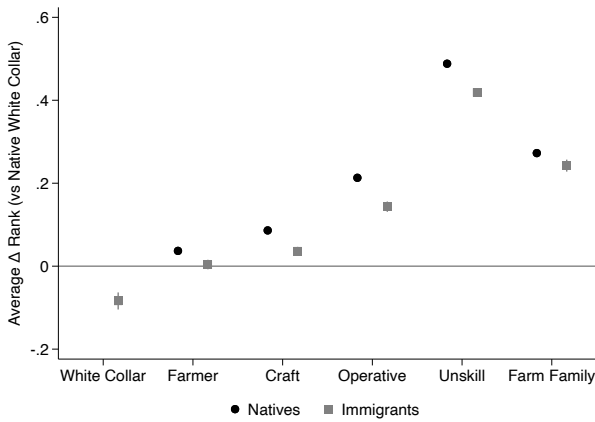


Figure J.11(b): 1870-1900

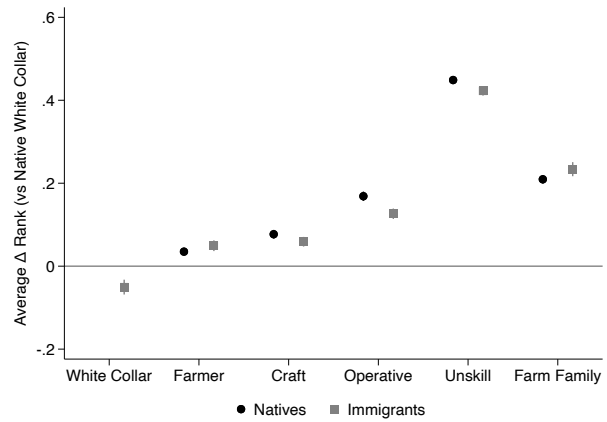


Figure J.11(c): 1880-1910

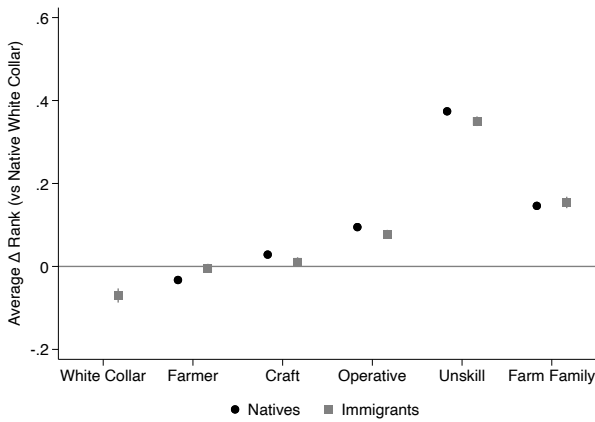


Figure J.11(d): 1900-1930

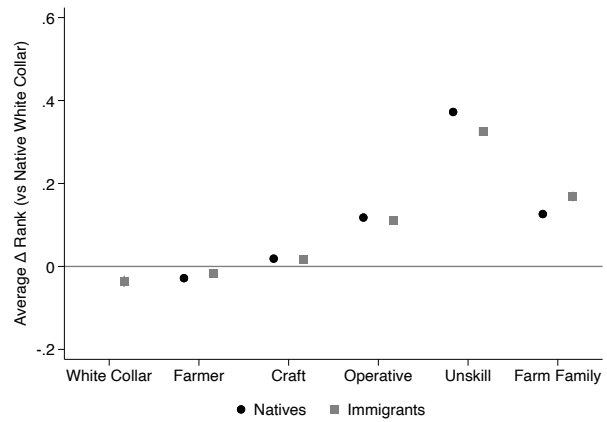
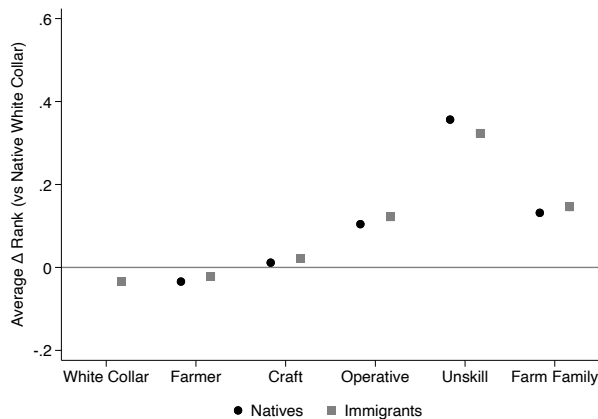
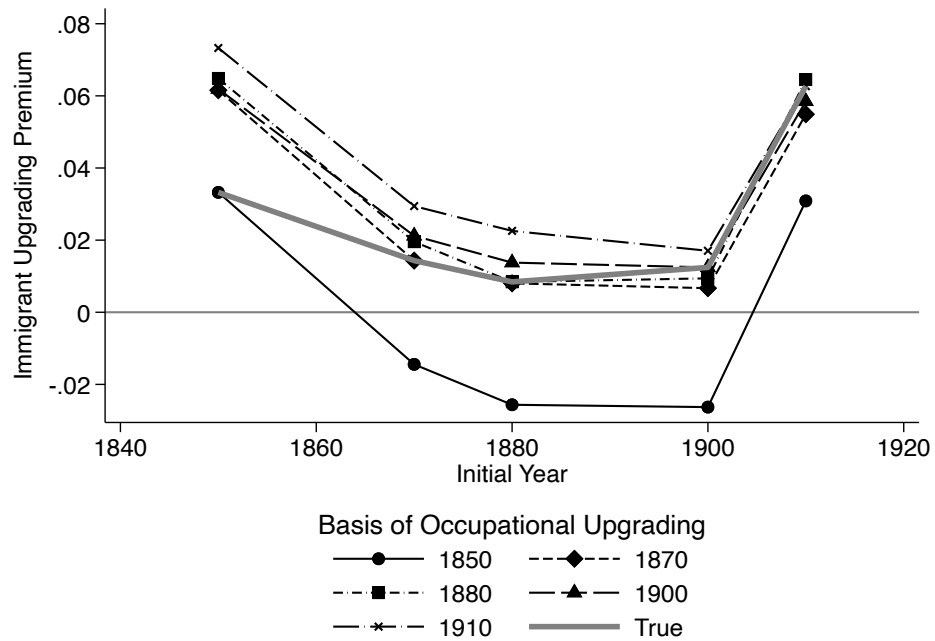


Figure J.11(e): 1910-1940



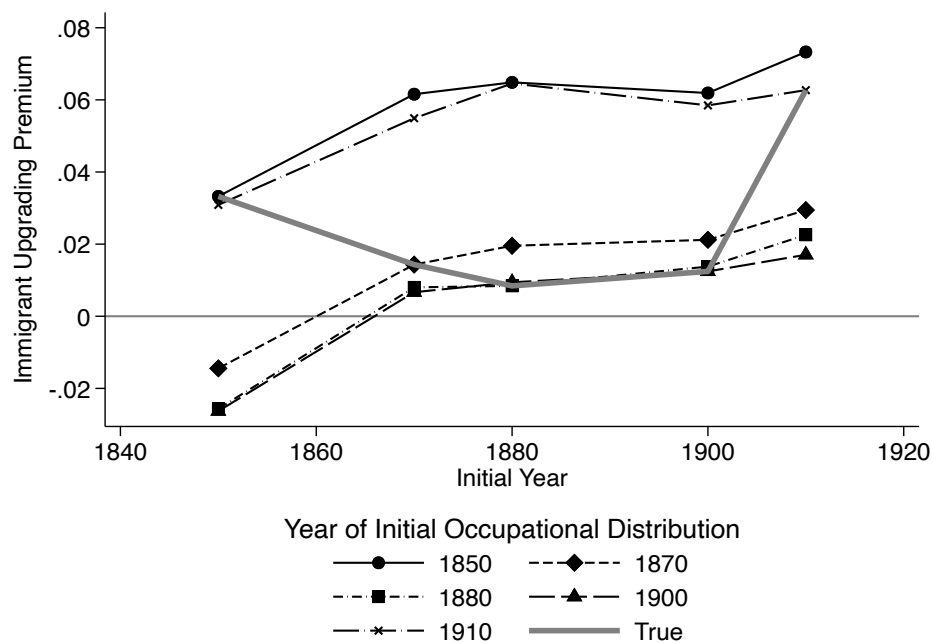
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure J.12: Unconditional assimilation, holding upgrading constant



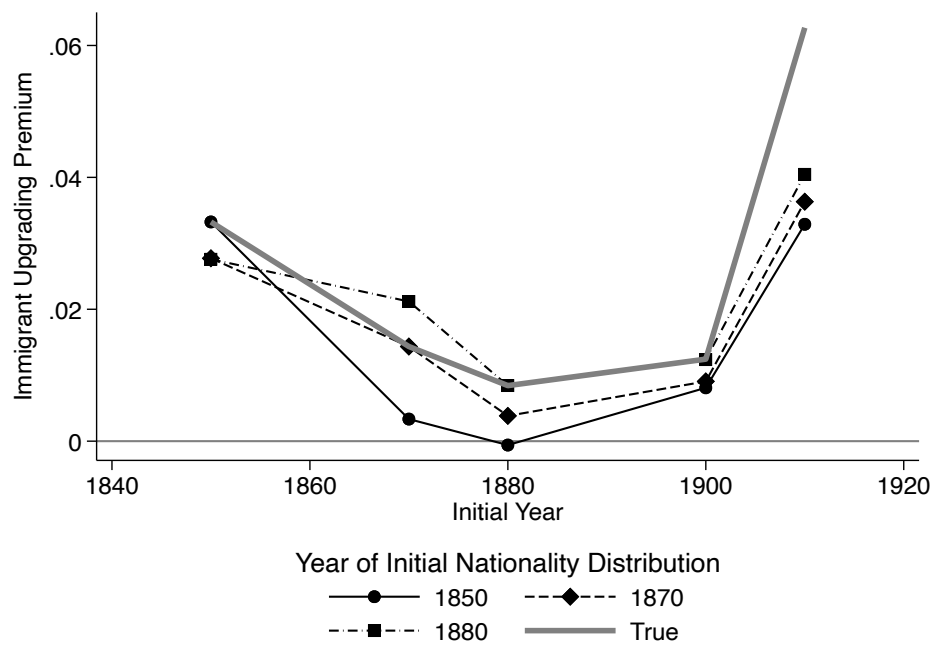
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure J.13: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure J.14: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Appendix K: Results limited to English speakers

This Appendix limits attention to individuals who report ability to speak English, which is a question solicited in the 1900-1930 censuses. The goal of this analysis is to remove individuals from the sample who might have had difficulty expressing their occupations to census enumerators, and whose occupations might therefore have been miscoded, similar to concerns raised by Ward (2021). This restriction is particularly biting in 1910 when the size of the Italian, Russian, and Austro-Hungarian immigrant groups was particularly large, and has the effect of essentially removing the bottom of the immigrant occupational distribution from this sample. Nonetheless, our main result of a U-shaped assimilation pattern driven primarily by changing initial occupational distribution remains robust.

Figure K.1: Occupational distributions

Figure K.1(a): Initial Year 1850

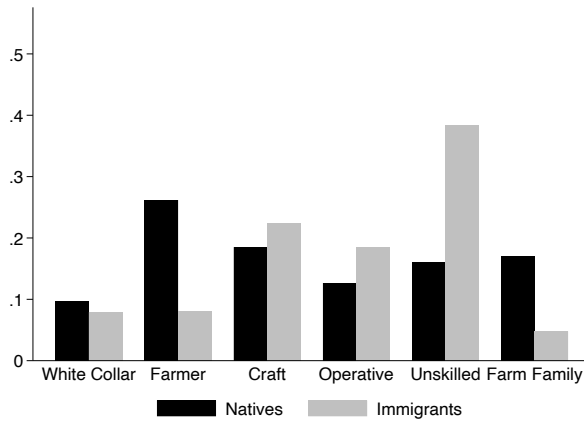


Figure K.1(b): Final Year 1880

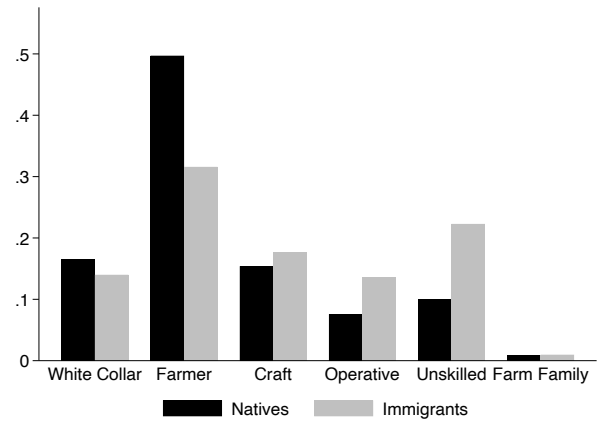


Figure K.1(c): Initial Year 1870

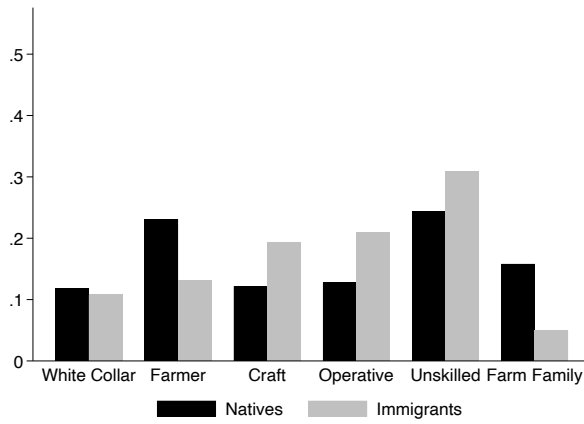


Figure K.1(d): Final Year 1900

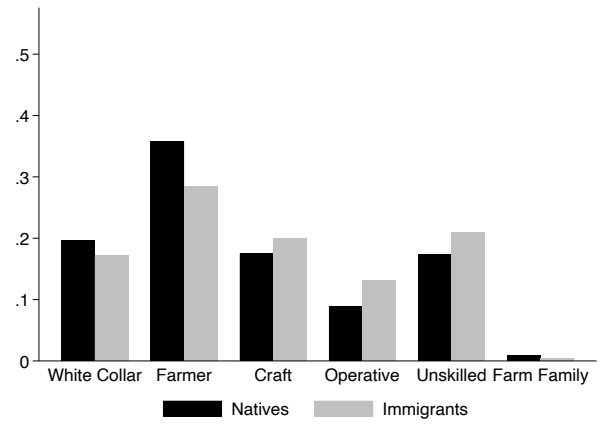


Figure K.1(e): Initial Year 1880

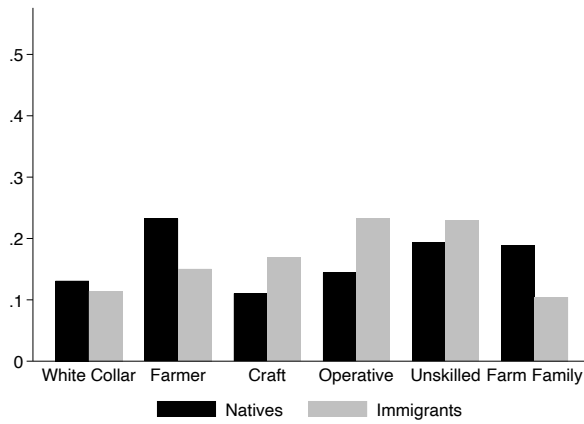


Figure K.1(f): Final Year 1910

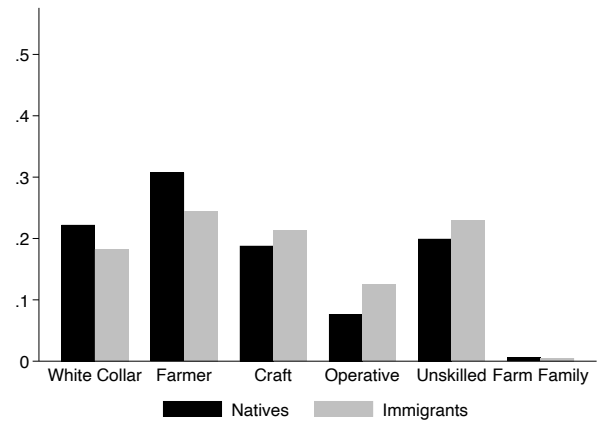


Figure K.1 (continued): Occupational distributions

Figure K.1(g): Initial Year 1900

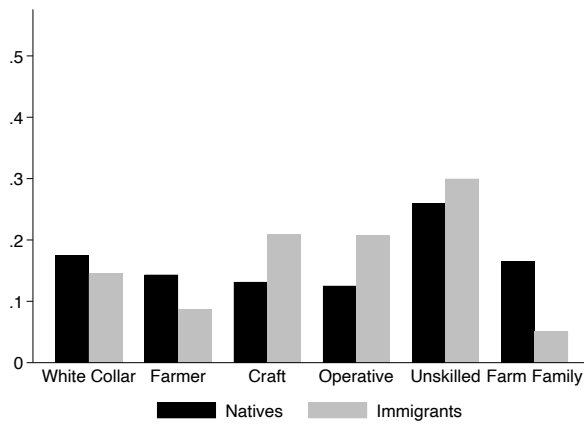


Figure K.1(h): Final Year 1930

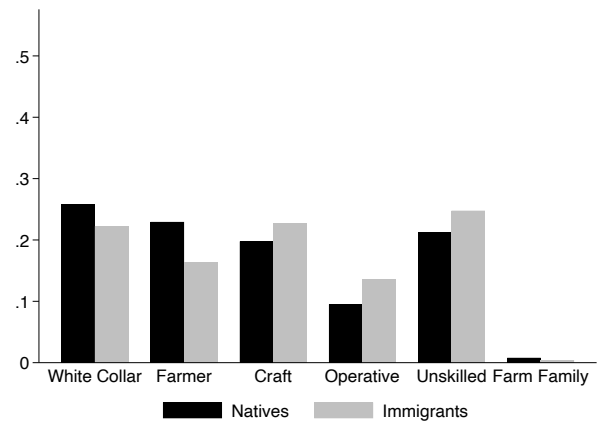


Figure K.1(i): Initial Year 1910

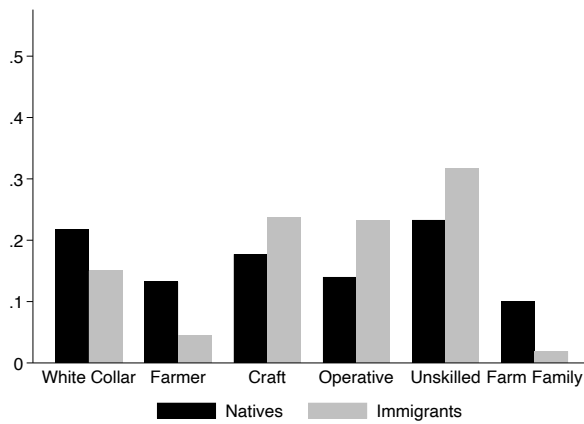
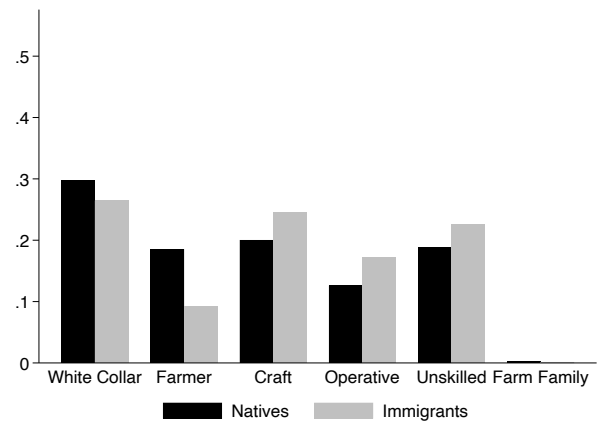


Figure K.1(j): Final Year 1940



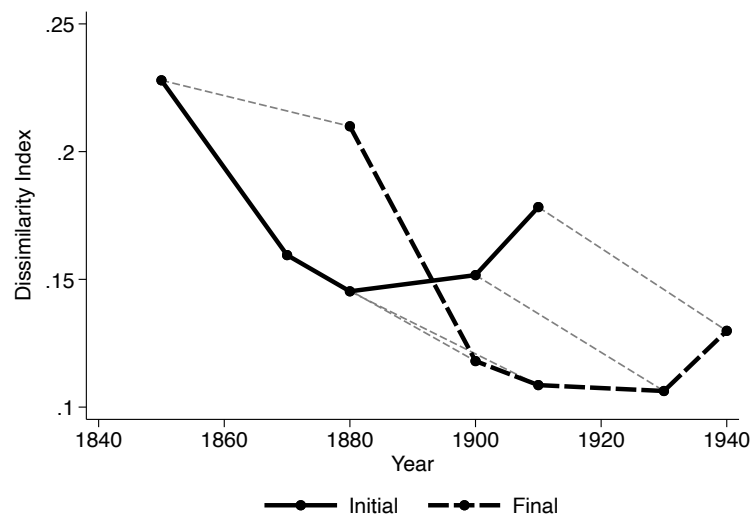
Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure K.2: Initial gaps in occupational rank



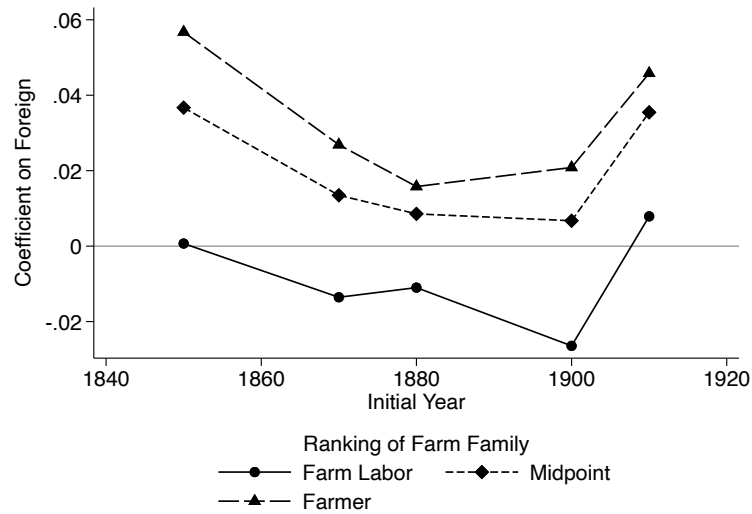
Notes: This figure presents the initial difference in the average occupational rank measure between immigrants and natives. The year on the x-axis is an initial year for one of our five linkage spans. The three lines each represent one potential ranking for farm family members. A negative coefficient corresponds to immigrants ranking behind natives. Observations corrected for selection into linkage.

Figure K.3: Dissimilarity indices



Notes: This figure presents age-adjusted dissimilarity indices for all initial and final years of linkage spans. The thin dashed lines linking the “initial” and “final” lines link the initial year of a linkage span to its final year. Observations weighted to correct for selection into linkage.

Figure K.4: Unconditional assimilation



Notes: This figure presents unconditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure K.5: Conditional assimilation



Notes: This figure presents conditional immigrant assimilation in terms of the average rank measure. The year on the x -axis represents the initial year of the linkage span. The three series represent the results for each of our three rankings of farm family members. Observations weighted to correct for selection into linkage.

Figure K.6: Occupational upgrading by initial occupation

Figure K.6(a): 1850-1880

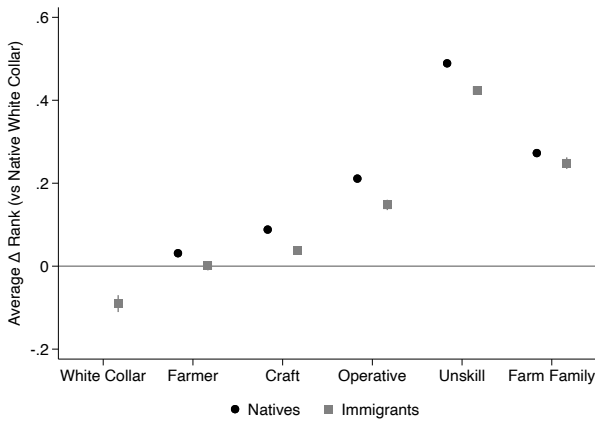


Figure K.6(b): 1870-1900

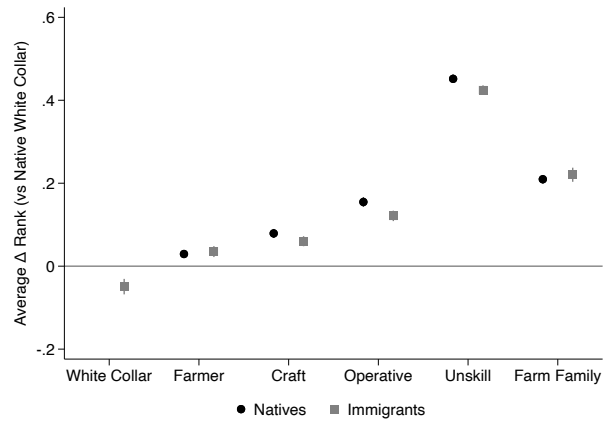


Figure K.6(c): 1880-1910

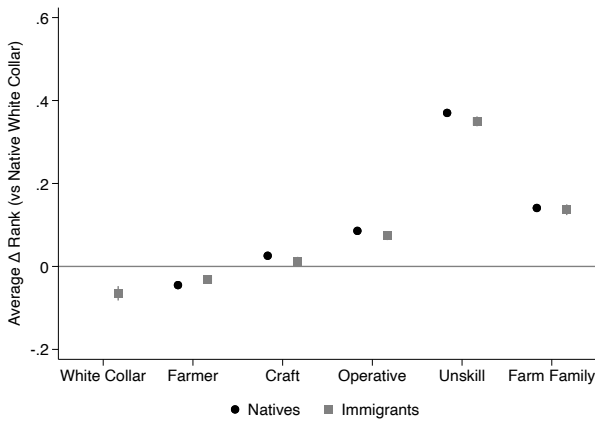


Figure K.6(d): 1900-1930

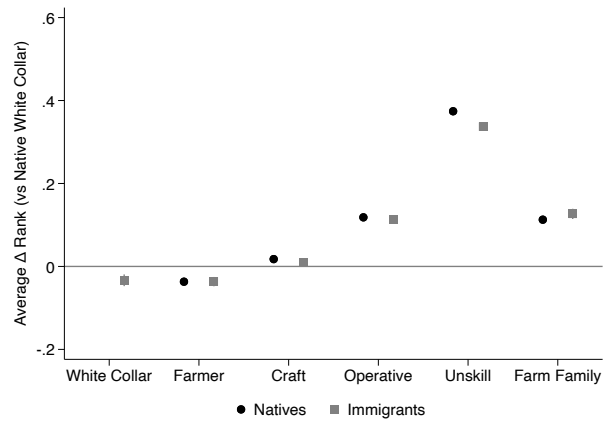
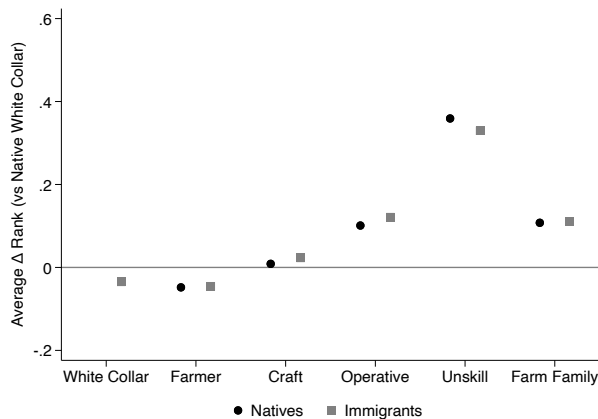
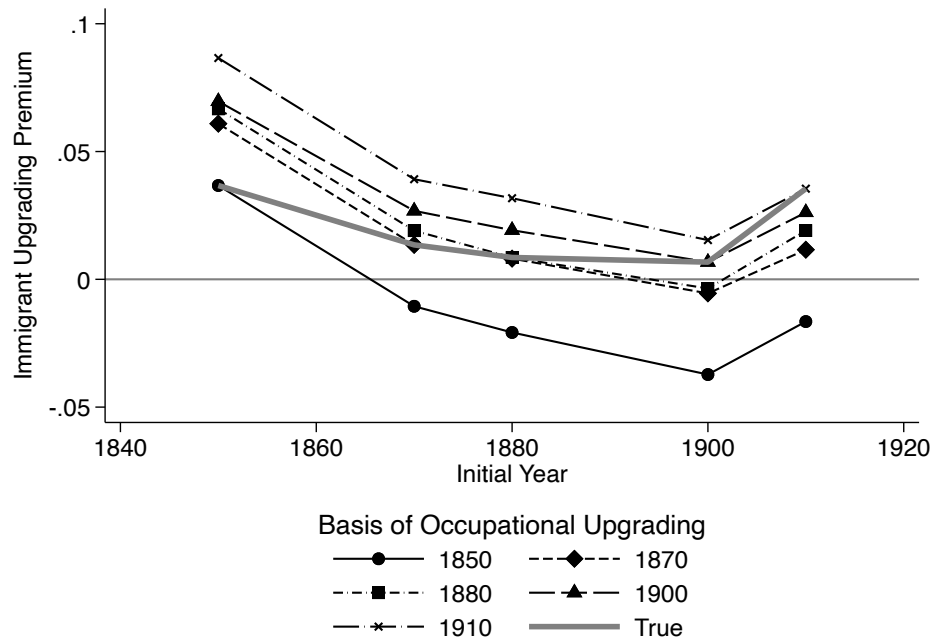


Figure K.6(e): 1910-1940



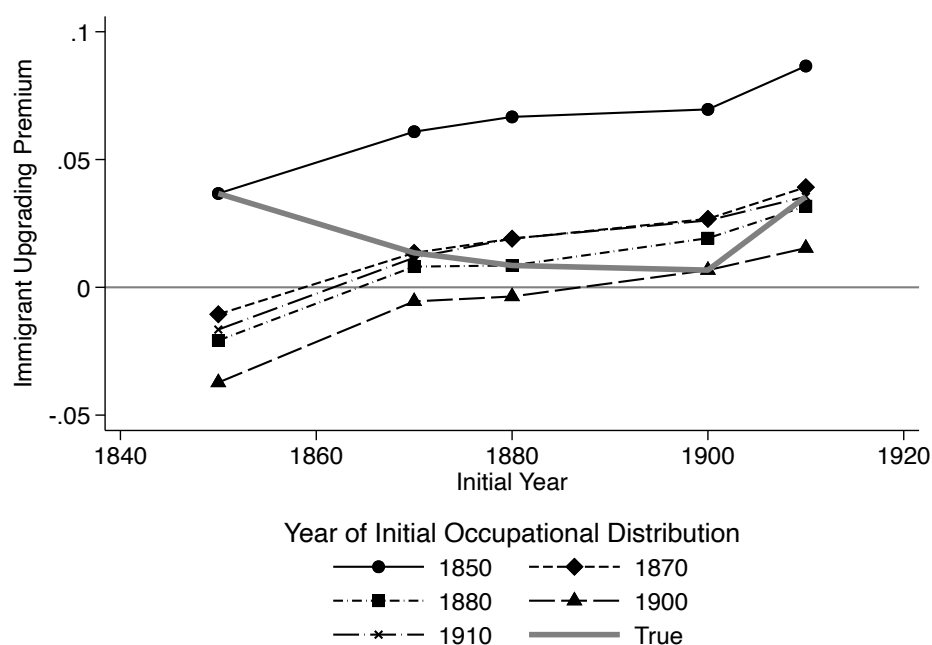
Notes: These graphs express the average upgrading experienced by immigrants or natives in each initial occupational category, expressed relative to natives initially in white collar occupations. Farm family members are ranked at the midpoint of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure K.7: Unconditional assimilation, holding upgrading constant



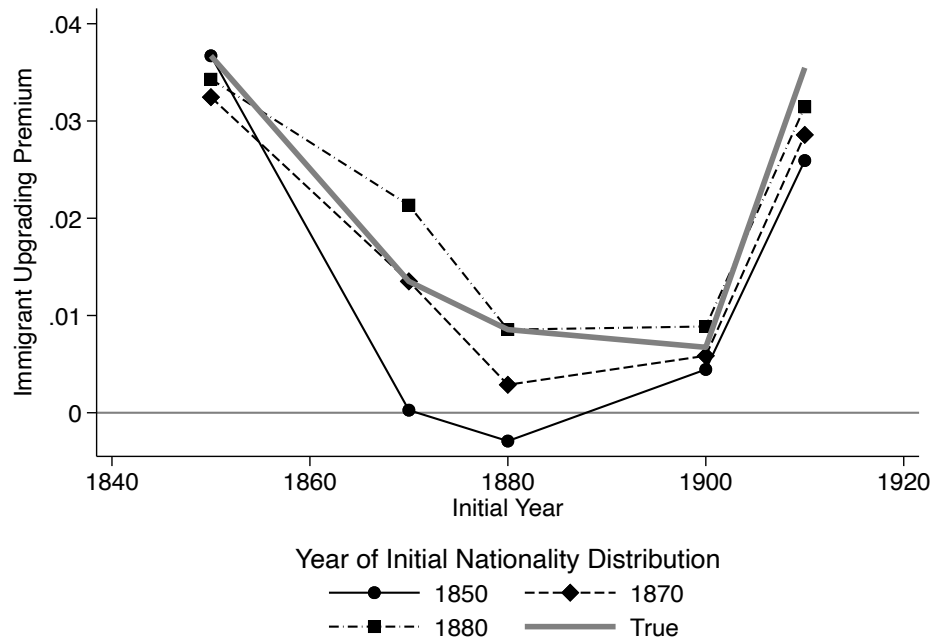
Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational upgrading, allowing the occupational weights to change over time. The legend indicates which initial year's occupational upgrading is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure K.8: Unconditional assimilation, holding initial occupational distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed occupational weights, allowing the occupational upgrading to change over time. The legend indicates which initial year's occupational weights are used. The x-axis indicates the initial year of the 30-year span from which the initial occupation-specific upgrading is taken. The y-axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Figure K.9: Unconditional assimilation, holding nationality distributions constant



Notes: Each line shows the occupational upgrading that would have occurred with fixed nationality distributions, allowing the occupational weights and occupational upgrading to change over time. The legend indicates which initial year's nationality distribution is used. The x -axis indicates the initial year of the 30-year span from which the initial occupation weights and initial occupation-specific upgrading are taken. The y -axis shows the difference in occupational upgrading between immigrants and natives over the period; positive numbers indicate that foreigners upgraded more than natives. Farm family members are ranked at the midpoint of the rankings of farmers and farm laborers. Observations weighted to correct for selection into linkage.

Appendix L: References

- Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James J. Feigenbaum, and Santiago Pérez (2021a). “Automated Linking of Historical Data.” *Journal of Economic Literature* 59:3, pp. 865-918.
- Abramitzky, Ran, Leah Platt Boustan, Elisa Jácome, and Santiago Pérez (2021b). “Intergenerational Mobility of Immigrants in the US over Two Centuries.” *American Economic Review* 111:2, pp. 580-608.
- Abramitzky, Ran, Leah Boustan, Katherine Eriksson, Santiago Pérez, and Myera Rashid (2020). *Census Linking Project: Version 2.0* [Machine-readable database]. <http://censuslinkingproject.org> (accessed December 16, 2021).
- Collins, William J. and Marianne H. Wanamaker (2021). “African American Intergenerational Mobility since 1880.” *American Economic Journal: Applied Economics*, Forthcoming.
- Collins, William J. and Ariell Zimran (2019). “The Economic Assimilation of Irish Famine Migrants to the United States.” *Explorations in Economic History* 74, 101302.
- Gershsteyn, Yefim (2000). “Use of SPEDIS Function in Finding Specific Values.” Proceedings of the Twenty-fifth Annual SAS Users Group International Conference, Indianapolis.
- Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek (2020). *IPUMS USA: Version 10.0* [Machine-readable database]. Minneapolis: University of Minnesota.
- Ward, Zachary (2021). “Intergenerational Mobility in American History: Accounting for Race and Measurement Error.” NBER Working Paper 29256.
- Zimran, Ariell (2021). “Immigrant Distribution in the United States during the Age of Mass Migration.” NBER Working Paper 28812.