

The Effects of Immigration Restriction Laws on Immigrant Segregation in the Early Twentieth Century U.S.*

Dafeng Xu[†]

Abstract

Several immigration restriction laws, passed in the background of anti-immigration populism in the early 1920s, put an end to the period of “open borders” in the U.S. Did immigration restriction laws affect immigrant segregation in the U.S.? In this paper, I use the linked county and individual sample constructed based on decennial censuses in the early twentieth century to examine how immigration restriction laws led to different trends in segregation across groups. Results shows that all immigrant groups followed similar trends in segregation before the passage of immigration restriction laws; afterwards, segregation among immigrants from more restricted countries declined relative to segregation among immigrants from less restricted countries. Such differences in segregation patterns were largely related to restrictions on new immigration from more restricted countries; selection on return migration could also partially explain the results. Further analyses suggest that de-segregation might not benefit immigrants’ assimilation.

Keywords: immigration restriction; law; immigrant segregation; early twentieth century; U.S.

*I thank Ran Abramitzky, Martha Bailey, Leah Boustan, Giovanni Facchini, Janna Johnson, Jason Kerwin, Daniel Lichter, Matt Nelson, Matthew Notowidigdo, Sam Peltzman, Margaret Peters, Hillel Rapoport, Evan Roberts, Steven Ruggles, Matthew Sobek, David Van Riper, Hans-Joachim Voth, Zachary Ward, John Robert Warren, and Ariell Zimran for their comments and suggestions. I also receive helpful comments from participants at numerous seminars and conferences.

[†]225 19th Ave. S., University of Minnesota, Minneapolis, MN 55455. Email: dafengxu@umn.edu.

1 Introduction

In the early twentieth century, the U.S. passed several immigration restriction laws in the background of anti-immigration populism, putting severe limitations on immigration from Europe, Asia, and Africa. Following the earlier Emergency Quota Act of 1921, the Immigration Act of 1924 marked the end of the “open borders period” in the U.S. history (Ngai, 1999, 2004). A quota system was established and the number of newcomers could not exceed the quota for each sending country. While quotas were assigned to all European countries, the effects of quota systems were heterogeneous. In particular, very low quotas were assigned to Southern and Eastern European countries. Using five U.S. decennial censuses from 1900 to 1940 (Ruggles et al., 2017), I link U.S. counties and individual immigrants across censuses, and examine effects of immigration restriction laws on immigrant segregation at the county level. Specifically, I study how immigration restriction laws led to different trends in segregation between immigrant groups that were less or not restricted, and immigrant groups that were severely restricted.

Immigrant segregation can be dynamic over time: on one hand, immigrants spatially assimilate to the host society (Massey and Denton, 1985); on the other hand, many new immigrants prefer ethnic enclave residence upon arrival (Bartel, 1989). Moreover, the spatial and social behaviors of the native-born population could also affect immigrant segregation. In the historical context of this paper, urbanization accelerated in the U.S. since the late nineteenth century (Ferrie, 2003), which could affect immigrant segregation because (a) many immigrant enclaves were located in cities, and (b) immigrants who previously lived in rural areas also started to move to cities (Hall and Ruggles, 2004). In addition, the early twentieth century was the era of Great Black Migration, which was partially delayed by mass immigration (Collins, 1997), but nonetheless occurred by around 1920 (Collins and Wanamaker, 2015). African Americans affected immigrants’ social outcomes (e.g., Boyd and Xu, 2003; Fox and Guglielmo, 2012): in particular, internal migration of African Americans led to racial segregation in many cities (Shertzer et al., 2016), which could

further change the pattern of immigrant segregation. Finally, immigrants usually faced hostility and discrimination, especially after the passage of immigration restriction laws (e.g., Higham, 1963; Brown and Warner, 1992; Goldin, 1994; Ngai, 1999; Haines, 2000; Moser, 2012; Abramitzky et al., 2017), which could also affect the dynamics of immigrant segregation. In general, immigrant segregation in the U.S. declined until 1920 but then started to rise (Cutler et al., 2008).

Among seven major immigrant groups that had over one million people in 1920, the British, Canadian, Irish, and Scandinavian group were less or not restricted by immigration restriction laws,¹ while the Italian, Polish, and Russian group were severely restricted. On the other hand, other social factors (e.g., Great Black Migration) might affect all immigrant groups. I compare county-level immigrant segregation across groups, before and after the passage of immigration restriction laws. I find that all major immigrant groups experienced similar trends in segregation before the 1920s. After the 1920s, however, segregation rose among four major less restricted groups, while the degree of segregation among three major more restricted groups became relatively lower.

Note that it is more controversial to study Polish and Russian segregation due to statistical issues in pre-1930 censuses. For example, the number of people reporting Poland as the birthplace in the 1910 census was abnormally low, making the number of Russian-born immigrants also unreliable.² This might be due to political situations in Eastern Europe before World War I. In this paper, the main analysis relies on the (imperfect) solution to these statistical issues, i.e., using the mother tongue to define the Russian and Polish group when the birthplace information are less reliable. That said, one might still worry that the results

¹Although there were more than one million German-born immigrants in 1920, in this paper I do not study German segregation, as the effects of immigration restriction laws (and other social factors) on German segregation might be confounded by the effects of World War I, which should not affect segregation of other groups that were not considered to be enemies of the U.S. in the 1910s.

²Only 34,554 people in the 1910 census reported Poland as the birthplace (the number of Russian-born immigrants: 1,562,045). Although this might be because Poland gained independence only in 1918, the 1900 census did document 384,058 Polish-born immigrants (the number of Russian-born immigrants: 424,648). Hence, it is more likely that there were statistical errors in the 1910 census. In any of these censuses, many Polish immigrants might report Russia as the birthplace, as they were indeed born in the Russian part of “geographic Poland”. This could even affect the 1920 census, as 1920 was very close to Poland’s independence.

should be interpreted with caution given that Polish and Russian segregation are calculated not as cleanly as Italian segregation. To further explore this, I redo the empirical analysis based on the sample without Polish and Russian immigrants, and find qualitatively similar results. This suggests that the main conclusion of this paper is not driven or affected by the construction of the “treatment group” or data issues in U.S. censuses.

I then study the “structure” of county-level immigrant segregation by analyzing individuals’ ethnic enclave residence at the sub-county level. To do so, I construct an individual-level panel dataset by linking the 1920 immigrant census sample to the 1930 census, and examine the effects of immigration laws on the share of immigrants of the same origin in the enumeration district. I find that immigration restriction laws decreased ethnic enclave residence of immigrants from more restricted countries, which is consistent with county-level results of trends in immigrant segregation.

There are several possible mechanisms behind the effects of immigration laws on immigrant segregation. Researchers observe that immigrants move to less segregated areas as a consequence of assimilation (e.g., Duncan and Lieberson, 1959; Gordon, 1964; Massey and Denton, 1985; Bartel, 1989). Hence, immigration laws generated direct effects on segregation, as the number of new immigrants from more restricted countries significantly declined after the 1920s. I indeed find that differences in arrivals of new immigrants between more and less restricted groups could lead to differences in segregation patterns.

In addition, it is theoretically possible that immigration laws could indirectly affect immigrant segregation through spatial behaviors of old immigrants who had settled down in the U.S. I call this as the “indirect effect”, as immigration laws did not directly restrict old immigrants. However, natives’ attitudes towards immigrants (e.g., hostility), especially after the passage of immigration restriction laws, could still affect old immigrants’ spatial behaviors. To study this, I use the linked individual panel and explore two types of spatial behaviors: internal migration and return migration. For the channel of internal migration, the segregation pattern would change when immigrants from more restricted countries be-

came more (or less) likely to move within the U.S. if they previously lived in segregated areas. For the channel of return migration, the segregation pattern would change when immigrants from more restricted countries became more (or less) likely to remain in the U.S. if they previously lived in segregated areas. Again, I focus on the enumeration district (i.e., sub-county) level and examine whether old immigrants' spatial behaviors were related to local demographic characteristics. Consistent with earlier historical research (e.g., Cerase, 1974), I find that selection on return migration might partially explains trends in immigrant segregation: immigrants from more restricted countries were more likely to stay in the U.S. after the 1920s if they previously lived in less segregated areas, especially compared to those from less restricted countries. On the other hand, I find no significant effects of internal migration on immigrant segregation.

This paper adds to the literature of population economics, economic history, public policy, and urban economics by using recently available and usable full-count census data to examine effects of immigration restriction laws on immigrant segregation, and specifically, how the spatial behaviors of U.S. immigrants were affected. The degree of ethnic segregation rose in the early twentieth century U.S. (Logan and Parman, 2017). Black migration (Vigdor, 2002; Collins and Wanamaker, 2014, 2015) contributed to changing segregation patterns, but the rise in immigrant segregation also played a crucial role (Cutler et al., 2008). In recent years, many researchers examine how immigration restriction laws affected European immigrants' return migration (e.g., Greenwood and Ward, 2015; Ward, 2017). On the other hand, less is known about the relationship between immigration restriction laws and trends in immigrant segregation, and furthermore, the possible mechanisms behind such trends in segregation. This paper fills the gap in the related literature.

In the remainder of this paper, Section 2 introduces the historical background. Section 3 introduces data and methods. Section 4 presents the main empirical analysis. Section 5 further discusses mechanisms behind empirical findings of this paper. Section 6 discusses potential effects of segregation on immigrants' assimilation. Section 7 concludes.

2 Background

This section introduces the historical background. I first present a brief overview of European immigration in the age of mass migration and immigration restriction laws in the U.S. I then focus specifically on immigrant segregation in the early twentieth century.

2.1 European Immigration and Immigration Restriction Laws

The U.S. absorbed nearly 30 million immigrants from Europe during the age of mass migration, when the U.S. generally remained its open borders with the exception that Chinese immigration was severely restricted by the Exclusion Act of 1882 (e.g., Kanazawa, 2005). There were several federal laws before 1920 that affected European immigration: the Naturalization Act of 1906 (and the following state-level laws) that required immigrants to learn English (Lleras-Muney and Shertzer, 2015), the Immigration Act of 1907 that limited immigrants with disability and disease, and the Immigration Act of 1917 that imposed literacy tests on immigrants. However, all of the above acts were more or less regulations rather than direct restrictions on immigration, and in particular, European immigration was not restricted based on the country of origin (van Nuys, 2002).

Researchers and the public have long discussed effects of immigration on natives' well-being (e.g., Borjas, 1987; Card, 1990; Borjas et al., 1992; Altonji and Card, 1994; Orrenius and Zavodny, 2006; Ottaviano and Peri, 2012; Akay et al., 2014). Similarly, this was one of the major issues discussed among the public in the age of mass migration: the social environment of anti-immigration populism had been developed since the late nineteenth century with the concern that the mass influx of uneducated and unskilled immigrants would result in economic downturn and harm the native-born population. The Immigration Restriction League was established in 1894, and its members believed that immigrants from Southern and Eastern Europe were “undesirable” relative to Anglo-Saxons (Higham, 1963; Brown and Warner, 1992; Hatton and Williamson, 1998). Despite the social environment of anti-

immigration populism, it was not until 1921 that the Emergency Quota Act was passed, and it was the first law that explicitly restricted European immigration by sending country with quotas. This act set the annual quota for any sending country to 3% of the number of residents of that origin residing in the U.S. in 1910 based on the census. Finally, restrictions on Southern and Eastern European immigration became even more severe in 1924, when a new Immigration Act was passed and the quota for any sending country was set to 2% of the number of residents of that origin residing in the U.S. in 1890. This generated immediate and sizable effects on the number of new immigrants from restricted countries.

Table 1: Immigration Flows in the Early Twentieth Century U.S.

	Italy	Russia	Poland	U.K.	Canada	Ireland	Scandinavia
A. Population size							
1900 census	490,883	424,648	384,058	1,182,239	1,229,923	1,641,393	1,078,478
1910 census	1,351,055	1,562,045	34,554	1,182,209	1,254,880	1,355,740	1,264,917
1920 census	1,608,841	1,450,734	1,133,710	1,172,036	1,213,953	1,355,740	1,183,451
1930 census	1,789,588	1,197,244	1,257,449	1,229,484	1,399,034	929,429	1,119,527
1940 census	1,636,574	1,069,635	1,003,129	955,918	1,119,087	678,074	856,085
B. Year of arrival, 1930							
Before 1900	245,088	204,647	199,602	451,099	503,813	411,353	541,859
1900 - 1909	629,055	456,244	447,188	229,694	218,318	188,827	266,401
1910 - 1919	543,363	364,145	452,566	220,711	217,308	129,798	152,529
1920 - 1924	277,372	140,348	116,534	194,473	255,300	82,812	73,377
1925 - 1929	94,065	32,344	42,209	134,234	204,229	116,724	85,384

Table 1 presents trends in U.S. immigration by country of origin in the early twentieth century. In Panel A, I first report the size of each population in five censuses. Panel A shows that all three “more restricted groups” that had over one million population in 1920—i.e., Italian, Russian, and Polish immigrants—had relatively few immigrants in 1900. However, the number of immigrants from these groups soon became much larger, while the number of immigrants from less restricted groups remained fairly stable. Panel A also points out crucial statistical issues concerning Russian and Polish immigrants. In particular, the number of Polish-born immigrants is abnormally low in the 1910 census, making the number of Russian-born immigrants also less reliable. This could be because that Poland did not gain independence before 1918, and census takers might document Polish immigrants’ birthplace as Russia even if they reported Poland as their birthplace. Since Poland’s inde-

pendence was close to 1920, such issues might still exist in the 1920 census.

In Panel B, I investigate the 1930 census and examine the year of arrival among immigrants by group. Of course, this neglects return migrants who once lived in the U.S., but left before 1930. However, the statistics still suggest the rough pattern of immigration flows for each group. Panel B presents the sharp differences in immigrant cohorts between more and less restricted groups. A large number of immigrants from more restricted countries of origin came to the U.S. during 1920 and 1924, but then immigration from these countries became substantially lower, after the passage of immigration restriction laws. On the other hand, trends in immigration from less restricted countries were relatively stable, and immigration restriction laws had no clear effects on immigration from these countries.

2.2 Immigrant Segregation in the Early Twentieth Century U.S.

Immigrant segregation was determined by several major social factors in the early twentieth century U.S. The period of 1900 to 1940 was the era of mass internal migration. Urbanization accelerated in the 1900s and 1910s (Ferrie, 2003; Hall and Ruggles, 2004), and in the late 1910s the urban population as the percentage of the total population exceeded 50%. Urbanization affected immigrant segregation as many ethnic enclaves were located in cities. These factors could potentially affect all immigrant groups.

On the other hand, immigration restriction laws generated heterogeneous effects on different immigrant groups through various channels. First, in general, new immigrants are more likely to reside within ethnic enclaves (Massey and Denton, 1985; Bartel, 1989), and thus the lack of newcomers caused by immigration laws might reduce immigrant segregation. This is the direct mechanism through which immigration laws could affect immigrant segregation. Moreover, immigration laws might generate indirect effects on immigrant segregation through changes in spatial behaviors of old immigrants in the U.S., even if old immigrants were not directly targeted. Such indirect effects might be related to natives' attitudes towards immigrants who had already in the U.S. One possible mechanism is inter-

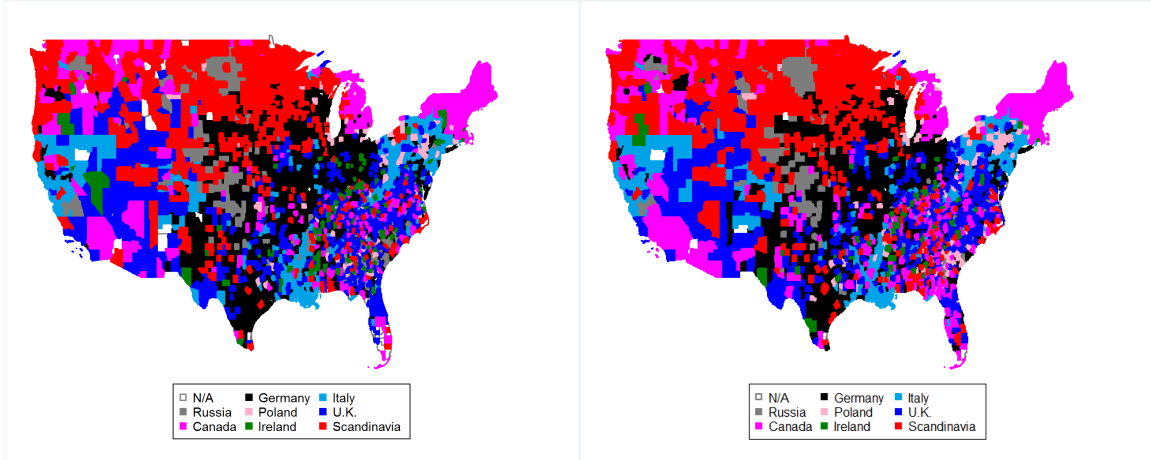


Figure 1: The Largest Immigrant Group by County, 1920 and 1930

nal migration: trends in immigrant segregation could be indirectly affected by immigration laws through internal migration, if old immigrants from countries that were targeted by immigration laws had unique outcomes of internal migration, e.g., if they moved to less segregated areas after immigration restriction laws were made. Another possible mechanism is selective return migration. The return migration rate was high among immigrants, and return migration was selected based on economic status (Abramitzky et al., 2014, 2017; Ward, 2017). While trends in return migration were similar among more or less restricted groups before immigration laws were made (Greenwood and Ward, 2015), the rate of return migration became lower among immigrants from more restricted countries afterwards (Ward, 2017). Immigrant segregation could be affected if return migration was statistically related to ethnic enclave residence prior to departure.

Figure 1 shows two county maps of the largest immigrant group originally from Europe and Canada in 1920 and 1930. The maps show that immigrant groups were concentrated in different areas in both years. While the pattern of ethnic settlements remained fairly stable, Figure 1 still shows declines in the number of counties where Italians and Russians (both were severely restricted) were the largest immigrant group during the 1920s.

In Appendix A, I consider a theoretical framework that analyzes three possible mechanisms behind the effects of immigration restriction laws on immigrants segregation. Here,

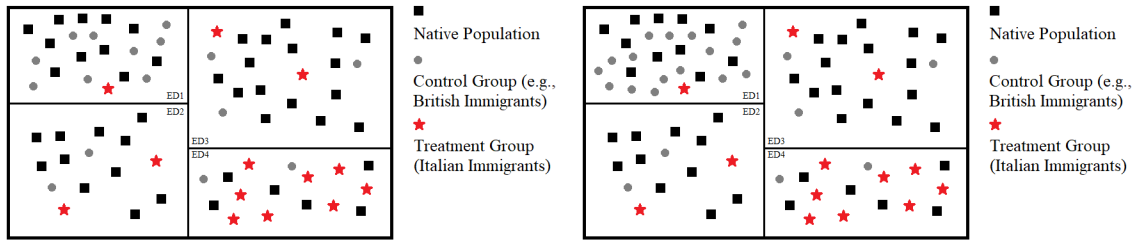


Figure 2: Illustration (1): Immigration Restriction Laws and Immigrant Segregation. (a): Segregation before 1920 (Baseline); (b): Trends in Segregation, Lack of Newcomers

I briefly summarize this framework by providing three illustrative examples from Figure 2 to 4. In each figure, I first present a “baseline” pattern of segregation in the left sub-figure: there are four sub-areas in one county, in which rectangles represent natives, circles represent immigrant groups that were less restricted (e.g., British immigrants), and stars represent immigrant groups that were more restricted (e.g., Italian immigrants). I then present the specific mechanism behind one type of the effect of immigration restriction laws on immigrant segregation in the right sub-figure.

I begin with the direct mechanism in Figure 2. Figure 2 illustrates that immigration restriction laws had a direct effect on, in this case, British and Italian segregation through differences in quotas on new immigration. British segregation rose as new British immigrants chose to reside in British enclaves, while Italian segregation remained stable, with no new immigration from Italy after the passage of immigration restriction laws. In Figure 3 and 4 I further present two indirect mechanisms. Figure 3 shows the relative increase in British segregation and decrease in Italian segregation, because old Italian immigrants previously living in more segregated areas were relatively more likely to move to less segregated areas after immigration restriction laws were made. Figure 4 shows how the mechanism of selective return migration led to the similar result: old Italian immigrants previously living in more segregated areas were less likely to stay in the U.S. after immigration restriction laws were made. In Appendix A, I explain that even if immigration restriction laws did not directly target immigrants who had settled down in the U.S., such mechanisms could still

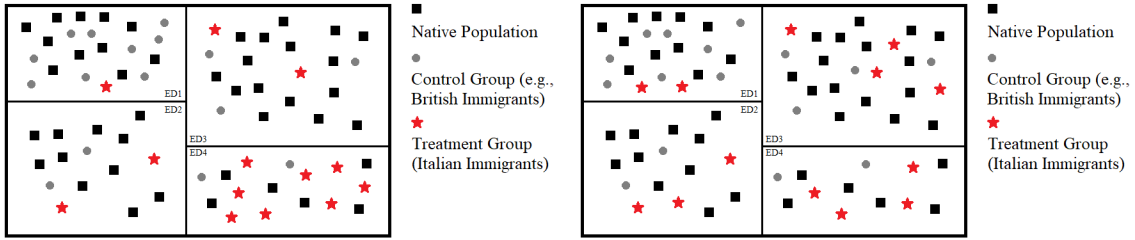


Figure 3: Illustration (2): Immigration Restriction Laws and Immigrant Segregation. (a): Segregation before 1920 (Baseline); (b): Trends in Segregation, Internal Migration

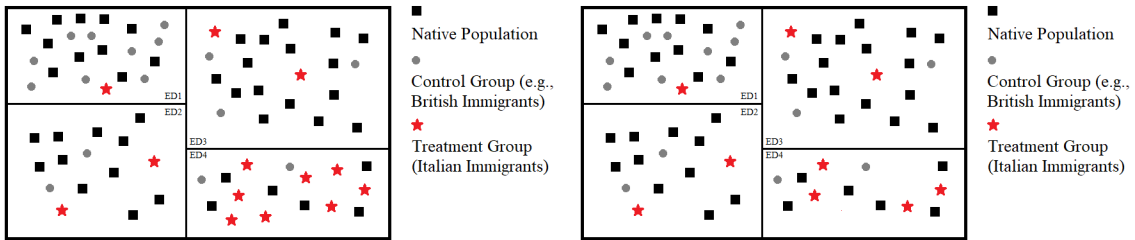


Figure 4: Illustration (3): Immigration Restriction Laws and Immigrant Segregation. (a): Segregation before 1920 (Baseline); (b): Trends in Segregation, Return Migration

be effective because old immigrants—in particular, old immigrants from more restricted countries—might adjust their spatial behaviors according to natives’ attitudes towards immigrants (e.g., Haines, 2000).

The theoretical framework (in Appendix A) shows that immigration laws led to the decline in segregation among immigrants from more restricted countries through restrictions on new immigration. Moreover, old immigrants who had settled down in the U.S. adjusted their spatial behaviors, and those who previously lived in less segregated areas were more likely to stay in the country. On the other hand, without assuming the exact form of old immigrants’ utility function, it is theoretically impossible to predict the sign of the selection on internal migration, and its association with trends in segregation.

To empirically study whether and how immigration restriction laws affected immigrant segregation, I calculate the degree of immigrant-native segregation for each immigrant group at the county level, and use enumeration districts (ED) as the subunits. An ED is

an area in which all households are surveyed by one census taker, which can be as large as a major part of a county in unpopulated rural areas, or as small as a few blocks in cities. On average, there are about 15 to 20 EDs in each county, and each ED contains approximately 1,500 residents. For a specific immigrant group in county j , I consider the dissimilarity index D_j (Duncan and Duncan, 1955; Duncan and Lieberson, 1959; Winship, 1977):

$$D_j = \frac{1}{2} \sum_{k=1}^{N_j} \left| \frac{n_k}{n_j} - \frac{e_k}{e_j} \right| \quad (1)$$

where N_j is the number of enumeration districts in county j , n_k is the number of natives in the enumeration district k , n_j is the total number of natives in county j . For this immigrant group, e_k and e_j can be similarly defined as n_k and n_j .

Table 2: County-Level Segregation by Group, 1900 - 1940

	1900	1910	1920	1930	1940
Italian-native	0.83	0.73	0.72	0.73	0.73
Russian-native	0.85	0.69	0.74	0.70	0.70
Polish-native	0.87	0.81	0.74	0.76	0.76
British-native	0.63	0.41	0.43	0.50	0.54
Canadian-native	0.64	0.46	0.47	0.53	0.56
Irish-native	0.67	0.48	0.52	0.61	0.69
Scandinavian-native	0.65	0.53	0.53	0.56	0.57
Norwegian-native	0.73	0.62	0.63	0.67	0.69
Danish-native	0.78	0.61	0.62	0.66	0.69
Swedish-native	0.71	0.56	0.56	0.61	0.63
Observations	2,819	2,960	3,065	3,099	3,106

Standard deviations are in parentheses.

Based on this index, Table 2 presents trends in segregation from 1900 to 1940. I aggregate immigrants from Scandinavia, but also list each country separately. Table 2 shows that trends were generally similar among all immigrant groups between 1900 and 1920 before immigration laws were made: segregation first declined in the 1900s, and then remained fairly stable during the 1910s. Although immigrants from more restricted countries had different “starting points” of segregation in 1900, trends in immigrant segregation were roughly parallel across all groups before 1920. This coincided with parallel *trends* in new immigration before the passage of immigration restriction laws (although more restricted

groups did have more newcomers), as reported in Table 1. Furthermore, Greenwood and Ward (2015) observe parallel trends in return migration between more or less restricted groups before immigration restriction laws were made. However, trends in segregation did change after the 1920s: there were significant differences in segregation patterns between more and less restricted groups. In particular, the degree of segregation among immigrants from less restricted countries increased after the 1920s, while the degree of segregation among immigrants from more restricted countries became relatively lower.

3 Data and Methods

This section introduces data and methods. I first discuss data and present descriptive statistics. I then analyze empirical strategies.

3.1 Data

I use two datasets in the main empirical analysis of this paper. First, I use a county-level panel dataset covering from 1900 to 1940. To construct this county panel, I aggregate demographic, social, and economic variables at the enumeration district (ED) or county level from individual census records. The descriptive statistics of individual characteristics by country of origin in five censuses are reported in Appendix B. I then use EDs as subunits of county, and calculate segregation for each immigrant group. I aggregate other control variables (e.g., the marriage rate, average occupational scores, and sex ratio) at the county level. I cluster three Scandinavian—Norwegian, Danish, and Swedish—group into one; these three groups had similar characteristics (see Appendix B for details). Note that all seven major immigrant groups are involved in the county panel (and individual records reported in Appendix B), including Polish and Russian immigrants.

Table 3 presents descriptive statistics by census year. The average age rose over time, from 25 years old in 1900 to nearly 30 years old in 1940. The marriage rate also became

higher, from 35.6% in 1900 to 44.5% in 1940. The homeownership rate slightly declined, especially after the 1920s when the Great Depression occurred. The rate of school attendance at the county level remained fairly stable. Finally, there was a decline in the average occupational score between 1910 and 1930, but then a rise between 1930 and 1940.

Table 3: County-Level Panel Data: 1900 - 1940

	1900	1910	1920	1930	1940
Average age	24.728 (3.039)	25.556 (2.981)	26.530 (2.968)	27.872 (3.035)	29.765 (2.992)
Marriage rate	0.356 (0.041)	0.379 (0.033)	0.399 (0.034)	0.414 (0.034)	0.445 (0.035)
Female rate	0.475 (0.044)	0.475 (0.033)	0.482 (0.024)	0.483 (0.023)	0.487 (0.020)
% homeownership	0.556 (0.170)	0.552 (0.154)	0.545 (0.143)	0.508 (0.135)	0.480 (0.125)
% school attendance	0.180 (0.047)	0.238 (0.038)	0.226 (0.034)	0.240 (0.029)	0.217 (0.033)
Average occupational score	6.452 (1.846)	7.070 (1.916)	5.444 (1.280)	5.979 (1.261)	7.338 (1.731)

Standard deviations are in parentheses.

To further examine the “structure” of county-level segregation, I use a linked individual panel dataset constructed based on the 1920 and 1930 census. I link two censuses using the algorithm *Reclink2*, which is used in the LIFE-M project at the University of Michigan (Wasi and Flaaen, 2015; Bailey et al., 2017). The linkage rate is close to 20%, which is comparable to other linkage projects. Results of record linkage suggest that the linked sample is a representative subsample of the linkable population in the 1930 census. I discuss the details of the linkage procedure and linkage quality in Appendix C.

Specifically, I construct this individual panel by linking between the 1920 census and the 1930 “linkable population” who arrived in the U.S. in or before 1920.³ The advantage of using this linked sample is that I can observe individuals across censuses and study changes in segregation and ethnic enclave residence at the sub-county level, and then analyze the “structure” of county-level segregation. One disadvantage of most linkage projects is that only male respondents can be linked, as women usually change surnames after marriage.

³Note that, however, I do not attempt to construct the linked sample of Russian and Polish immigrants due to statistical issues in censuses concerning Polish and Russian immigrants (one cannot determine Polish and Russian origins in the 1920 and 1930 census).

That said, sub-county findings based on the individual panel could still provide additional support for county-level results.

Table 4: Individual-Level Panel Data: 1920 - 1930

	All	Italian	British	Canadian	Irish	Scandinavian
Age, 1920	40.018 (15.143)	34.820 (12.730)	41.498 (15.622)	40.180 (16.792)	43.862 (13.759)	43.877 (14.184)
Years since migration, 1920	22.057 (13.572)	15.784 (8.467)	23.687 (14.969)	20.060 (14.735)	24.727 (14.138)	25.566 (13.280)
Occupational scores, 1920	17.446 (13.774)	17.755 (12.596)	17.790 (14.775)	17.239 (15.017)	17.967 (13.841)	16.674 (12.956)
Occupational scores, 1930	17.964 (14.180)	18.635 (13.057)	18.600 (15.071)	18.043 (15.257)	17.134 (14.036)	16.685 (12.991)
Married, 1920	0.687	0.673	0.708	0.683	0.658	0.713
Married, 1930	0.753	0.805	0.756	0.729	0.694	0.716
Citizenship, 1920	0.568	0.344	0.657	0.559	0.736	0.789
Citizenship, 1930	0.767	0.631	0.847	0.731	0.889	0.802
Urban status, 1920	0.726	0.836	0.751	0.703	0.871	0.529
Urban status, 1930	0.745	0.861	0.760	0.718	0.873	0.532
Farm status, 1920	0.122	0.034	0.084	0.131	0.042	0.295
Farm status, 1930	0.113	0.038	0.077	0.118	0.037	0.292
Homeownership, 1920	0.439	0.343	0.433	0.454	0.356	0.625
Homeownership, 1930	0.577	0.553	0.566	0.564	0.517	0.633
% same origin in ED, 1920	0.095 (0.108)	0.161 (0.146)	0.036 (0.041)	0.082 (0.086)	0.056 (0.051)	0.094 (0.090)
% same origin in ED, 1930	0.076 (0.085)	0.127 (0.112)	0.030 (0.028)	0.070 (0.072)	0.044 (0.048)	0.086 (0.077)
% natives in ED, 1920	0.748 (0.125)	0.673 (0.133)	0.789 (0.112)	0.774 (0.109)	0.745 (0.113)	0.766 (0.104)
% natives in ED, 1930	0.783 (0.108)	0.722 (0.146)	0.815 (0.041)	0.800 (0.086)	0.763 (0.051)	0.796 (0.098)
Segregation with native-born whites, 1920	0.36 (0.162)	0.56 (0.105)	0.26 (0.078)	0.25 (0.091)	0.33 (0.071)	0.34 (0.123)
Segregation with native-born whites, 1930	0.35 (0.141)	0.52 (0.093)	0.28 (0.086)	0.25 (0.085)	0.35 (0.073)	0.33 (0.117)
Observations	490,349	135,346	101,099	115,078	40,080	98,746

Standard deviations are in parentheses. Standard deviations of binary variables are omitted.

Table 4 presents descriptive statistics of this linked individual panel. Italian immigrants were relatively younger and arrived in the U.S. later, but had similar occupational scores with immigrants from less restricted countries. Italian immigrants were less likely to be married in 1920, but then had the highest marriage rate in 1930. In both censuses, the Italian group had the lowest citizenship rate. Italian immigrants had somewhat similar locational characteristics with Irish immigrants: they were more likely to live in cities, and less likely to live in farms; they were also less likely to be homeowners. As Italians (as well

as Poles and Russians) were more likely to reside in cities, it is useful to study immigrant segregation in the urban sample. I will examine this in Section 4.

In the rest of the table I study settlement patterns by country of origin. Italian immigrants lived in enumeration districts with significantly higher shares of immigrants of the same origin, and lower shares of the native-born population. Italians also lived in more segregated U.S. counties. However, Table 4 suggests significant changes in ethnic enclave residence and segregation patterns among Italian immigrants between the 1920 and 1930 U.S. census. The Italian population experienced sharp declines in both the share of immigrants of the same origin at the enumeration district level, and the degree of segregation at the county level. The magnitudes of such demographic changes in the Italian immigrant group appeared to be largest compared to other less restricted groups.

3.2 Empirical Strategies

I first discuss the county panel dataset, in which I collapse the data to origin-county-year cells. This is an unbalanced panel, as a small number of counties cannot be linked due to border changes and, in particular, a few counties in latter censuses did not even exist in earlier censuses. However, the geography of U.S. counties generally remained fairly stable in the early twentieth century, and it is thus possible to create the panel dataset with the highly consistent geographic units. In robustness checks I will also conduct the empirical analysis in the panel with only consistent counties. I estimate the following specification:

$$D_{ect} = \alpha + \sum_{t \neq 1920} \beta^{(t)} S_e \mathbb{1}(t > 1920) + \mathbf{X}_{ct} \gamma + \tau_e + \eta_c + \kappa_t + \varepsilon_{ect} \quad (2)$$

where e indexes the country of origin, c indexes the county, and t indexes the census year. D_{ect} is the dissimilarity index of segregation between the immigrant group e and natives, as introduced in Equation 1. S_e is a binary variable indicating that the immigrant group was the more restricted group in immigration restriction laws. Hence, $S_e \mathbb{1}(t > 1920)$

represents whether group e was the more restricted group in the specific census year t . \mathbf{X}_{ct} is the vector of time-varying county characteristics. I further control for country-of-origin fixed effects τ_e , county fixed effects η_c , and year fixed effects κ_t . Essentially, this is a difference-in-differences framework in which I compare segregation patterns between more and less restricted groups (represented by S_e) before and after immigration restriction laws were made (represented by $\mathbb{1}(t > 1920)$).

Equation 2 is a baseline specification. Based on this, I consider several types of specifications. In the main specification in Section 4.1, I add the following sets of controls: (a) county-by-year fixed effects, which account for all observable and unobservable county time-varying characteristics; (b) continent-by-year fixed effects, which account for quota differences between Canadian immigrants (not restricted) and less restricted groups. I also run weighted regressions base on county population characteristics.

I further consider several types of changes to sample: (a) I exclude the Russian and Polish group (with the origin defined based on the mother tongue in some censuses) and redo the analysis. (b) I study the sample without Irish immigrants, as Table 2 shows slightly different trends in Irish segregation. (c) I study the sample without Canadian immigrants, who were never restricted. (d) I exclude the 1900 census, as there were relatively fewer immigrants from more restricted countries in 1900. (e) I study the sample with only consistent counties, in which county borders remained unchanged.

I finally conduct the placebo tests, in which I suppose that immigration restriction laws had been imposed against other major immigrant groups that were less or not restricted in actual history. I then reestimate the above specification for these “fake more restricted groups”. If the conclusion is similar to that observed among immigrants from actual more restricted countries, one may worry that there were other factors affecting immigrant segregation, and immigration restriction laws were not the major factors, because actual more restricted groups were not specifically targeted and affected overall.

Following the county-level analysis, I turn to analyze the structure of segregation by

studying ethnic enclave residence in enumeration districts using the linked individual panel. I employ the difference-in-difference strategy by replicating Equation 2:

$$D_{iect} = \alpha + \beta S_{ie} \mathbb{1}(t > 1920) + \mathbf{Z}_{it} \gamma + \tau_i + \eta_{ct} + \varepsilon_{iejct} \quad (3)$$

where i indexes the individual. S_{ie} is the indicator that i 's group was the more restricted group in immigration restriction laws. I include individual controls and fixed effects, and time-varying characteristics. In this specification, β reflects the effects of immigration restriction laws on immigrants' settlement patterns in terms of county-level segregation. I further estimate effects of immigration restriction laws on the share of immigrants of the same origin $(\%SO)_{iejct}$, where j indexes the enumeration district in which i lived:

$$(\%SO)_{iejct} = \alpha + \beta S_{ie} \mathbb{1}(t > 1920) + \mathbf{Z}_{it} \gamma + \tau_i + \eta_{ct} + \varepsilon_{iejct} \quad (4)$$

4 Empirical Analysis: Trends in Immigrant Segregation

This section studies the main empirical question of this paper: did immigration restriction laws affect trends in immigrant segregation in the early twentieth century U.S.? In Section 4.1, I present main findings. In Section 4.2, I conduct several robustness checks. In Section 4.3, I use the linked individual panel sample to study effects of immigration laws on individuals' settlement patterns at the enumeration district (i.e., sub-county) level.

4.1 Main Results: County Panel

In Table 5, I report the main findings of this paper. I start with the full sample, in which all seven major immigrant groups are included. Four less restricted groups were British, Canadian, Irish, and Scandinavian immigrants, and three more restricted groups were Italian, Polish, and Russian immigrants.

Table 5: Immigration Restriction Laws and Immigrant Segregation, Full Sample

	Segregation (D_{ect}) between Natives/Native Whites and Immigrants (of the Specific Origin)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1920 county-level avg.: 0.73 (treatment-native) 0.49 (control-native)				1920 county-level avg.: 0.73 (treatment-whites) 0.49 (control-whites)			
Restricted, 1900 (pre-legislation)	-0.028 (0.026)	-0.026 (0.025)	-0.034 (0.032)	0.024 (0.037)	-0.029 (0.026)	-0.027 (0.026)	-0.035 (0.032)	0.020 (0.038)
Restricted, 1910 (pre-legislation)	0.031 (0.037)	0.033 (0.037)	0.027 (0.051)	0.056 (0.036)	0.031 (0.037)	0.032 (0.037)	0.027 (0.051)	0.053 (0.035)
Restricted, 1930 (post-legislation)	-0.082*** (0.021)	-0.082*** (0.021)	-0.093** (0.027)	-0.077*** (0.014)	-0.083*** (0.021)	-0.082*** (0.021)	-0.093** (0.027)	-0.073*** (0.015)
Restricted, 1940 (post-legislation)	-0.137*** (0.030)	-0.137*** (0.030)	-0.151*** (0.035)	-0.128*** (0.022)	-0.137*** (0.030)	-0.137*** (0.030)	-0.152*** (0.034)	-0.125*** (0.022)
Sample	Full	Consistent counties	Excluding Italian	Weighted ct. pop.	Full	Consistent counties	Excluding Italian	Weighted ct. pop.
Adjusted R ²	0.653	0.654	0.672	0.761	0.651	0.652	0.668	0.761
Observations	76,936	74,460	68,087	76,936	76,936	74,460	68,087	76,936

Standard errors are in parentheses. *: $p < .05$; **: $p < .01$; ***: $p < .001$.



Figure 5: Immigration Restriction Laws and Immigration Segregation: (a) Immigrants and Natives; (b) Immigrants and Native-Born Whites

From Column 1 to 4 I study segregation between immigrants and the native-born population. I control for county-by-year fixed effects, country-of-origin fixed effects, and standard errors are clustered at the country-of-origin level. Column 1 shows that more and less restricted groups followed similar trends in segregation before the 1920s, with no statistically significant pre-legislation coefficients. After the 1920s, however, segregation among immigrants from more restricted countries declined relative to segregation among immigrants from less restricted countries. This is consistent with the descriptive findings shown in Table 2, i.e., the passage of immigration restriction laws led to differences in segregation patterns between immigrant groups that were more and less restricted.

Column 2 shows similar results when I focus only on counties with consistent borders (and thus the county panel is now balanced). In Column 3 I exclude Italian immigrants, and in Column 4 I run a weighted regression based on the county population. I find the similar empirical conclusion. In general, the effect size shown in Table 5 appears to be large relative to the average degree of county-level immigrant segregation in 1920.

The next question related to the main results is: if immigration restriction laws did lead to the decline in segregation, who moved to immigrants' enclaves? Prior studies show some indirect evidence that the native-born white population—rather than African Americans—might contribute most to trends in immigrant segregation. Cutler et al. (2008) document low interactions between whites and African Americans in the early twentieth century. Specifically, Collins (1997) argue that African Americans' migration was not parallel with European immigration. To further analyze this, I rerun the regressions from Column 5 to 8, and find very similar quantitative results. While not reported here, I find no significant effects on immigrant-black segregation, suggesting that trends in immigrant segregation were most associated with migration among the native-born white population.

I present illustrative results of event-study coefficients in Figure 5. In the left figure, I focus on effects of immigration laws on immigrant-native segregation; in the right figure, I focus on segregation between immigrants and native-born whites. Figure 5 suggests graphic evidence that (a) pre-trends are indistinguishable from zero, and (b) the degree of segregation among immigrants from more restricted countries became relatively lower after the passage of immigration laws.

In Section 2 and 3, I discuss the statistical issues concerning Polish and Russian immigrants in earlier censuses. While I do find the effects of immigration restriction laws on immigrant segregation in Table 5, one might worry that the results should be interpreted with caution. I redo the analysis based on the sample without Polish and Russian immigrants, i.e., Italian immigrants are the only more restricted group in this sample.

Figure 6 presents graphic results. Compared with the main table, Figure 6 shows dif-

ferent coefficients in both the pre- and post-legislation period. Indeed, Table 5 similarly shows that the magnitude of the effect appears to be larger in the sample that excludes Italian immigrants (see Column 3 and 7). Nevertheless, Figure 6 presents similar evidence of the effects of immigration restriction laws in the sample without Polish and Russian immigrants, suggesting that the main results are unlikely to be driven by data issues in censuses or the construction of the treatment group.

I conclude the main analysis by discussing standard errors. Clustering at the country-of-origin level is the standard strategy (Bertrand et al., 2004) in a typical difference-in-differences setting similar to that in this paper. However, one might worry that there are relatively few groups in the context of this paper. In particular, in the sample that excludes Polish and Russian immigrants, there is essentially only one treatment group (i.e., Italians). An alternative strategy is to cluster standard errors at the origin-by-year level. While not reported here, this strategy yields similar results of pre-trends (pre-legislation coefficients are insignificant), and actually smaller standard errors for post-legislation coefficients.

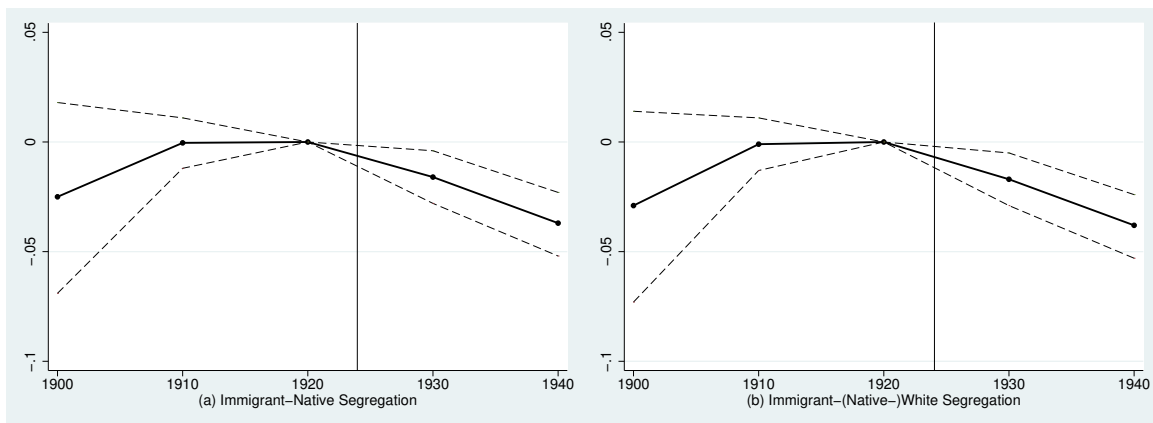


Figure 6: Immigration Restriction Laws and Immigration Segregation, the Italian Sample: (a) Immigrants and Natives; (b) Immigrants and Native-Born Whites

4.2 Robustness Checks

I now conduct several additional tests to check the robustness of the main results. I start with a variety of changes to samples and specifications in Table 6. In all regressions I

control for county-by-year fixed effects. Results are very similar when I control for other sets of fixed effects (e.g., including county and year fixed effects separately, plus time-varying county controls) and cluster standard errors at different levels. Again, I first study segregation between immigrants of the specific origin and the native-born population from Column 1 to 4, and then study immigrant-white segregation from Column 5 to 8.

Table 6: Immigration Restriction Laws and Immigrant Segregation, Changes to Sample

	Segregation (D_{ect}) between Natives/Native Whites and Immigrants (of the Specific Origin)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Immigrant-native segregation				Immigrant-white segregation			
Restricted, 1900 (pre-legislation)	-0.021 (0.017)	-0.025 (0.023)		0.007 (0.030)	-0.026 (0.016)	-0.026 (0.023)		0.003 (0.030)
Restricted, 1910 (pre-legislation)	0.044 (0.041)	0.031 (0.046)	0.028 (0.045)	0.053 (0.042)	0.041 (0.041)	0.031 (0.046)	0.027 (0.045)	0.050 (0.042)
Restricted, 1930 (post-legislation)	-0.063** (0.015)	-0.082* (0.028)	-0.078** (0.025)	-0.073** (0.017)	-0.059* (0.015)	-0.082* (0.028)	-0.079** (0.026)	-0.069** (0.017)
Restricted, 1940 (post-legislation)	-0.102*** (0.011)	-0.135** (0.035)	-0.130** (0.029)	-0.121** (0.020)	-0.099*** (0.011)	-0.135** (0.035)	-0.131** (0.029)	-0.118** (0.020)
Sample	Excluding Irish	Excluding Canadian	Excluding 1900	Weighted, % citizen	Excluding Irish	Excluding Canadian	Excluding 1900	Weighted, % citizen
Adjusted R ²	0.758	0.632	0.677	0.759	0.759	0.631	0.675	0.759
Observations	64,688	64,183	68,528	76,936	64,688	64,183	68,528	76,936

Standard errors are in parentheses. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

In Column 1, I exclude Irish immigrants from the sample. Descriptive statistics presented in Section 2 and 3 show that Irish immigrants might have different settlement patterns in the U.S. Therefore, it is useful to conduct the robustness check in a sample without Irish immigrants. Column 1 of Table 6 shows different pre-legislation estimates (and are still insignificant), but post-legislation coefficients are quantitatively very similar, suggesting that the inclusion of Irish immigrants should not affect the main conclusion of this paper. In Column 2 I exclude Canadian immigrants, who were not restricted in the 1920s. I find similar results based on this subsample. In Column 3 I drop observations in the 1900 census. The coefficient for 1910 (pre-legislation) is still insignificant, while magnitudes of effects in the post-legislation period are similar to those in main findings. In Column 4 I run a weighted regression, in which weights are constructed based on the citizenship rate at the county level. I again observe quantitatively similar effects of immigration restriction laws in this model. From Column 5 to 8, I redo the analysis using degrees of immigrant-white

segregation as dependent variables, and find very similar results.

Section 3 and Appendix B mention that more restricted groups are generally more likely to reside in cities. This suggests that more and less restricted groups might have different “starting points” of segregation in terms of urban residence. Even if two types of groups did follow parallel trends in segregation before the 1920s, it is useful to examine the heterogeneous effects by urban status, and it is particularly interesting to estimate the effects of immigration restriction laws on immigrant segregation in urban areas.

Table 7: Immigration Restriction Laws and Immigrant Segregation, Urban/Rural Sample

	Segregation (D_{ect}) between Natives/Native Whites and Immigrants (of the Specific Origin)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Immigrant-native segregation				Immigrant-white segregation			
Restricted, 1900 (pre-legislation)	-0.024 (0.031)	-0.113** (0.021)	-0.024 (0.031)	0.007 (0.040)	-0.027 (0.043)	-0.115** (0.020)	-0.026 (0.031)	0.002 (0.040)
Restricted, 1910 (pre-legislation)	0.032 (0.038)	0.015 (0.037)	0.033 (0.038)	0.057 (0.033)	0.031 (0.038)	0.015 (0.037)	0.031 (0.038)	0.054 (0.033)
Restricted, 1930 (post-legislation)	-0.079*** (0.020)	-0.065* (0.025)	-0.079** (0.020)	-0.073*** (0.013)	-0.079*** (0.020)	-0.065* (0.025)	-0.078*** (0.020)	-0.069*** (0.013)
Restricted, 1940 (post-legislation)	-0.132*** (0.028)	-0.119** (0.030)	-0.132*** (0.028)	-0.122*** (0.019)	-0.132*** (0.028)	-0.120** (0.030)	-0.132*** (0.028)	-0.118*** (0.019)
Sample	Urban	Rural	Urban consistent	Urban weighted	Urban	Rural	Urban consistent	Urban weighted
Adjusted R ²	0.706	0.614	0.707	0.775	0.706	0.611	0.706	0.777
Observations	45,655	31,281	44,980	45,655	45,655	31,281	44,980	45,655

Standard errors are in parentheses. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

I begin with the urban county sample in Column 1, Table 7. I find very similar effects of immigration restriction laws on immigrant-native segregation in urban areas. I observe insignificant pre-legislation coefficients, and find the relative decline of segregation among more restricted groups after the passage of immigration restriction laws. However, results based on the rural sample are less robust (the pre-legislation coefficient in 1900 is significant). This might be because relatively fewer immigrants from more restricted countries resided in rural areas even in 1900, and the estimates might be less reliable. I then return to the urban sample: in Column 3 I focus on urban counties with consistent borders, and in Column 4 I introduce weights constructed based on county populations. Both columns show similar effects of immigration restriction laws in the urban sample. From Column 5 to 8 I repeat the exercise using immigrant-white segregation as dependent variables, and

find that the results remain unchanged. Hence, the main findings of this paper are at least robust in the sample of urban counties, in which most immigrants from more restricted countries resided in the early twentieth century U.S.

I conclude robustness checks by discussing placebo tests. In Appendix D, I conduct several placebo tests in which I suppose that immigrants from (one of the) less restricted countries were affected by immigration restriction laws. If one can observe similar “effects” of immigration restriction laws, then the main results reported in this section should be unreliable because it was probably not immigration laws—but other social factors—that led to differences in segregation patterns between more and less restricted groups after the 1920s. Results reported in Appendix D, however, show that immigration restriction laws had the particular effects on more restricted groups, and different trends in immigrant segregation should be indeed associated with immigration restriction laws.

4.3 Sub-County Evidence from Individual Records

I conclude this section by presenting sub-county evidence of trends in immigrant segregation based on individual panel data in the 1920 and 1930 census. The individual panel data contain all four major less restricted groups (British, Canadian, Irish, and Scandinavian), and Italian immigrants who were mainly restricted. Although some studies use individual-level panel data from 1900 to 1940 (e.g., Abramitzky et al., 2014), this idea of data construction might not work in this paper, as the long individual panel might make the linked sample biased towards old immigrants (who had been in the U.S. in 1900), and thus the linked sample is no longer representative to immigrants in 1930. I also do not link Russian and Polish immigrants due to possible inconsistency of the birthplace information.

To explore the “structure” of immigrant segregation, I examine demographic changes by immigrant group at the sub-county level. Specifically, in Table 8, I focus on the enumeration district (ED) level from 1920 to 1930, and estimate Equation 3 and 4. In Column 1 and 2, I regress the share of immigrants of the same origin in the ED on the interaction

Table 8: Immigration Restriction Laws and Immigrant Segregation, Individuals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ethnic enclave characteristics:				Segregation measure:			
	% same origin, ED		% native, ED		with native-born whites		with natives	
Average in 1930:	0.127		0.722		0.52		0.52	
$S_{ie1}(t > 1920)$	-0.025** (0.008)	-0.027** (0.010)	0.025*** (0.004)	0.029*** (0.002)	-0.049** (0.013)	-0.052** (0.018)	-0.054*** (0.013)	-0.057*** (0.017)
Controls	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
State FE	Yes	No	Yes	No	Yes	No	Yes	No
County FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R ²	0.288	0.372	0.393	0.558	0.726	0.856	0.719	0.854
Observations	980,698	980,698	980,698	980,698	980,698	980,698	980,698	980,698

Standard errors are clustered at the origin-by-year level and are in parentheses. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

between the indicator of the more restricted group and the year dummy, with state and county fixed effects, respectively. I find that after the 1920s, immigrants from more restricted countries (in this table, Italians) who remained in the U.S. lived in EDs with lower shares of compatriots. The effects were large relative to the average share of compatriots (0.127). Column 3 and 4 present similar results: after the 1920s, remaining immigrants from more restricted countries lived in EDs with higher shares of natives. From Column 5 to 8, I turn to focus on county-level immigrant segregation, and find significantly different effects of immigration restriction laws between more and less restricted groups in terms of county-level segregation. The sizes of the effects were large relative to the average degree of segregation. These sub-county findings are consistent with the main conclusion.

5 Discussions: Mechanisms

In Section 4, I show that immigration restriction laws had different effects on immigrant segregation among more and less restricted groups. There might be multiple mechanisms behind trends in segregation. In Appendix A I present a theoretical framework that analyzes three possible mechanisms, and illustrate them in Section 2. In this section, I empirically test these mechanisms.

I first investigate the direct mechanism: immigration laws could directly affect trends in immigrant segregation through restrictions on new immigration. I use the cross-sectional 1920 and 1930 census to study restrictions on new immigration. I then focus on two indi-

rect mechanisms related to immigrants who had been in the U.S. before 1920. Although immigration laws did not restrict old immigrants, they might adjust their settlement patterns after immigration laws were made. Specifically, I examine internal migration of remaining immigrants, and selective return migration. To study the latter two mechanisms, I use linked individual panel data and analyze spatial outcomes of remaining immigrants.

5.1 Restrictions on New Immigration

I first examine the direct effects of immigration laws on trends in immigrant segregation. Immigration restriction laws had immediate and sizable impacts on the number of new immigrants from more restricted countries. In contrast, the number of immigrants from other countries remained fairly stable. Classical findings in labor economics and demography point out that new immigrants are generally more likely to reside in ethnic enclaves (Massey and Denton, 1985; Bartel, 1989; Altonji and Card, 1991), and with the increase in years since migration, immigrants follow trajectories of “spatial assimilation” by moving outside ethnic enclaves. If this was similarly true in the early twentieth century, then one should expect the decline in segregation in groups where new immigration was restricted.

I study this question in Table 9. Here, I only include Italian immigrants as the restricted group due to statistical issues regarding Russian and Polish immigrants in the 1920 census. While not reported here, I find similar results with the inclusion of Russian and Polish immigrants defined based on the mother tongue. On top of Table 9 I first show the average share of immigrants of the same origin in the ED. The descriptive findings suggest Italian immigrants resided in areas with significantly higher shares of immigrants of the same origin, compared to less restricted groups. However, there was a sharp decline in the local share of immigrants of the same origin between 1920 and 1930 (from 17.1% to 13.3%), compared to that among less restricted groups.

I then regress the share of immigrants of the same origin in the ED on years since migration (divided by 10, hence is “decades since migration”), and other individual-level

Table 9: Years since Migration and Ethnic Enclave Residence (1): State Controls

	% of Immigrants of the Same Origin, at Enumeration District (ED) Level					
	(1)	(2)	(3)	(4)	(5)	(6)
	Italian	British	Canadian	Irish	Scandinavian	Pooled
1920 census:						
Average:	0.171	0.036	0.080	0.055	0.087	0.092
Decades since migration	-0.013*** (0.003)	-0.003** (0.001)	-0.002* (0.001)	-0.002** (0.001)	-0.003*** (0.001)	-0.003*** (0.000)
(Decades since migration)*Italy						-0.008** (0.002)
Adjusted R ²	0.181	0.228	0.418	0.224	0.267	0.305
Observations	1,609,346	1,159,148	1,271,325	1,050,668	1,183,603	6,220,090
1930 census:						
Average:	0.133	0.033	0.072	0.050	0.070	0.078
Decades since migration	-0.007*** (0.001)	-0.003*** (0.000)	-0.002*** (0.000)	-0.004*** (0.001)	-0.005*** (0.001)	-0.003*** (0.000)
(Decades since migration)*Italy						-0.005** (0.001)
Adjusted R ²	0.169	0.189	0.428	0.204	0.206	0.278
Observations	1,788,943	1,230,211	1,398,968	929,514	1,119,550	6,467,186
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors are in parentheses and are clustered at the state level. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

I report coefficients of decades since migration (i.e., years since migration are divided by 10). State FE are included.

control variables introduced in Appendix B. In Table 9, I run regressions in each group from Column 1 to 5; I then pool the sample, and examine differences in ethnic enclave residence between more and less restricted groups in Column 6. Results show evidence of spatial assimilation over time: immigrants who lived in the U.S. longer also resided in areas with lower shares of immigrants of the same origin. The rate of spatial assimilation, reflected by the coefficient of “decades since migration”, was largest among Italian immigrants, i.e, the more restricted group. Column 1 also presents a sharp decrease in the rate of spatial assimilation among Italian immigrants. This is expected, as the average years since migration among Italian immigrants should become substantially larger after new Italian immigration was severely restricted by immigration laws, and thus the rate of spatial assimilation dropped. Furthermore, Column 6 shows particularly large effects of years since migration on ethnic enclave residence among Italian immigrants, suggesting that immigration restriction laws directly affected immigrant segregation through restrictions on new

immigration. which is predicted by the analytical framework presented in Appendix A. In Appendix E I redo the analysis based on county fixed effects, and find very similar results.

5.2 Internal Migration

Immigration laws in the 1920s did not put limitations on old immigrants who had settled down in the U.S. However, old immigrants might adjust their settlement patterns after immigration laws came into effect. Hence, immigration laws might generate indirect effects through changes in spatial behaviors of old immigrants. I start with the discussion of internal migration of old immigrants who remained in the U.S. during the 1920s.

Table 10: Internal Migration Rate during the 1920s, County or State Level

	All	Italian	British	Canadian	Irish	Scand.	Norwegian	Danish	Swedish	Control
Migration type:										
Any	44.7%	52.1%	39.1%	30.0%	55.8%	41.1%	38.6%	46.1%	41.5%	41.8%
Inter-state	32.5%	38.2%	27.6%	28.1%	43.6%	30.4%	27.0%	35.3%	31.7%	30.3%
Within-state	17.9%	22.4%	15.9%	16.4%	21.4%	15.4%	15.9%	16.7%	14.4%	16.4%
Observations	490,349	135,146	101,099	115,078	40,080	98,746	39,486	15,137	43,943	355,003

Table 10 presents the internal migration rate by group during the 1920s. I calculate the internal migration rate using linked individual panel data, by comparing places of residence in the two censuses. I consider internal migration along three dimensions: (a) any type of migration, including both inter-state and within-state inter-county migration; (b) inter-state migration; (c) within-state inter-county migration. Table 10 reports that the overall internal migration rate was high among immigrants between 1920 and 1930, and was significantly higher than that of the full U.S. population (e.g., Molloy et al., 2011). Note that, however, the traditional measure of internal migration is based on inter-state migration, as researchers are only able to observe whether individuals reside outside their state of birth. Indeed, only 32.5% of all immigrants moved to other states in the 1920s, which is close to the internal migration rate documented in earlier research. Finally, for immigrants who remained in the same state, 20% of them moved to other counties.

In Appendix A, I argue that without assuming the exact form of immigrants' utility

function, it is theoretically impossible to predict the sign of the effect of immigration restriction laws on ethnic enclave residence (and furthermore, segregation) through internal migration among old immigrants remaining in the U.S. It is thus an empirical question how internal migration of remaining immigrants was associated with trends in segregation after the passage of immigration restriction laws. To examine the effects of internal migration on segregation, I first estimate the following specification within the more restricted group (in this section, Italians) in the linked sample:

$$D_{it} = \alpha + \beta M_i \mathbb{1}(t > 1920) + \gamma M_i + \delta \mathbb{1}(t > 1920) + \mathbf{X}_{ij} \mu + \tau_{i(s)} + \varepsilon_{it} \quad (5)$$

where D_{it} is the degree of segregation in i 's county of residence. M_i is a binary indicator of internal migration. This is essentially a two-period difference-in-differences structure, in which I compare the degree of segregation in 1920 and 1930 between movers and non-movers in the more restricted group. I control for individual time-varying characteristics introduced in Section 3, state fixed effects, and cluster standard errors by year. I then include immigrants from less restricted countries in the regression, and further estimate a difference-in-difference-in-differences specification as follows:

$$D_{iet} = \alpha + \beta S_e M_i \mathbb{1}(t > 1920) + \gamma S_e M_i + \delta S_e \mathbb{1}(t > 1920) + \lambda M_i \mathbb{1}(t > 1920) \\ + \rho S_e + \phi M_i + \psi \mathbb{1}(t > 1920) + \mathbf{X}_{iet} \mu + \tau_{i(s)t} + \kappa_e + \varepsilon_{iet} \quad (6)$$

where e indexes the country of origin. Similar to the main empirical analysis, I cluster standard errors at the origin-by-year level.

In Table 11 I estimate Equation 5 and 6. I first focus only on Italian immigrants. The first panel shows that Italian immigrants who migrated within the U.S. indeed moved to less segregated areas. However, the effect was very small. After including less restricted groups in the second panel, I find no significant relationship between segregation and internal migration. In Appendix F, I further examine within-state migration and inter-state

Table 11: The Effects of Internal Migration: Overall (Any Type)

	Degree of segregation		% same origin, ED		% native, ED	
	(1)	(2)	(3)	(4)	(5)	(6)
Italian:						
$M_i \mathbb{1}(t > 1920)$	-0.007*	-0.004*	-0.006*	-0.007**	0.004*	0.008***
	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)
Adjusted R ²	0.092	0.367	0.108	0.169	0.184	0.326
Observations	270,692	270,692	270,692	270,692	270,692	270,692
All immigrants:						
$S_e M_i \mathbb{1}(t > 1920)$	0.001	-0.001	-0.0001	-0.003	0.001	0.001
	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.002)
Adjusted R ²	0.628	0.727	0.206	0.287	0.256	0.388
Observations	980,698	980,698	980,698	980,698	980,698	980,698
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	No	Yes	No	Yes

Standard errors are in parentheses and clustered at the origin-by-year level. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

migration in more or less groups. Although I find significant results in a few models, results are not consistent and suggest no overall effects on internal migration on immigrant segregation. In sum, internal migration should not be the major channel through which immigration restriction laws affected immigrant segregation.

5.3 Return Migration

I now study return migration. The rate of return migration among European immigrants in the early twentieth century was very high (Greenwood and Ward, 2015), especially after the 1920s when immigrant laws were made (Ward, 2017). This could influence trends in segregation if return migration was selective based on individual characteristics. In Appendix A, I argue that more restricted groups would become less segregated due to the decline in ethnic enclaves, as (a) after the passage of immigration restriction laws, recent immigrants were less likely to remain in the U.S. than immigrants who were socially and economically more established (Ward, 2017), and (b) recent immigrants were more likely to reside in ethnic enclaves (Bartel, 1989).

The direct way to study return migration is to obtain (a) departure records (Ward, 2017), or (b) records in European countries that documented migratory experiences (Abramitzky

Table 12: The 1920 Census and the Linked Sample (Appearing in Both Censuses)

	Italian		British		Canadian		Irish		Scandinavian	
	1920 census	Linked sample	1920 census	Linked sample	1920 census	Linked sample	1920 census	Linked sample	1920 census	Linked sample
Segregation, with native-whites	0.57 (0.106)	0.56 (0.105)	0.26 (0.085)	0.26 (0.078)	0.25 (0.097)	0.25 (0.091)	0.33 (0.071)	0.33 (0.061)	0.34 (0.123)	0.34 (0.122)
Segregation, with natives	0.57 (0.106)	0.56 (0.104)	0.27 (0.086)	0.27 (0.079)	0.26 (0.098)	0.26 (0.091)	0.33 (0.074)	0.33 (0.064)	0.35 (0.123)	0.35 (0.123)
# of the same origin, ED level	333 (436)	327 (427)	63 (122)	64 (104)	141 (202)	144 (205)	84 (99)	92 (102)	120 (136)	130 (146)
# of natives, ED level	1275 (731)	1270 (704)	1296 (678)	1294 (664)	1235 (630)	1231 (605)	1215 (633)	1224 (619)	1094 (590)	1120 (592)
% of the same origin, ED level	0.165 (0.149)	0.161 (0.144)	0.035 (0.039)	0.036 (0.038)	0.080 (0.086)	0.082 (0.086)	0.052 (0.052)	0.056 (0.051)	0.088 (0.082)	0.094 (0.090)
% of natives, ED level	0.671 (0.139)	0.673 (0.133)	0.791 (0.121)	0.789 (0.112)	0.772 (0.117)	0.774 (0.109)	0.751 (0.125)	0.745 (0.113)	0.772 (0.110)	0.766 (0.104)
Occupational score	17.275 (12.243)	17.755 (12.596)	16.991 (14.667)	17.740 (14.677)	16.397 (14.937)	17.239 (15.017)	16.399 (13.934)	17.966 (13.841)	16.007 (12.564)	17.472 (12.820)
Observations	955,017	135,346	606,868	99,368	595,914	115,078	460,886	40,080	663,798	43,943

Standard deviations are in parentheses.

et al., 2017). However, there are currently no nationwide departure records, or migratory records for all European countries. Another indirect way to study return migration is to link two censuses and observe the individual characteristics among those who remained in the country (Abramitzky et al., 2014). Table 12 compares immigrants in the 1920 census and the linked individual panel. This presents differences between all immigrants who lived in the U.S. in 1920 and immigrants who remained in the U.S. during the 1920s, and thus the linked individual panel excludes return migrants. The first two rows of the table show that remaining immigrants from the more restricted countries (here, Italians) lived in slightly less segregated counties in 1920.

To study who left the country and how return migration could affect segregation patterns, one needs to examine local ethnic composition at the sub-county level, as county-level segregation is calculated based on demographic characteristics of subunits within the county. In the next four rows I compare ethnic composition between the 1920 census and the linked individual panel in the ED. Results show that Italian immigrants who remained in the U.S. lived in EDs with relatively fewer compatriots and more natives. On the contrary, immigrants from less restricted countries who remained in the U.S. after the 1920s lived in EDs with more compatriots. This is consistent with the theoretical prediction in Appendix

A, and could partially explain differences in segregation between more and less restricted groups. Indeed, immigrants who lived in areas with fewer immigrants of the same origin and more natives were generally more assimilated in terms of socioeconomic status, and scholars do find evidence of negative selection of out-migration and positive selection of in-migration (Abramitzky et al., 2014, 2017), especially after immigration laws were made (Greenwood and Ward, 2015; Massey, 2016; Ward, 2017). I finally study differences in occupational scores. In all immigrant groups I find that immigrants who remained in the U.S. had higher occupational scores. This is consistent with findings that return migration of European immigrants was negatively selected based on economic status (Greenwood and Ward, 2015; Ward, 2017; Abramitzky et al., 2017).

I further study effects of return migration on immigrant segregation by estimating the following specification in the 1920 census:

$$(\%SO)_{iej} = \alpha + \beta S_e R_i + \gamma S_e + \delta R_i + \mathbf{X}_{iej} \mu + \tau_{i(s)} + \kappa_e + \varepsilon_{iej} \quad (7)$$

where i indexes the individual, e indexes i 's country of origin, and j indexes the enumeration district. I examine demographic variables at the sub-county level: $(\%SO)_{iej}$ is the share of immigrants of the same origin as i in the enumeration district j . S_e is a binary variable indicating that i 's group e was the more restricted group, and R_i is a binary variable indicating that i was found in the 1930 census (i.e., i remained in the U.S. during the 1920s). \mathbf{X}_{iej} is the vector of individual characteristics, and $\tau_{i(s)}$ are state fixed effects. I cluster standard errors by country of origin. The interaction term $S_e R_i$ reflects differences between (a) all immigrants in the 1920 and immigrants who remained in the U.S. after the 1920s, and (b) immigrants from more and less restricted countries.

In Column 1, Table 13, I present the baseline estimation of Equation 7, using the share of immigrants of the same origin in the enumeration district as the dependent variable. I include individual controls and country-of-origin fixed effects, but do not control for any

Table 13: The Effects of Return Migration: the 1920 Census

	% same origin, ED			# same origin, ED‡		
	(1)	(2)	(3)	(4)	(5)	(6)
Average, all:		0.094			171	
Average, Italian:		0.165			333	
$S_e R_i$	-0.006** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)	-7.663** (1.649)	-8.863** (2.320)	-8.791*** (1.573)
S_e	0.082*** (0.001)	0.097*** (0.013)	0.098*** (0.015)	180.573*** (3.141)	215.127*** (21.171)	220.345*** (24.904)
R_i	0.003* (0.001)	0.002 (0.001)	0.001* (0.000)	2.445 (1.717)	-0.987 (2.012)	-0.801 (1.379)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	No	No	Yes	No
County FE	No	No	Yes	No	No	Yes
Adjusted R ²	0.202	0.289	0.377	0.213	0.276	0.340

Standard errors are in parentheses and are clustered at the country-of-origin level. Observations: 3,282,483.

*: $p < .05$; **: $p < .01$; ***: $p < .001$. ‡: controlling for ED population in the regression.

geographic unit fixed effects. Results show that Italian immigrants who remained in the U.S. lived in areas with relatively lower shares of immigrants of the same origin in 1920. I find similar results in Column 2 in which I control for state fixed effects, and in Column 3 in which I control for county fixed effects. The magnitudes of the effects, however, appear to be fairly small relative to the average size of ethnic enclaves. Subsequently, I investigate the number of immigrants of the same origin from Column 4 to 6, with the additional control of the population size of the enumeration district. I again find the similar results. In sum, Table 12 and 13 present empirical evidence that supports the theoretical prediction in Appendix A, i.e., selective return migration based on ethnic enclave residence did exist, and could partially contribute to differences in immigrant segregation between more and less restricted groups after the 1920s, although its magnitude was moderate.

6 Further Discussions

In this section, I briefly discuss potential effects of segregation on immigrants' assimilation. Researchers find that de-segregation is related to immigrants' socioeconomic assimilation (Massey and Denton, 1985), and in particular, there is a significant relationship between

immigrants’ de-segregation and economic outcomes when they actively make locational choices based on individual characteristics (e.g., Bleakley and Chin, 2010). In other words, the decline in immigrant segregation might suggest potential economic benefits. On the other hand, immigrants might choose to stay in ethnic neighborhoods to form ethnic social networks within small geographic areas, and further benefit from such within-group social relationships (Munshi, 2003). Hence, de-segregation might not necessarily lead to better outcomes for immigrants, in particular in the contexts of this paper that trends in segregation were not the consequences of immigrants’ choices, but the consequences of legislation against immigration that could not be decided by immigrants.

The above discussions lead to a policy-relevant question: did de-segregation lead to immigrants’ assimilation? Before immigration restriction laws were made, a major concern among the natives—especially politicians—was that immigrants refused to assimilate into the mainstream society, and one of the main reasons was segregation. In this section, I investigate effects of segregation on literacy and English proficiency at the county level, which are two major measures of immigrants’ assimilation.

Table 14: The Effects of Segregation: Literacy Rate

	First-generation immigrants			Second-generation immigrants		
	Italian	Polish	Russian	Italian	Polish	Russian
	(1)	(2)	(3)	(4)	(5)	(6)
Average, 1920:	0.808	0.892	0.898	0.988	0.989	0.993
De-segregated counties, 1900	-0.032* (0.016)	0.010 (0.014)	0.004 (0.010)	-0.001 (0.007)	0.009 (0.006)	0.003 (0.003)
De-segregated counties, 1910	-0.003 (0.015)	0.010 (0.019)	0.007 (0.009)	0.001 (0.006)	0.001 (0.008)	0.001 (0.003)
De-segregated counties, 1930	0.003 (0.014)	0.008 (0.011)	0.009 (0.009)	0.002 (0.005)	0.001 (0.004)	0.006* (0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,492	5,353	8,188	5,363	4,822	7,249

Standard errors are in parentheses and are clustered at the county level. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

To examine this, I focus on specific immigrant groups that are restricted (i.e., Italian, Polish, and Russian), and compare the literacy (both speak and read) rate and the ability of English speaking between counties experiencing the rise and decline in segregation,

before and after the 1920s. In other words, I run difference-in-differences estimations for counties grouped by trends in segregation. Table 14 shows results of the literacy rate for first-generation and second-generation immigrants (literacy was not surveyed in the 1940 census): in general, I find no significant effects of segregation on the literacy rate at the county level in both models. The only exception is that de-segregation improved the literacy rate among second-generation Russian immigrants, but the effect size is very small.

Table 15: The Effects of Segregation: English Speaking

	First-generation immigrants			Second-generation immigrants		
	Italian	Polish	Russian	Italian	Polish	Russian
	(1)	(2)	(3)	(4)	(5)	(6)
Average, 1920:	0.859	0.893	0.908	0.968	0.966	0.972
De-segregated counties, 1900	-0.027 (0.014)	0.001 (0.013)	0.009 (0.009)	0.001 (0.007)	-0.001 (0.007)	-0.002 (0.005)
De-segregated counties, 1910	0.001 (0.017)	0.046*** (0.010)	-0.023 (0.012)	0.004 (0.006)	-0.001 (0.009)	0.003 (0.004)
De-segregated counties, 1930	0.008 (0.009)	0.021* (0.008)	0.016* (0.006)	0.001 (0.006)	-0.004 (0.005)	-0.001 (0.004)
De-segregated counties, 1940	0.032*** (0.009)	0.043*** (0.008)	0.021** (0.007)	0.001 (0.005)	-0.004 (0.005)	-0.001 (0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,151	6,901	10,280	6,610	6,018	8,855

Standard errors are in parentheses and are clustered at the county level. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

In Table 15 I turn to examine English speaking, and find that de-segregation had positive effects on Russian immigrants' English skills. I also observe a higher rate of English speaking proficiency among Italian immigrants, but only in 1940. The results are less clear among Polish immigrants (pre-trends are also significant, and have larger coefficients). No effects are found among second-generation immigrants. In general, Table 14 and 15 suggest that de-segregation might affect assimilation patterns for some immigrants, but the overall effects of segregation on immigrants' assimilation appear to be unclear. This is consistent with the earlier argument: the decline in immigrant segregation was the result of the passage of immigration restriction laws in the early twentieth century U.S., which was not chosen by immigrants themselves; because of this, de-segregation might not necessarily accelerate immigrants' assimilation into the U.S. society.

7 Conclusion

Immigrant segregation is an important topic in population economics and urban economics. Cutler et al., (2008) find that immigrant segregation in the U.S. declined in the first two decades of the twentieth century, but started to rise afterwards. Was immigration legislation in the early twentieth century U.S. associated with immigration segregation? In this paper, I focus on immigration restriction laws that came into effect in the 1920s, and study differences in segregation patterns between more and less restricted groups.

The empirical analysis of this paper shows that both more and less restricted groups experienced similar trends in segregation before the passage of immigration restriction laws. Afterwards, however, two types of groups followed different trends in segregation. Specifically, segregation among immigrants from more restricted countries declined relative to segregation among immigrants from less or not restricted countries. The results are robust to changes to samples and specifications.

In addition, I discuss three possible behavioral mechanisms behind the effects of immigration restriction laws on immigrant segregation. A clear explanation of the differences in segregation patterns between more and less restricted groups after the 1920s is that immigration restriction laws had different quota rules across groups, and immigration restriction laws thus generated direct effects on immigrant segregation through restrictions on new immigrants, who were historically more likely to reside in ethnic enclaves (Haines, 2000). I further investigate two possible indirect mechanisms related to spatial behaviors of old immigrants who had settled down in the U.S., through which immigration laws might indirectly affect trends in segregation. First, I examine the effects of internal migration among immigrants who had arrived in the U.S. before the 1920s. I find no significant evidence that internal migration of remaining immigrants led to differences in segregation patterns between more and less restricted groups. Second, I find some evidence that return migration was negatively selected in terms of ethnic enclave residence, i.e., after immigration restriction laws were made, immigrants from more restricted countries were slightly more likely

to remain in the U.S. if they previously lived in areas with lower shares of immigrants of the same origin (and higher shares of natives). This is similar to selective return migration based on economic status (e.g., Ward, 2017; Abramitzky et al., 2017).

Finally, I briefly discuss consequences of the decline in segregation caused by the passage of immigration restriction laws. Focusing on the literacy rate and the rate of English speaking proficiency, I find effects of de-segregation in some models, but the overall effects of trends in segregation on immigrants' assimilation appear to be less clear. This might suggest that de-segregation, as a result of the passage of immigration restriction laws rather than immigrants' choices, might not result in immigrants' assimilation in the early twentieth century U.S.

Appendix A: A Theoretical Framework of the Effects of Immigration Restriction Laws on Immigrant Segregation

I consider a simple analytical framework of immigrants' settlement patterns following the analysis in Section 2.2. Suppose that an immigrant i considers his utility $U_i = u_i - u_{i0}$, where u_i is i 's utility in the U.S., u_{i0} is the "reservation utility" for i in his home country, and i stays in the U.S. if $U_i \geq 0$. Following the theoretical model of Edin et al. (2003) and Munshi (2003), I assume that

$$u_i = f(e_i) + z_i \tag{8}$$

where e_i is the measure of ethnic enclave residence, say, a continuum of locational choice such that $e_i \in (0, 1)$: under this setting, e_i represents the share of immigrants of the same origin at the locale. $f(e_i)$ is the utility that i gains from ethnic enclave residence, and individual i maximizes u_i subject to a "market opportunity locus" (Edin et al., 2003):

$$z_i = \mu_i + g(e_i, y_i) \tag{9}$$

where μ_i reflects i 's individual characteristics unrelated to e_i ; $g(e_i, y_i)$ measures the utility that i gains (e.g., labor market outcomes, marriage outcomes) through interactions—with both immigrants of the same origin and natives—in the U.S., where y_i is i 's years since migration. I assume that $\frac{\partial f(e_i)}{\partial e_i} > 0$ and $\frac{\partial g(e_i, y_i)}{\partial y_i} > 0$. Furthermore, for $g(e_i, y_i)$, I assume that $\frac{\partial g(e_i, y_i)}{\partial e_i} \Big|_{y_i=0} > 0$, $\frac{\partial g(e_i, y_i)}{\partial e_i} \Big|_{y_i=+\infty} < 0$, and $\frac{\partial g(e_i, y_i)}{\partial e_i \partial y_i} < 0$.⁴ These assumptions are based on that (a) immigrants have incentives to reside in ethnic enclaves (hence $\frac{\partial f(e_i)}{\partial e_i}$ is positive); (b) immigrants socioeconomically assimilate over time (hence $\frac{\partial g(e_i, y_i)}{\partial y_i}$ is positive); (c) When immigrants are more assimilated, they are more likely to benefit from interactions with the majority group, instead of compatriots (hence $\frac{\partial g(e_i, y_i)}{\partial e_i \partial y_i}$ is negative). The optimal e_i^* satisfies:

$$\frac{\partial f(e_i^*)}{\partial e_i^*} + \frac{\partial g(e_i^*, y_i)}{\partial e_i^*} = 0 \quad (10)$$

The above equation suggests e_i^* should not exist if y_i is sufficiently small, and u_i is an increasing function of e_i . Therefore, newcomers are more likely to reside in ethnic enclaves. If so, restrictions on new immigration should lead to the decrease in segregation. Figure 2 in Section 2.2 presents illustrations of this direct mechanism.

To further study two indirect mechanisms (i.e., internal migration and return migration) behind effects of immigration laws, I now consider a more general framework, in which $U_i' = u_i - u_{i0} - c(e_i, y_i)$. This new term $c(e_i, y_i)$ represents (negative) shocks of immigration laws against more restricted groups, where $c(e_i, y_i) > 0$. I assume that $\frac{\partial c(e_i, y_i)}{\partial y_i} < 0$ and $\frac{\partial c(e_i, y_i)}{\partial e_i} < 0$: in other words, I assume that (a) less assimilated immigrants should be more likely to be affected; (b) immigrants living in ethnic enclaves are less exposed to the majority group (and thus their hostility or discrimination). Now the new optimal e_i^+ satisfies:

$$\frac{\partial f(e_i^+)}{\partial e_i^+} + \frac{\partial g(e_i^+, y_i)}{\partial e_i^+} - \frac{\partial c(e_i^+, y_i)}{\partial e_i^+} = 0 \quad (11)$$

⁴Upon arrival, immigrants first rely on their ethnic social networks (e.g., Edin et al., 2003; Munshi, 2003; Damm, 2009) for support, but then receive higher utility by interacting with the majority group in the host society (e.g., Massey and Denton, 1985; Bartel, 1989; Bleakley and Chin, 2010); such a pattern of spatial assimilation is a function of years since immigration.

Is internal migration associated with e_i ? If y_i is sufficiently small, U'_i is still an increasing function of e_i . This suggests recent immigrants might be better off by moving to ethnic enclaves following the passage of immigration laws. However, if both e_i^* (in Equation 10) and e_i^+ (in Equation 11) exist, it is theoretically impossible to compare e_i^* with e_i^+ without knowing f , g , and c . In fact, one can construct different functional forms such that $e_i^* > e_i^+$ or $e_i^* < e_i^+$. Hence, it is unclear whether and how immigration laws could affect segregation through the channel of internal migration.

On the other hand, it is possible to predict selective return migration. Immigrant i will not remain in the U.S. if $U'_i < 0$. Since $\frac{\partial U'_i}{\partial y_i} > 0$, more recent immigrants (for whom e_i^* or e_i^+ might not exist) have lower (maximized) utility than older immigrants, and are more likely to leave the U.S. If so, return migration should be positively related to recent immigrants (who prefer ethnic enclave residence), and thus immigration laws should further lead to the decrease in segregation, as ethnic enclaves decline.

Following the above analyses, Figure 3 and 4 in Section 2.2 present illustrations of two indirect mechanisms. In the right figure of Figure 3, some immigrants from more restricted countries (stars) moved to other areas as the response to immigration restriction laws. In the right figure of Figure 4, immigrants from more restricted countries were more likely to remain in the U.S. if they previously lived in less segregated areas. In both figures, segregation among immigrants from more restricted countries became relatively lower. In Section 5 I empirically discuss all of the above three mechanisms, and test whether they could explain differences in segregation patterns between more and less restricted groups.

Appendix B: Individual Characteristics

In Section 3 I report descriptive statistics of the aggregate county-level data. In this appendix, I report descriptive statistics of individual census records by country of origin. In this paper I cluster three Scandinavian populations into one; statistics of both the whole

Scandinavian population and each sub-population are reported.

Table 16: Descriptive Statistics of Individual Characteristics: the 1900 Census

	Italian	Polish	Russian	British	Canadian	Irish	Scand.	Norwegian	Danish	Swedish
Age	31.035 (14.206)	31.766 (14,315)	29.305 (13.997)	42.065 (17.095)	35.367 (16.441)	45.548 (16.581)	38.635 (15.159)	40.994 (16.046)	38.379 (14.873)	37.341 (14.531)
Years since migration	8.749 (7.701)	11.048 (8.054)	10.035 (6.894)	23.676 (14.922)	17.794 (13.602)	26.887 (15.871)	17.596 (10.438)	20.247 (11.686)	17.446 (10.025)	16.105 (9.429)
Married	0.572	0.569	0.543	0.626	0.559	0.538	0.613	0.628	0.659	0.593
Female	0.351	0.422	0.461	0.453	0.485	0.541	0.441	0.443	0.412	0.448
Homeowner	0.115	0.275	0.171	0.393	0.344	0.373	0.546	0.640	0.543	0.492
Citizen	0.628	0.720	0.784	0.865	0.800	0.903	0.832	0.821	0.832	0.838
Occupational scores	13.627 (12.504)	12.745 (12.207)	13.074 (14.111)	13.969 (14.004)	12.451 (13.549)	11.506 (12.695)	11.457 (11.728)	10.457 (11.228)	11.812 (12.171)	11.920 (11.863)
Urban	0.752	0.763	0.813	0.626	0.592	0.759	0.445	0.308	0.427	0.528
Farm	0.029	0.086	0.115	0.146	0.163	0.093	0.367	0.496	0.398	0.284

Standard deviations are in parentheses. Standard deviations of binary variables are omitted.

Table 17: Descriptive Statistics of Individual Characteristics: the 1910 Census

	Italian	Polish	Russian	British	Canadian	Irish	Scand.	Norwegian	Danish	Swedish
Age	30.582 (13.488)	30.337 (13.402)	29.585 (13.492)	43.048 (17.610)	39.400 (16.932)	46.694 (16.785)	41.148 (15.692)	41.542 (16.755)	41.351 (15.505)	40.852 (15.046)
Years since migration	9.138 (8.365)	7.377 (8.272)	9.27 (8.441)	23.851 (16.408)	21.169 (14.587)	27.345 (16.817)	20.645 (13.023)	21.333 (14.380)	21.002 (12.852)	20.129 (12.134)
Married	0.574	0.583	0.572	0.619	0.607	0.521	0.611	0.586	0.657	0.613
Female	0.343	0.414	0.421	0.451	0.495	0.546	0.428	0.424	0.395	0.439
Homeowner	0.177	0.261	0.218	0.390	0.381	0.369	0.542	0.579	0.547	0.517
Citizen	0.574	0.552	0.671	0.852	0.834	0.905	0.837	0.821	0.838	0.847
Occupational scores	14.606 (12.509)	11.597 (11.032)	11.620 (13.414)	14.721 (14.789)	13.631 (14.440)	12.720 (13.919)	13.035 (13.056)	12.447 (12.826)	13.455 (13.334)	13.284 (13.109)
Urban	0.772	0.811	0.865	0.706	0.687	0.830	0.523	0.423	0.478	0.596
Farm	0.028	0.034	0.055	0.100	0.130	0.061	0.290	0.369	0.332	0.230

Standard deviations are in parentheses. Standard deviations of binary variables are omitted.

I start with the 1900 census in Table 16. In the 1900 census, immigrants from more restricted countries were significantly younger than immigrants from less restricted countries. They also arrived in the U.S. much later. This explains the large effects of the 1924 Act on immigration by country of origin, as the quota system used in the 1924 Act was based on the number of immigrants by country of origin in 1890, when most immigrants from more restricted countries had not arrived in the U.S. yet. The marriage rate was around 60% in all immigrant groups. Immigrants from more restricted countries, especially Italian immigrants, were substantially more likely to be male, less likely to own houses, and less

likely to be U.S. citizens. Note that in contrast to all other groups, more than half of Irish immigrants were female. The homeownership rate was substantially higher among Scandinavian immigrants. Among all immigrant groups, Italian and British immigrants had higher occupational scores, while Irish and Scandinavian immigrants had relatively low occupational scores. Immigrants from more restricted countries were much more likely to reside in cities. Immigrants from less restricted countries were more likely to live in rural areas, and in particular, less than half of Scandinavian immigrants lived in cities. They also had the higher rate of farm living.

Table 18: Descriptive Statistics of Individual Characteristics: the 1920 Census

	Italian	Polish	Russian	British	Canadian	Irish	Scand.	Norwegian	Danish	Swedish
Age	35.340 (14.001)	36.833 (13.954)	35.206 (13.866)	45.505 (17.245)	42.434 (18.176)	48.729 (15.846)	46.494 (15.515)	46.812 (16.259)	46.089 (15.752)	46.427 (14.989)
Years since migration	14.995 (8.792)	16.705 (10.796)	15.811 (8.972)	26.196 (16.419)	24.629 (15.785)	29.453 (16.140)	26.637 (13.867)	27.290 (14.865)	26.246 (14.178)	26.373 (13.139)
Married	0.686	0.750	0.692	0.646	0.617	0.556	0.666	0.651	0.688	0.668
Female	0.406	0.435	0.446	0.476	0.510	0.561	0.439	0.444	0.396	0.449
Homeowner	0.115	0.361	0.276	0.393	0.344	0.373	0.546	0.640	0.543	0.492
Citizen	0.628	0.736	0.522	0.865	0.800	0.903	0.832	0.821	0.832	0.838
Occupational scores	11.256 (12.788)	11.159 (12.606)	11.768 (14.647)	10.248 (13.763)	9.768 (13.593)	8.757 (12.468)	9.761 (12.317)	9.215 (11.907)	10.056 (12.599)	9.993 (12.457)
Urban	0.846	0.829	0.884	0.757	0.723	0.863	0.564	0.471	0.534	0.626
Farm	0.030	0.059	0.057	0.076	0.111	0.043	0.257	0.328	0.282	0.209

Standard deviations are in parentheses. Standard deviations of binary variables are omitted.

In Table 17 I present statistics of the 1910 census. The average age of immigrants from more restricted countries slightly declined, in contrast to that in all other groups. This suggests that a large number of new (and younger) immigrants might move to the U.S. from more restricted countries between 1900 and 1910. Similarly, I find no significant increase in years since migration among immigrants from more restricted countries. In 1910, on average, most of individual characteristics were similar to those in 1900. It is worth mentioning that Scandinavian immigrants became more likely to live in urban areas, and less likely to lead farming lives in 1910, compared with the 1900 statistics. Many Scandinavian immigrants were traditionally farmers after arrival, and the rise in urban residence among Scandinavian immigrants reflects the process of urbanization.

Table 19: Descriptive Statistics of Individual Characteristics: the 1930 Census

	Italian	Polish	Russian	British	Canadian	Irish	Scand.	Norwegian	Danish	Swedish
Age	41.024 (13.488)	42.269 (12.761)	41.744 (13.451)	46.160 (17.610)	41.718 (19.434)	48.157 (16.505)	50.216 (15.970)	50.005 (16.579)	49.644 (16.080)	50.511 (15.582)
Years since migration	20.891 (10.190)	22.549 (10.233)	22.637 (10.133)	26.353 (17.903)	24.496 (17.796)	28.711 (17.511)	30.871 (15.880)	31.005 (16.716)	30.269 (15.921)	30.975 (15.354)
Married	0.774	0.778	0.767	0.652	0.580	0.540	0.541	0.639	0.681	0.648
Female	0.418	0.463	0.468	0.490	0.518	0.562	0.432	0.440	0.389	0.441
Homeowner	0.489	0.528	0.385	0.462	0.443	0.452	0.597	0.591	0.588	0.602
Citizen	0.515	0.513	0.631	0.666	0.553	0.707	0.752	0.739	0.777	0.751
Occupational scores	11.433 (13.186)	11.410 (13.372)	12.462 (15.610)	10.810 (13.979)	9.922 (13.722)	9.219 (12.550)	10.338 (12.783)	9.861 (12.422)	10.899 (13.159)	10.448 (12.867)
Urban	0.878	0.863	0.902	0.803	0.767	0.890	0.631	0.558	0.595	0.685
Farm	0.030	0.055	0.053	0.054	0.086	0.029	0.211	0.266	0.240	0.169

Standard deviations are in parentheses. Standard deviations of binary variables are omitted.

I report descriptive statistics of the 1920 and 1930 census in Table 18 and 19, respectively. It is particularly useful to compare individual characteristics in the 1920 and 1930 census, before and after immigration restriction laws came into effect. I observe the significant increase in years since migration in most immigrant groups. However, the magnitude of the increase in years since migration was largest among immigrants from more restricted countries. Similarly, the average age of immigrants from more restricted countries significantly rose during the 1920s, which was opposite to the small or insignificant change in the average age in less restricted groups. These indirectly suggest the lack of new immigrants from Eastern and Southern European countries due to restrictions, and most immigrants from more restricted countries who were in the U.S. after the 1920s were old immigrants who had settled down before immigration restriction laws were passed. Immigration laws could have long-term impacts on differences in demographic characteristics among immigrant groups: Table 20 exactly presents this pattern in the 1940 census.

Appendix C: Record Linkage between Two Censuses

Record linkage has recently been a widely adopted technique in social statistics and economic history, and scholars develop various methods to link individual records across sam-

Table 20: Descriptive Statistics of Individual Characteristics: the 1940 Census

	Italian	Polish	Russian	British	Canadian	Irish	Scand.	Norwegian	Danish	Swedish
Age	48.178 (13.481)	49.916 (12.235)	49.513 (12.509)	51.679 (16.097)	46.648 (18.423)	53.197 (15.394)	56.363 (14.784)	55.641 (15.252)	55.583 (14.951)	57.020 (14.411)
Married	0.781	0.764	0.783	0.684	0.637	0.585	0.652	0.641	0.684	0.649
Female	0.426	0.474	0.474	0.502	0.536	0.570	0.437	0.444	0.395	0.449
Homeowner	0.464	0.467	0.333	0.451	0.436	0.402	0.574	0.561	0.566	0.583
Citizen	0.644	0.623	0.719	0.763	0.677	0.803	0.820	0.809	0.843	0.819
Occupational scores	13.473 (13.621)	13.162 (13.928)	15.036 (12.268)	13.099 (14.734)	13.360 (14.419)	10.493 (13.396)	11.736 (13.648)	11.349 (13.379)	12.746 (14.041)	11.652 (13.667)
Urban	0.854	0.838	0.887	0.753	0.699	0.846	0.624	0.573	0.594	0.664
Farm	0.030	0.059	0.047	0.052	0.081	0.027	0.185	0.226	0.206	0.155

Standard deviations are in parentheses. Standard deviations of binary variables are omitted.

The 1940 U.S. census did not survey the year of immigration, hence I do not report years since migration in this table.

ples (e.g., Jaro, 1989; Winkler, 1990; Ferrie, 1996; Ruggles, 2008; Goeken et al., 2011; Abramitzky et al., 2014; Collins and Wanamaker, 2015; Feigenbaum, 2017). In this paper, I use the stata command *Reclink2*, developed by the University of Michigan (Wasi and Flaaen, 2015; Bailey et al., 2017), to link individual immigrants in the 1920 and 1930 U.S. census. *Reclink2* is a command for multivariate linkage, which allows the comparison of various individual characteristics between potential matches.

I first construct subsamples of “linkable populations” by immigrant group in the 1930 census (i.e., those who had been in the U.S. before 1920), and search these linkable individuals in the 1920 census. I only attempt to link male immigrants between two censuses, as women usually change their surnames after marriage. I use the following linking variables in the linkage process: first name, last name (also with the truncated parts of the last name), age, year of immigration, and mother tongue. In the algorithm, I set that the mother tongue and the first letter of the last name are required to be perfectly matched between individuals in two samples. However, a two-year window of mismatch is allowed for age and year of immigration. The core variable of linkage is name; *Reclink2* uses the bi-gram distance to measure two strings. I include the first and last name in the algorithm separately, and also include the truncated parts of the last name—the first and last few characters—as linking variables. After obtaining the potential links for each 1930-census individual, I only accept the best match from the 1920 census that has a substantially higher linkscore than the

second-best match.⁵

Table 21: Checking Linkage Quality (1)

	Degree of segregation			# same origin, ED			# native, ED		
	Linkable population	Linked sample	<i>p</i> -value	Linkable population	Linked sample	<i>p</i> -value	Linkable population	Linked sample	<i>p</i> -value
Italian	0.52 (0.092)	0.52 (0.091)	1.000	228 (228)	228 (224)	1.000	1,329 (777)	1,321 (777)	<0.001
Observations	842,104	135,369		842,104	135,369		842,104	135,369	
British	0.27 (0.087)	0.27 (0.076)	1.000	55 (77)	55 (68)	0.008	1,417 (798)	1,421 (784)	0.150
Observations	463,796	99,368		463,796	99,368		463,796	99,368	
Canadian	0.25 (0.094)	0.25 (0.085)	1.000	130 (174)	129 (168)	0.078	1,354 (769)	1,357 (767)	0.235
Observations	449,843	115,078		449,843	115,078		449,843	115,078	
Irish	0.35 (0.081)	0.34 (0.070)	< 0.001	76 (92)	75 (89)	0.040	1,334 (698)	1,340 (675)	0.104
Observations	312,816	40,083		312,816	40,083		312,816	40,083	
Norwegian	0.39 (0.144)	0.38 (0.137)	< 0.001	88 (115)	83 (96)	0.008	977 (725)	943 (727)	0.003
Observations	163,791	39,490		163,791	39,490		163,791	39,490	
Danish	0.43 (0.034)	0.43 (0.036)	1.000	57 (81)	58 (82)	0.153	1,144 (14,879)	1,165 (15,068)	0.002
Observations	92,418	15,317		92,418	15,317		92,418	15,317	
Swedish	0.039 (0.114)	0.039 (0.115)	1.000	105 (147)	101 (137)	< 0.001	1,201 (747)	1,196 (768)	0.189
Observations	332,067	43,943		332,067	43,943		332,067	43,943	

Standard deviations are in parentheses.

The linkage rate in each subsample is: 16.1% for Italian immigrants, 21.4% for British immigrants, 25.6% for Canadian immigrants, 12.8% for Irish immigrants, 24.1% for Norwegian immigrants, 16.6% for Danish immigrants, and 13.2% for Swedish immigrants (overall, 16.8% for Scandinavian immigrants). The overall linkage rate is comparable to similar linkage projects (e.g., Goeken et al., 2011; Abramitzky et al., 2014; Bailey et al., 2017). Note that in this paper I report the “crude” linkage rate calculated based on the full linkable population, while some papers report the linkage rate after excluding individuals with the same names whose potential matches cannot be distinguished and selected in the linking algorithm (e.g., Collins and Wanamaker, 2014), which will be higher than the crude linkage rate reported here.

While it is generally impossible to link all linkable individuals, the linked sample should be a representative subsample of the linkable population (Collins and Wanamaker, 2014;

⁵The highest linkscore is 1.00 (perfectly matched) and I set the lowest possible linkscore as 0.7 (the lower threshold). To be selected in the linked sample, the best match must have a linkscore that is 0.1 higher than that of the second-best match.

Table 22: Checking Linkage Quality (2)

	% same origin, ED			% native, ED			Occ. scores		
	Linkable population	Linked sample	<i>p</i> -value	Linkable population	Linked sample	<i>p</i> -value	Linkable population	Linked sample	<i>p</i> -value
Italian	0.127 (0.114)	0.127 (0.112)	1.000	0.725 (0.119)	0.722 (0.114)	< 0.001	18.659 (12.614)	18.636 (13.056)	0.536
Observations	842,104	135,369		842,104	135,369		842,104	135,369	
British	0.030 (0.034)	0.030 (0.028)	1.000	0.815 (0.110)	0.815 (0.101)	1.000	18.443 (14.879)	18.482 (14.885)	0.453
Observations	463,796	99,368		463,796	99,368		463,796	99,368	
Canadian	0.071 (0.075)	0.070 (0.072)	< 0.001	0.800 (0.101)	0.800 (0.097)	1.000	18.145 (15.178)	18.280 (15.598)	0.007
Observations	449,843	115,078		449,843	115,078		449,843	115,078	
Irish	0.044 (0.049)	0.044 (0.048)	1.000	0.765 (0.122)	0.763 (0.108)	0.002	16.951 (14.879)	17.134 (15.068)	0.014
Observations	312,816	40,083		312,816	40,083		312,816	40,083	
Norwegian	0.083 (0.070)	0.083 (0.061)	1.000	0.808 (0.103)	0.822 (0.087)	< 0.001	15.979 (12.803)	15.981 (13.059)	0.978
Observations	163,791	39,490		163,791	39,490		163,791	39,490	
Danish	0.047 (0.054)	0.048 (0.056)	0.035	0.830 (0.101)	0.830 (0.097)	1.000	16.556 (13.364)	16.587 (13.399)	0.791
Observations	92,418	15,317		92,418	15,317		92,418	15,317	
Swedish	0.030 (0.034)	0.031 (0.036)	< 0.001	0.816 (0.110)	0.817 (0.105)	0.008	18.443 (14.879)	18.598 (15.068)	0.003
Observations	332,067	43,943		332,067	43,943		332,067	43,943	

Standard deviations are in parentheses.

Bailey et al., 2017). In other words, ideally, there should be very small differences in individual characteristics between the linked individual panel and the full linkable population, so that results based on the linked sample can be safely “extrapolated” to the linkable population in the 1930 census. In Table 21 and 22 I compare several key variables relevant to the historical context of this paper. In Table 21 I examine the degree of segregation, the number of immigrants of the same origin in the enumeration district, and the number of natives in the enumeration district. In Table 22 I examine the share of immigrants of the same origin and the share of natives in the enumeration district, and occupational scores.

In Table 21 I first compare the share of immigrants of the same origin in the enumeration district between the linkable population in the 1930 census and the linked individual panel. I find almost no significant differences in the degree of segregation between the linkable population and the linked sample, with the exception of the Irish and Norwegian group. I similarly find small differences in the number of immigrants of the same origin and the number of natives in the enumeration district; although some differences are significant, the magnitudes of the differences are very small. In Table 22 I focus on the share of immigrants

of the same origin, the share of natives in the enumeration district, and occupational scores. Again, the differences in these characteristics between two samples are either insignificant or of very small magnitudes. These two tables suggest the linked individual panel dataset is a fairly representative subsample of the linkable population in the 1930 census.

Appendix D: Placebo Tests in Section 4.2

In this appendix, I present placebo tests introduced in Section 4.2. The basic idea of placebo tests is to examine whether immigration laws indeed led to different trends in segregation between more and less restricted groups. If so, then one should expect results of placebo tests different from the main findings, as “fake” restrictions were “assigned” to other groups that were less or not restricted in actual history.

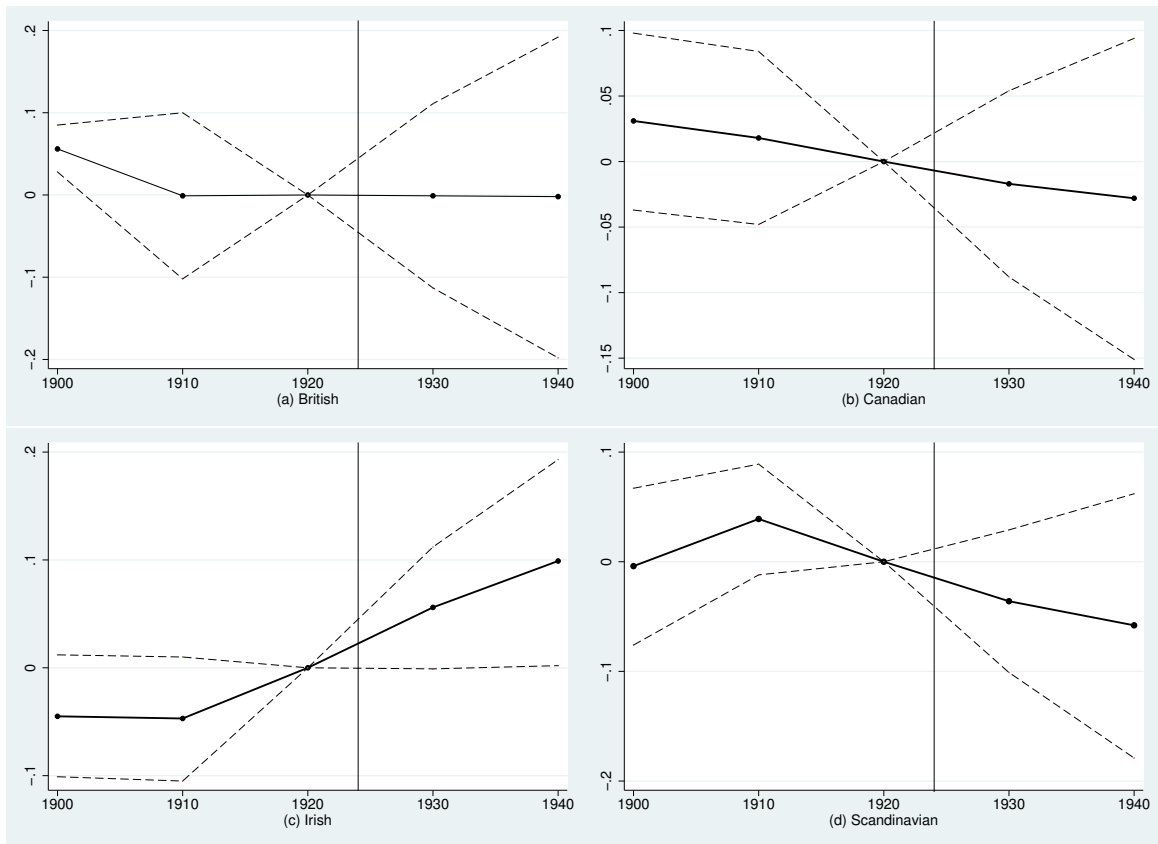


Figure 7: Placebo Tests: “Fake” Restrictions on Other Immigrant Groups

Figure 7 presents the results. I run regressions similar to the model in the main table, but assuming that British, Canadian, Irish, and Scandinavian immigrants as the specific more restricted group. I find almost all estimates insignificant, suggesting that immigration laws had the particular effect on the Italian immigrants. Figure 7(c) does present some significant results for Irish immigrants. However, the shape of the trend in Irish segregation appears to be opposite to that shown in Figure 5. At the very least, immigration restriction laws had very different effects on segregation between Irish immigrants and immigrants from more restricted countries in actual history. Also, In Section 4 I show that excluding Irish immigrants from the sample does not affect the conclusion of this paper. These suggest that after the 1920s, differences in segregation patterns between more and less restricted groups did exist, and were very likely to be related to immigration restriction laws.

Appendix E: Additional Tests in Section 5.1

I now present an additional table following the discussion in Section 5.1, where I analyze spatial assimilation, i.e., new immigrants started to move out and spatially assimilated with the increase in years in migration. In Section 5.1, I control for state fixed effects in the regressions. Table 23 presents similar results when I control for county fixed effects.

Appendix F: Additional Tests in Section 5.2

In Section 5.2, I show that internal migration—including both inter-state and within-state migration—had no overall effects on trends in segregation by group. I now further test inter-state migration and within-state inter-county migration separately.

In Table 24 I examine inter-state migration. In both panels, I find no significant association between inter-state migration and the degree of segregation. I do find some significant but small effects of migration from Column 3 to 6: Italian immigrants moved to places with relatively higher shares of immigrants of the same origin, and lower shares of natives.

Table 23: Years since Migration and Ethnic Enclave Residence (2): County Controls

	% of Immigrants of the Same Origin, at Enumeration District (ED) Level					
	(1)	(2)	(3)	(4)	(5)	(6)
	Italian	British	Canadian	Irish	Scandinavian	Pooled
1920 census:						
Average:	0.171	0.036	0.080	0.055	0.087	0.092
Decades since migration	-0.013*** (0.003)	-0.002*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.003*** (0.000)	-0.002*** (0.001)
(Decades since migration)*Italy						-0.009*** (0.002)
Adjusted R ²	0.303	0.411	0.579	0.370	0.538	0.386
Observations	1,609,346	1,159,148	1,271,325	1,050,668	1,183,603	6,220,090
1930 census:						
Average:	0.133	0.033	0.072	0.050	0.070	0.078
Decades since migration	-0.007*** (0.001)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.004*** (0.001)	-0.002*** (0.000)
(Decades since migration)*Italy						-0.005*** (0.001)
Adjusted R ²	0.264	0.341	0.578	0.354	0.456	0.352
Observations	1,788,943	1,230,211	1,398,968	929,514	1,119,550	6,467,186
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors are in parentheses and are clustered at the county level. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

I report coefficients of decades since migration (i.e., years since migration are divided by 10). County FE are included.

This actually contradicts the main findings that the degree of Italian segregation became relatively lower over time.

I conclude the discussion of internal migration by focusing on within-state inter-county migration in Table 25. Results show that Italian immigrants who moved within the state generally resided in less segregated areas, as well as enumeration districts with lower shares of immigrants of the same origin and higher shares of natives. I find less significant effects after including less restricted groups. Note that this appears to be opposite to the effects of inter-state migration reported in Section 5.2. In sum, although I find significant results in some models, the overall effects of internal migration were insignificant and could not explain the main empirical conclusion of this paper.

Table 24: The Effects of Internal Migration: Inter-State Migration

	Degree of Segregation		% same origin, ED		% native, ED	
	(1)	(2)	(3)	(4)	(5)	(6)
Italian:						
$M_i \mathbb{1}(t > 1920)$	-0.002 (0.001)	0.002 (0.001)	0.001 (0.001)	-0.007** (0.001)	-0.002 (0.003)	0.003*** (0.001)
Adjusted R ²	0.093	0.367	0.109	0.169	0.186	0.326
Observations	270,692	270,692	270,692	270,692	270,692	270,692
All immigrants:						
$S_e M_i \mathbb{1}(t > 1920)$	0.003 (0.004)	0.001 (0.003)	0.011** (0.003)	0.007* (0.003)	-0.006*** (0.001)	-0.003* (0.002)
Adjusted R ²	0.628	0.723	0.205	0.286	0.257	0.388
Observations	980,698	980,698	980,698	980,698	980,698	980,698
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	No	Yes	No	Yes

Standard errors are in parentheses and clustered at the origin-by-year level. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

Table 25: The Effects of Internal Migration: Within-State Inter-County Migration

	Degree of Segregation		% same origin, ED		% native, ED	
	(1)	(2)	(3)	(4)	(5)	(6)
Italian:						
$M_i \mathbb{1}(t > 1920)$	-0.013*** (0.000)	-0.013*** (0.000)	-0.018*** (0.001)	-0.018** (0.004)	0.012** (0.002)	0.013* (0.003)
Adjusted R ²	0.095	0.361	0.109	0.161	0.200	0.320
Observations	167,194	167,194	167,194	167,194	167,194	167,194
All immigrants:						
$S_e M_i \mathbb{1}(t > 1920)$	-0.003 (0.004)	-0.004 (0.003)	-0.018*** (0.004)	-0.018*** (0.002)	0.005 (0.007)	0.006 (0.004)
Adjusted R ²	0.629	0.724	0.206	0.286	0.270	0.389
Observations	661,848	661,848	661,848	661,848	661,848	661,848
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	No	Yes	No	Yes

Standard errors are in parentheses and clustered at the origin-by-year level. *: $p < .05$; **: $p < .01$; ***: $p < .001$.

References

- [1] Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy*, 122(3), 467 - 506.
- [2] Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2017. "To the New World and Back Again: Return Migration and Upward Mobility." *Industrial and Labor Relations Review*. DOI: <https://doi.org/10.1177/0019793917726981>
- [3] Akay, Alpaslan, Amelie Constant, and Corrado Giulietti. 2014. "The Impact of Immigration on the Well-Being of Natives." *Journal of Economic Behavior & Organization*, 103, 72 - 92.

- [4] Altonji, Joseph G., and David Card. 1994. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives." in *Immigration, Trade, and the Labor Market*, edited by John M. Abowd and Richard B. Freeman. Chicago: University of Chicago Press.
- [5] Avitabile, Ciro, Irma Clots-Figueras, and Paolo Masella. 2013. "The Effect of Birthright Citizenship on Parental Integration Outcomes." *Journal of Law and Economics*, 56(3), 777 - 810.
- [6] Bartel, Ann P. 1989. "Where Do the New U.S. Immigrants Live?" *Journal of Labor Economics*, 7(4), 371 - 391.
- [7] Bailey, Martha J., Connor Cole, Morgan Henderson, and Catherine Massey. 2017. "How Well Do Automated Linking Methods Perform in Historical Samples? Evidence from New Ground Truth." working paper.
- [8] Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004 "How Much Should We Trust Differences-In-Differences Estimates?" *Quarterly Journal of Economics*, 119(1), 249 - 275.
- [9] Bleakley, Hoyt, and Aimie Chin. 2010. "Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants." *American Economic Journal: Applied Economics*, 2(1), 165 - 192.
- [10] Borjas, George J. 1987. "Immigrants, Minorities, and Labor Market Competition." *Industrial and Labor Relations Review*, 40(2), 382 - 392.
- [11] Borjas, George J., Richard B. Freeman, and Lawrence F. Katz. 1992. "On the Labor Market Effects of Immigration and Trade." in *Immigration and the Workforce: Economic Consequences for the United States and Source Areas*, edited by George J. Borjas and Richard B. Freeman. Chicago: University of Chicago Press.
- [12] Boyd, Robert L., and Xiaohe Xu. 2003. "Did Retail Enterprise Among White Immigrants Benefit from the Residential Segregation of Blacks? A Study of Large Northern Cities in the Early 20th Century." *Social Science Quarterly*, 84(4), 934 - 945.
- [13] Brown, M. Craig, and Barbara D. Warner. 1992. "Immigrants, Urban Politics, and Policing in 1900." *American Sociological Review*, 57(3), 293 - 305.
- [14] Cannato, Vincent J. 2010. *American Passage: The History of Ellis Island*, New York: Harper Perennial.
- [15] Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review*, 43(2), 245 - 257.
- [16] Cerase, Francesco P. 1974. "Expectations and Reality: A Case Study of Return Migration from the United States to Southern Italy." *International Migration Review*, 8(2), 245 - 262.
- [17] Collins, William J. 1997. "When the Tide Turned: Immigration and the Delay of the Great Black Migration." *Journal of Economic History*, 57(3), 607 - 632.

- [18] Collins, William J., and Marianne H. Wanamaker. 2014. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." *American Economic Journal: Applied Economics*, 6(1), 220 - 252.
- [19] Collins, William J., and Marianne H. Wanamaker. 2015. "The Great Migration in Black and White: New Evidence on the Selection and Sorting of Southern Migrants." *Journal of Economic History*, 75(4), 947 - 992.
- [20] Cutler, David M, Edward L. Glaeser, and Jacob L. Vigdor. 2008. "Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation." *Review of Economics and Statistics*, 90(3), 478 - 497.
- [21] Damm, Anna P. 2009. "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence." *Journal of Labor Economics*, 27(2), 281 - 314.
- [22] Dubin, Lois C. 1999. *The Port Jews of Habsburg Trieste: Absolutist Politics and Enlightenment Culture*. Palo Alto: Stanford University Press.
- [23] Duncan, Otis Dudley, and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indexes." *American Sociological Review*, 20(2), 210 - 217.
- [24] Duncan, Otis Dudley, and Stanley Lieberson. 1959. "Ethnic Segregation and Assimilation." *American Journal of Sociology*, 64(4), 364 - 374.
- [25] Edin, Per-Anders, Peter Fredriksson and Olof Åslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment." *Quarterly Journal of Economics*, 118(1), 329 - 357.
- [26] Feigenbaum, James. 2017. "Automated Census Record Linking: A Machine Learning Approach." manuscript.
- [27] Ferrie, Joseph P. 1996. "A New Sample of Males Linked from the Public Use Microdata Sample of the 1850 U.S. Federal Census of Population to the 1860 U.S. Federal Census Manuscript Schedules." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 29(4), 141 - 156.
- [28] Ferrie, Joseph P. 2003. "Internal Migration." in *Historical Statistics of the United States, Millennial Edition*, edited by Susan Carter et al. New York: Cambridge University Press.
- [29] Fox, Cybelle, and Thomas A. Guglielmo. 2012. "Defining Americas Racial Boundaries: Blacks, Mexicans, and European Immigrants, 1890 - 1945." *American Journal of Sociology*, 118(2), 327 - 379.
- [30] Goeken, Ron, Lap Huynh, Thomas Lenius, and Rebecca Vick. 2011. "New Methods of Census Record Linking." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 44(1), 7 - 14.
- [31] Goldin, Claudia. 1994. "The Political Economy of Immigration Restriction in the United States, 1890 to 1921." in *The Regulated Economy: A Historical Approach to Political Economy*, edited by Claudia Goldin and Gary D. Libecap. Chicago: University of Chicago Press.

- [32] Gordon, Milton M. 1964. *Assimilation in American Life: The Role of Race, Religion, and National Origins*. New York: Oxford University Press.
- [33] Greenwood, Michael J., and Zachary Ward. 2015. "Immigration Quotas, World War I, and Emigrant Flows from the United States in the Early 20th Century." *Explorations in Economic History*, 55, 76 - 96.
- [34] Gross, Dominique M., and Nicolas Schmitt. 2003. "The Role of Cultural Clustering in Attracting New Immigrants." *Journal of Regional Science*, 43(2), 295 - 318.
- [35] Haines, Michael R. 2000. "The Population of the United States, 1790 - 1920." In *The Cambridge Economic History of the United States*, eds., Stanley L. Engerman and Robert E. Gallman. New York: Cambridge University Press.
- [36] Hall, Patricia Kelly, and Steven Ruggles. 2004. "Restless in the Midst of Their Prosperity: New Evidence on the Internal Migration of Americans, 1850 - 2000." *Journal of American History*, 91(3), 829 - 846.
- [37] Hatton, Timothy J., and Jeffrey G. Williamson. 1998. *The Age of Mass Migration: Causes and Economic Impact*, New York: Oxford University Press.
- [38] Higham, John. 1963. *Strangers in the Land: Patterns of American Nativism, 1860 - 1925*, New Brunswick: Rutgers University Press.
- [39] Jaro, Matthew A. 1989. "Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida." *Journal of the American Statistical Association*, 84(406), 414 - 420.
- [40] Kanazawa, Mark. 2005. "Immigration, Exclusion, and Taxation: Anti-Chinese Legislation in Gold Rush California." *Journal of Economic History*, 65(3), 779 - 805.
- [41] Lleras-Muney, Adriana, and Allison Shertzer. 2015. "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory Schooling Laws on Immigrants." *American Economic Journal: Economic Policy*, 7(3), 258 - 290.
- [42] Logan, Trevon D., and John M. Parman. 2017. "The National Rise in Residential Segregation." *Journal of Economic History*, 77(1), 127 - 170.
- [43] Massey, Catherine G. 2016. "Immigration Quotas and Immigrant Selection." *Explorations in Economic History*, 60, 21 - 40.
- [44] Massey, Douglas S., and Nancy A. Denton. 1985. "Spatial Assimilation as a Socioeconomic Outcome." *American Sociological Review*, 50(1), 94 - 106.
- [45] Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. "Internal Migration in the United States." *Journal of Economic Perspectives*, 25(3), 173 - 196.
- [46] Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *Quarterly Journal of Economics*, 118(2), 549 - 599.

- [47] Moser, Petra. 2012. "Taste-Based Discrimination Evidence from a Shift in Ethnic Preferences after WWI." *Explorations in Economic History*, 49(2), 167 - 188.
- [48] Ngai, Mae M. 1999. "The Architecture of Race in American Immigration Law: A Reexamination of the Immigration Act of 1924." *Journal of American History*, 86(1), 67 - 92.
- [49] Ngai, Mae M. 2004. *Impossible Subjects: Illegal Aliens and the Making of Modern America*, Princeton: Princeton University Press.
- [50] Orrenius, Pia M., and Madeline Zavodny. 2006. "Does Immigration Affect Wages? A Look at Occupation-Level Evidence." *Labour Economics*, 14(5), 757 - 773.
- [51] Ottaviano, Gianmarco I. P., and Giovanni Peri. 2012. "Rethinking the Effects of Immigration on Wages." *Journal of the European Economic Association*, 10(1), 152 - 197.
- [52] Ruggles, Steven. 2008. "Linking Historical Censuses: a New Approach." *History and Computing*, 14(1-2), 213 - 224.
- [53] Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. *Integrated Public Use Microdata Series: Version 7.0* [Machine-readable database]. Minneapolis: University of Minnesota.
- [54] Sammartino, Annemarie H. 2010. *The Impossible Border: Germany and the East, 1914-1922*. Ithaca: Cornell University Press.
- [55] Shertzer, Allison, Randall P. Walsh, and John R. Logan. 2016. "Segregation and Neighborhood Change in Northern Cities: New Historical GIS Data from 1900 - 1930." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 49(4), 187 - 197.
- [56] van Nuys, Frank. 2002. *Americanizing the West: Race, Immigrants, and Citizenship, 1890-1930*, Lawrence: University Press of Kansas.
- [57] Vigdor, Jacob L. 2002. "The Pursuit of Opportunity: Explaining Selective Black Migration." *Journal of Urban Economics*, 51(3), 391 - 417.
- [58] Ward, Zachary. 2017. "Birds of Passage: Return Migration, Self-Selection and Immigration Quotas." *Explorations in Economic History*, 64, 37 - 52.
- [59] Wasi, Nada, and Aaron Flaaen. 2015. "Record Linkage using STATA: Pre-processing, Linking and Reviewing Utilities." *Stata Journal*, 15(3), 672 - 697.
- [60] Winkler, William E. 1990. "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage." *Proceedings of the Section of Survey Research Methods*, 354 - 359.
- [61] Winship, Christopher. 1977. "A Reevaluation of Indexes of Residential Segregation." *Social Forces*, 55(4), 1058 - 1066.

[62] Wyman, Mark. 1993. *Round-Trip to America: The Immigrants Return to Europe, 1880 - 1930*, Ithaca: Cornell University Press.