Tian Lou[1,*] *and Tao Song*[2]

Ethnic Segregation and Immigrants' Labor Market Outcomes: The Role of Education

Abstract

Given the conflicting findings about whether ethnic segregation benefits or hurts immigrants' labor market outcomes in the existing literature, this paper investigates how segregation effects vary within young immigrants' education levels. We also test the differential segregation effects for young immigrants with different education levels and from ethnic groups with different average education levels. We find that on average, ethnic segregation negatively impacts earnings of highly educated immigrants but benefits lower educated immigrants. Additionally, the net effects of ethnic segregation depend on whether immigrants' own education levels match their group average education levels. Specifically, being segregated with many highly educated co-ethnics can reverse the negative segregation effects on highly educated immigrants' earnings. However, a highly educated ethnic enclave can reduce the positive isolation effects for lower educated immigrants. Finally, we do not find significant segregation effects on immigrants' employment.

Current version: February 6, 2023
Keywords: Immigrants, Segregation, Labor market outcomes, Ethnic enclaves
JEL-codes: J15, J24, J31, J61
Corresponding author: Tian Lou
lou.151@osu.edu

1 Ohio Education Research Center, John Glenn College of Public Affairs, The Ohio State University, CHRR at The Ohio State University, 921 Chatham Lane, Suite 200, Columbus, OH 43211, United States

2 Department of Economics, Sewanee: The University of the South, 735 University Avenue, Sewanee, TN 37383, United States

© The Author(s). 2023 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
Cite as: Lou and Song. *IZA Journal of Labor Economics* (2023) 12:05.
https://doi.org/10.2478/izajole-2023-0005

1 Introduction

The number of immigrants in the U.S. has increased substantially since the 1960s. Researchers have observed that immigrants from the same ethnic groups tend to cluster together in the same geographic areas (Cutler et al., 2008). It is however a puzzle as to whether this residential segregation could positively or negatively affect immigrants' labor market outcomes. While some studies find that ethnic segregation is beneficial for immigrants (Cutler et al., 2008; Damm, 2009; Edin et al., 2003; Munshi, 2003), others show that ethnic segregation either leads to deteriorated labor market outcomes (Beaman, 2012; Borjas, 2000) or has no impact at all (Grönqvist, 2006, Song and Xu, 2020). What is more important is that ethnic segregation could impact immigrant members differently depending on their group characteristics. Edin et al. (2003) find that a higher average income of the ethnic group could positively affect enclave members' wages, and Cutler et al. (2008) find a similar wage effect, but only for those from the highly educated groups.

In this paper, we explore whether immigrants' own education levels and the ethnic groups' education levels might individually and interactively explain the puzzles on the effects of ethnic segregation. We first conjecture that the impacts of ethnic segregation might vary according to immigrants' individual education levels. For example, language barriers (Carliner, 2000), labor market discrimination (Edin et al., 2003), and overreliance on informal social networks in job searches (Borjas, 1998; Ioannides and Loury, 2004; Wahba and Zenou, 2005) could impede low education immigrants from seeking job opportunities outside of ethnic enclaves (Glitz, 2014). As a result, these immigrants could be more susceptible to segregation effects, relative to high education immigrants. On the other hand, high education immigrants, who for various reasons become ethnically segregated, could also be more negatively affected by segregation if their better matched jobs are outside of the enclaves.

Second, the differential segregation effects mentioned above may depend on the average education level of the immigrants' ethnic group. The direction of the effect, however, is also ambiguous. On the one hand, if immigrants' individual education levels are similar to their ethnic enclaves' average education levels, they may benefit from better job networks, which would either alleviate the negative impacts of segregation or reinforce the positive effects of segregation. For example, it is possible that highly educated immigrants' high group average education level may be correlated with their occupational specific skills such as computer programming or data analytical expertise. If both education and these talents are highly demanded in areas where certain industries concentrate, then ethnic segregation might create positive network effects or synergies that could help the highly educated immigrants (Cutler et al., 2008). On the other hand, matches between immigrants' individual education levels and group average education levels may negatively impact immigrants due to increased competition for jobs (Beaman, 2012; Cortes, 2008).

Since theories cannot reconcile which of the aforementioned effects dominate, this study uses empirical strategies to examine the puzzle of ethnic segregation. We argue that it remains difficult to empirically estimate impacts of segregation because immigrants are likely to endogenously sort geographically based on observed and unobserved factors. Research using neighborhood-level segregation intensity measures (such as percentage of immigrants from the same ethnic group in the same tract/block) often find that segregation harms immigrants' labor

market outcomes. This finding, however, might be driven by lower educated immigrants sorting into ethnically concentrated neighborhoods due to housing market discrimination (Ihlanfeldt and Sjoquist, 1998) or a lack of fluency in the host country's language (Carliner, 2000), among other factors. Therefore, a number of papers adopt an alternative method by exploiting policies that exogenously place refugees in host countries (Beaman, 2012; Damm, 2009; Edin et al., 2003). They nevertheless still find conflicting effects of ethnic enclaves on immigrants' labor market outcomes. While these results are less prone to endogeneity issues, they may not be generalizable enough to all immigrants because of potential unobserved factors specific to the refugee groups, the placement policies, and the host countries.

In this paper, instead of exploiting such exogenous variations in segregation intensity, we utilize two strategies to address the issues posed by immigrants' sorting behaviors. First, we use a Metropolitan Statistical Area (MSA) level isolation index, as opposed to neighborhood level indices, to measure segregation intensity. Within the same MSA, lower educated immigrants are more likely to stay in ethnic enclave neighborhoods, while highly educated immigrants may move to more diversified neighborhoods with better economic conditions.[1] Studies also suggest that it is more difficult for immigrants to move between MSAs (Cutler et al., 2008; Cutler and Glaeser, 1997; Evans et al., 1992). Therefore, the comparisons between MSAs could reduce the bias caused by endogenous residential location choices.

Our second strategy is to adopt a triple differences model to minimize the selection bias associated with MSA-level segregation intensity. Specifically, we control for MSA-specific time trends, group-specific time trends, and ethnicity-MSA-specific factors. Thus, our model takes into account any MSA or ethnic group factors that influence immigrants' outcomes over time, such as the overall wage increase in a MSA and the expansion of industries that hire a large number of workers from a specific ethnic group. Our model also minimizes any time invariant MSA factors that impact outcomes of immigrants from a certain ethnic group, for example, local job training programs that target specific immigrant groups.

Following the convention in the literature (e.g., Cutler et al., 2008), we use a sample of immigrants who are between ages 20-30 and who migrated to the U.S. at age 17 or later. The reason we focus on young immigrants is to minimize the bias caused by older immigrants endogenously sorting into ethnic enclaves of different education levels and different economic conditions.[2] Moreover, we focus on immigrants who arrived close to adulthood to reduce the likelihood of ethnic enclaves impacting immigrants' educational attainment.

To estimate the differential segregation effects, we introduce a set of double and triple interaction terms between the segregation intensity, immigrants' own education, and ethnic group average education. We find that, given the education level of the co-ethnic peers, the average segregation effect is positive for lower educated immigrants and negative for highly educated immigrants. However, when estimating the differential effects of segregation by both individual education and group average education, we find that low education immigrants only receive the benefits from ethnic segregation when being isolated with other lower educated

1 For example, Danzer and Yaman (2016) find that immigrants who have better host country language skills are less likely to ethnically segregate. Additionally, Furtado and Song (2015) find that immigrants with better observed and unobserved human capital tend to assimilate into host countries faster.

2 A balancing test using immigrants between ages 30-40 confirms our concern. We found that 30-40 years old immigrants with higher levels of education are less likely to segregate. More details about the balancing test will be discussed in Section 5.

ethnic peers. The negative segregation effects on highly educated immigrants can be reversed if they are living with many other highly educated co-ethnics. These findings support the theory of positive network effects from ethnic segregation. Finally, we do not find any significant segregation effects on young immigrant's employment prospects.[3]

This paper proceeds as follows. In Section 2, we discuss in greater detail how immigrants' individual education levels and ethnic group average education levels may influence segregation effects. In Section 3, we introduce the data used for our empirical estimations and the measure of segregation intensity. In Section 4, we discuss the triple interaction and triple differences model. In Section 5, we test the validity of our identification strategy. In Section 6, we present the results of our main model. In Section 7, we test the robustness of the earnings results. Section 8 presents the employment results. We conclude in the last section.

2 Why Immigrants' Individual Education Levels and Group Average Education Levels Matter?

2.1 Inconsistent Segregation Effects

In the current literature, there is no definite conclusion as to how ethnic segregation influences immigrants' labor market outcomes. Some papers showed evidence that immigrants gain labor market advantages from ethnic segregation. For example, using the natural experiment of the Swedish immigration policy, Edin et al. (2003) showed that, when sorting is taken into account, living in enclaves improves labor market outcomes for lower educated immigrants. Munshi (2003) used an instrumental variable approach and found that Mexican immigrants with larger network sizes in the host country have better labor market outcomes. Damm (2009) showed that immigrants' hourly wages increase with enclave size when refugees were randomly dispersed throughout Denmark.

On the other hand, other research found negative or no segregation effects on immigrants' outcomes. For instance, Borjas (2000) showed that ethnic segregation slows down wage growth for immigrants in the U.S. Beaman (2012) found that an increase in network size can negatively impact new arrivals yet benefit tenured members. Grönqvist (2006) used Swedish data and found that the size of enclave has no effect on earnings.

Moreover, segregation effects might change with the average human capital of ethnic groups. For example, Edin et al. (2003) found that immigrants from high income groups benefit more from ethnic segregation. Cutler et al. (2008) showed that ethnic group average education levels negatively impact segregation effects. However, this relationship was reversed when they instrument the endogenous segregation intensity.

In this paper, we propose that ethnic segregation has different impacts on immigrants with different education levels and from ethnic groups with different average education levels. If our hypothesis is true, it may help explain why findings regarding segregation effects are inconsistent in the previous literature. We will analyze different cases in the following three sub-sections.

3 In the rest of the manuscript, we use the phrase "immigrants" in replace for "young immigrants" for brevity purposes.

2.2 Individual Education and Segregation Effects

Given the same ethnic group average education level, the effects of ethnic segregation might be different for immigrants with different education levels. The direction of these effects, however, is uncertain. Living in ethnic enclaves may lead to worse outcomes for lower educated immigrants than for highly educated immigrants. Lower educated immigrants who are living in ethnic enclaves have fewer chances to speak the host country language, which slows down their assimilation process. In addition, jobs in ethnic enclaves might be of a lower quality (Cutler et al., 2008). Since highly educated immigrants may have broader networks extending beyond their local neighborhoods, if ethnic segregation has negative effects on immigrants' labor market outcomes in general, education may in some part protect highly educated immigrants from these effects.

On the contrary, it is also likely that segregation may be more harmful to highly educated immigrants because their opportunity costs of staying in ethnic enclaves might be larger than those of lower educated immigrants'. More interactions with natives and job searching in native-dominated labor markets may bring highly educated immigrants more job opportunities and higher returns to education. Lower educated immigrants might instead benefit from ethnic segregation because ethnic enclaves could provide job networks and language environments where lower educated immigrants can obtain job market information more easily (Bayer et al., 2008).

2.3 Ethnic Group Average Education and Segregation Effects

Similarly, ethnic segregation may have different impacts on immigrants from ethnic groups with different average education levels. The direction of the segregation effects might still be unclear.[4] Immigrants from well-educated groups may receive higher wages if they assimilate into native communities than if they live in ethnic enclaves, since group average education might correlate with unobserved talents which have higher returns in native-dominated labor markets. For example, immigrants from Asian countries may have higher levels of analytical and quantitative skills than other immigrants (Zong and Batalova, 2016), which may complement natives' interactive and communication skills (Peri and Sparber 2011). Thus, if they choose to work with natives rather than compete for lower educated jobs in ethnic enclaves, they might have higher wages.

Alternatively, high group average education levels may also strengthen the positive segregation effects for immigrants. With more highly educated immigrants in the same ethnic enclave, immigrants of all education levels may benefit from better quality in professional networks. Also, highly educated immigrants may create more job opportunities near ethnic enclaves by building targeted businesses, such as ethnicity-specific companies (Borjas, 2000).

2.4 Differential Segregation Effects for Different Individual and Group Education Levels

Finally, the differential segregation effects by individual education levels and ethnic group average education levels are again indeterminant for the following reasons. Ethnic segregation

4 Cutler et al. (2008) concluded that group average education positively influences segregation effects. However, this result is based on an instrumental variable which we argue might violate the exclusion restriction. Our analysis suggesting this is in Appendix A.

may be beneficial to immigrants who are living in ethnic enclaves with many similarly educated co-ethnics, because the positive network effects of ethnic segregation could amplify and improve job matching quality. For example, it might be easier for a lower educated Chinese immigrant to find a table waiting job in Chinatown, and a highly educated Chinese computer engineer to find better job information through the Chinese community near Silicon Valley.

Conversely, ethnic segregation may lead to adverse labor market outcomes as competition between immigrants with similar education levels drives wage rates down. This negative effect may be especially harmful for lower educated immigrants who are constricted to the labor markets near ethnic enclaves.[5] For example, lower educated Chinese immigrants living in Chinatown may face a fiercer competition from lower educated co-ethnics than if they were living in a Chinese community near Silicon Valley. Also, as was mentioned above, both education and unobserved talents related to high group average education levels may have higher returns in the native-dominated labor market. Therefore, the costs of segregation could be especially high for highly educated immigrants who are living in ethnic enclaves with high levels of group average education.

In conclusion, segregation effects may vary with immigrants' individual education levels, group average education levels, and the two factors combined. However, the direction of these effects remains unclear. This paper therefore empirically tests these differential segregation effects.

3 Data

3.1 The U.S. Census and the American Community Survey

Our study uses the Integrated Public Use Micro Series (Ruggles et al., 2015) from the 2000 U.S. Census and the American Community Survey (ACS) five-year sample for 2006-2010. The census data are 5-percent samples of the U.S. population. The 2006-2010 ACS sample is a five-year estimate of all geographical areas using the 1-percent samples from 2006, 2007, 2008, 2009, and 2010.[6] Individual weights are adjusted appropriately.

The outcome variables are inflation adjusted log of annual earnings and a dummy variable indicating whether an immigrant is currently employed. Immigrants who are currently students, who are not living in metropolitan areas, or who are not in the labor force were excluded from our sample.

We also restrict our sample to immigrants between the ages of 20 and 30 and migrated to the U.S. when they were 17 or older. There are several reasons for this sample restriction. First, for immigrant youths, it is more likely that their parents selected their residential locations for them (Cutler et al., 2008), unlike older immigrants who have more time to decide where to live in the host country based on their own preferences.[7] Therefore, immigrant youths are the least

5 For example, Patel and Vella (2013) find that newly arrived immigrants are more likely to choose the same occupations as their co-ethnics. Immigrants from non-English-speaking countries tend to have jobs which require less communication skills (Chiswick and Taengnoi, 2007).

6 The advantage of using this single combined 5-year estimate dataset is that it increases statistical reliability of the data for less populace areas and small subgroups. Considering this paper includes all identifiable ethnic groups, where some of them have small population sizes, we believe this is the best sample to use for more accurate identification (U.S. Census, 2022).

7 We run balancing tests by using samples from different age groups. We find that samples of older immigrants are more likely to fail the balancing tests. Also, when we do not restrict the age of migration to the U.S., results of the balancing tests also show systematic selections into ethnic enclaves.

likely to endogenously sort into ethnic enclaves. Second, we are concerned that ethnic segregation may influence immigrants' individual education levels if they move to the host country before having graduated from high school. It is therefore difficult to separate direct segregation effects on immigrants' labor market outcomes from indirect segregation effects (through the influence of ethnic segregation on immigrants' individual education levels). Immigrants who migrated to the U.S. at age 17 or later would normally have made their education decisions regarding high school completion and college attendance before arriving. Then, even if their ethnic group impacted their education, the influence would be more limited. The summary statistics are shown in Table 1.

3.2 **Isolation Index**

In this paper, we use the isolation index as a measure of ethnic segregation intensity. It thereby shows the extent of an immigrant's exposure to other immigrants from the same ethnic group in the same area. We use Equation (1) to calculate the isolation index,

$$ISO_{gmt} = \frac{\sum_i \frac{group_i}{group_{total}} \times \frac{group_i}{population_i} - \frac{group_{total}}{population_{total}}}{1 - \frac{group_{total}}{population_{total}}} \tag{1}$$

where $group_i$ is the population of immigrants from ethnic group g living in tract i; $group_{total}$ is the population of immigrants from ethnic group g living in MSA m; $population_i$ is the total population in tract i; $population_{total}$ is the total population in MSA m. ISO_{gmt} is the MSA-level isolation index according to country of origin, calculated based on tract-level data. The isolation index ranges from 0 to 1, where 0 indicates a complete absence of co-ethnics and 1 is the presence of only co-ethnics.

Table 1 Descriptive Statistics

	Earning regression sample			**Employment regression sample**		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**
Log earnings	104,577	9.656	0.738	-	-	-
Employment	-	-	-	129,966	0.919	0.273
Isolation index	8,149	0.014	0.021	8,874	0.014	0.020
Group average education	8,149	13.136	2.241	8,874	13.153	2.241
Immigrant share	8,149	0.004	0.015	8,874	0.004	0.014
Education	104,577	11.023	4.170	129,966	10.922	4.218
Age	104,577	26.395	2.829	129,966	26.330	2.874
Male	104,577	0.680	0.467	129,966	0.655	0.475
Married	104,577	0.383	0.486	129,966	0.385	0.486
Have children	104,577	0.301	0.459	129,966	0.305	0.461
White	104,577	0.121	0.326	129,966	0.121	0.326
Black	104,577	0.025	0.157	129,966	0.026	0.159
Asian	104,577	0.182	0.386	129,966	0.176	0.381
Hispanic	104,577	0.672	0.469	129,966	0.677	0.468

This method considers the effects of the total population and the ethnic group population on isolation intensity. For example, if we use the traditional equation to calculate isolation intensity, even if the population of an ethnic group is very small, we may still get a very large isolation index.[8] However, it is inevitable that immigrants of that group will come into contact with individuals from other ethnic groups. Therefore, the traditional approach of calculating isolation index may not be an ideal measure for our purposes.

We utilize the tract-level information in the 2006-2010 ACS five-year sample (U.S. Census Bureau, 2016) and the 2000 Summary Tape File 3 tables to construct the isolation index. Ethnic group is defined by an immigrant's country of origin. We have 60 different ethnic groups for the 2000 sample and 74 groups for the 2010 sample. Since we are interested in using multiple cross-sections of data across time, we will, after merging, use the 59 ethnic groups that are consistently defined across all sample years.

Following previous literature (Cutler et al., 2008; Cutler and Glaeser, 1997), we use the MSA-level isolation index instead of neighborhood level measures, such as percentage of immigrants from an ethnic group in a tract/block. We focus on comparing the outcomes of immigrants living in highly segregated MSAs with those of immigrants living in less segregated MSAs. This is because, within the same MSA, immigrants with higher socioeconomic status are more likely to move to better neighborhoods. Hence, the selection bias associated with neighborhood-level segregation intensity might be more substantial than the selection bias associated with MSA-level segregation intensity. The comparison between neighborhoods within MSAs will lead to biased estimations of the segregation effects.

4 Model

In this section, we begin by discussing our use of a triple differences model to deal with the selection bias caused by immigrants' sorting behaviors, then we discuss how we estimate the differential isolation effects.

$$y_{igmt} = \alpha + \beta ISO_{gmt} + X_{igmt}\Gamma + Edu_{igmt}\Lambda + \theta share_{gmt} + \mu_{gt} + \delta_{mt} + \tau_{gm} + u_{igmt} \tag{2}$$

Equation (2) shows the triple differences model. Here, y_{igmt} represents either the log annual earnings of immigrant who belongs to ethnic group g living in MSA m in year t, or a dummy variable indicating whether this person is employed. The independent variable of

8 The traditional equation for the isolation index is $\sum_i \frac{group_i}{group_{total}} \times \frac{group_i}{population_i}$ (Bell, 1954). It is the percentage of immigrants from a specific ethnic group in a given tract where the average number of immigrants from the same ethnic group live. In an extreme case, where all of the immigrants from ethnic group g all live in one tract, both $\frac{group_i}{group_{total}}$ and $\frac{group_i}{population_i}$ can be very large. However, the percent of immigrants from ethnic group g among the total population in MSA m might be very small. Therefore, immigrants from ethnic group g must interact with non-group members. For more details about the isolation index, see Cutler et al. (1999).

Another method that is often used to measure ethnic segregation is the dissimilarity index. The equation is $\frac{1}{2}\Sigma\left(\frac{group_i}{group_{total}} - \frac{non-group_i}{non-group_{total}}\right)$. Again, the value of the dissimilarity index might be driven by a few neighborhoods within an MSA. In other words, if segregation intensities are very high in a few neighborhoods, the dissimilarity index might show that immigrants from a certain group are highly concentrated within an MSA, even if group members in most neighborhoods are being extensively exposed to natives (Cutler et al. 2008).

interest is ISO_{gmt}, which is the isolation index of ethnic group g in MSA m in year t. If isolation has a positive effect on immigrants' labor market outcomes, then β will be positive and significant. The percentage of immigrants from group g in MSA m in year t is denoted by $share_{gmt}$. Individual characteristics, including age, gender, marital status, whether having children in the household, and race, are included in X. Individual years of schooling is represented by Edu_{igmt}.

Since we utilize multiple cross-sections of data across years, we can adopt the triple differences technique in order to control for ethnic group-specific time effects (μ_{gt}), time-varying MSA effects (δ_{mt}), and time-invariant MSA-specific ethnic group effects (τ_{gm}). The triple differences model removes unobserved factors that influence the labor market outcomes of immigrants from different ethnic groups, living in different MSAs, and in different years. In addition, the triple differences model could also eliminate other types of unobserved factors such as if immigrants from some ethnic groups experience faster wage growth rates over time, if economic conditions in some MSAs deteriorate or improve faster than in others over time, or if local policies and job training programs benefit some immigrant groups but not others.[9] However, if unobserved ability factors cause lower educated (or highly educated) immigrants to be more likely to sort into their ethnic enclaves (Cutler et al., 2008), our identification strategy might not remove all of the selection bias. For example, if immigrants with lower levels of social skills are more likely to be lower educated and are more likely to live in ethnic enclaves, then social skills will to some extent determine their labor market outcomes. In this situation, the triple differences model cannot eliminate the bias associated with social skills. As such, we opt for the method used in Bifulco et al. (2011) to test the validity of the triple differences model. We discuss this technique in detail in Section 5.

We are also interested in measuring how segregation effects vary with immigrants' individual education levels and ethnic group average education levels. Cutler et al. (2008) investigated the differential segregation effects on incomes of immigrants from ethnic groups with different group average education levels. The estimation equation can be written as the following:

$$y_{igmt} = \alpha + \beta ISO_{gmt} + \lambda S_{igmt} + \omega \overline{S}_{gmt} + \phi\left(ISO_{gmt} \cdot \overline{S}_{gmt}\right) + X_{igmt}\Gamma + \theta share_{gmt} + \mu_{gt} + \delta_{mt} + \tau_{gm} + u_{igmt} \quad (3)$$

where S_{igmt} and $\overline{S}_{gmt}$ are the immigrant's individual and group average education levels (both variables measured in years).[10] The coefficient of the interaction term between the isolation index and ethnic group average education level (ϕ) shows how the ethnic group average education level influences the segregation effects. It is likely that ϕ is positive, since exposure to more highly educated individuals from the same ethnic groups may provide immigrants with better job opportunities.[11]

In this paper, we complement the existing literature by introducing additional two-way interaction terms among the isolation index, individual education, and ethnic group average

9 More examples of the triple differences model can be found in Yelowitz (1995) and Ravallion et al. (2005). Angrist and Pischke (2008) also briefly discuss the triple differences model.

10 Cutler et al. (2008) used MSA and group fixed effects model to identify segregation effects. In Equation (3), we modify their estimation equation by using the triple differences model.

11 Cutler et al. (2008) found that ϕ is negative in the MSA and group fixed effects model, and that ϕ is reversed after using mean years since immigration of members from the same ethnic group within a city to correct for the endogeneity of the isolation index.

education, as well as the triple interaction term of the three variables. The following represents the main regression:

$$\begin{aligned} y_{igmt} = {} & \alpha + \beta ISO_{gmt} + \lambda S_{igmt} + \omega \overline{S}_{gmt} + \rho\left(ISO_{gmt} \cdot S_{igmt}\right) + \phi\left(ISO_{gmt} \cdot \overline{S}_{gmt}\right) + \chi\left(\overline{S}_{gmt} \cdot S_{igmt}\right) \\ & + \psi(ISO_{gmt} \cdot \overline{S}_{gmt} \cdot S_{igmt}) + X_{igmt}\Gamma + \theta share_{gmt} + \mu_{gt} + \delta_{mt} + \tau_{gm} + u_{igmt} \end{aligned} \tag{4}$$

where $ISO_{gmt} \cdot S_{igmt}$ and $ISO_{gmt} \cdot \overline{S}_{gmt} \cdot S_{igmt}$ are the variables of interest. The coefficient ρ shows how the segregation effects are different for highly educated and lower educated immigrants. For example, a negative ρ would imply that segregation hurts highly educated immigrants more than lower educated immigrants, given the same level of ethnic group average education. The coefficient of the triple interaction term (ψ) illustrates how immigrants' individual education and their ethnic group average education levels simultaneously influence the segregation effect. The advantage of this estimation is that it allows us to investigate how the effect of ethnic segregation changes when immigrants' own education levels match (or do not match) their ethnic group average education levels. For example, a negative ρ and a positive ψ indicate that, although segregation might hurt highly educated individuals, being isolated with well-educated co-ethnics can alleviate the negative segregation effects. The parameter χ represents how ethnic group average education levels influence the effects of immigrants' individual education levels. For example, if χ is positive, the implication is that highly educated immigrants will be more productive if they are exposed to more highly educated co-ethnics, when the segregation intensity is held constant. When taken entirely, the net effect of ethnic segregation for an immigrant *i* from ethnic group *g* living in MSA *m* in year *t* is $\beta + \rho + \phi + \psi$.

5 Validity of Identification Strategy

Before addressing our results, we need to test whether our primary identification strategy—the triple differences model—eliminates, or at least attenuates, the potential biases that are caused by immigrants' sorting behaviors. Immigrants are likely to choose their residential areas based on observed and unobserved characteristics, such as language, social connections, dietary habits, etc. It is thus likely that immigrants in segregated areas have different attributes from immigrants in less segregated areas. These attributes may influence their labor market outcomes and may exaggerate or undermine the effects of segregation. For example, immigrants who cannot speak native languages may prefer to reside within their ethnic enclaves (Bleakley and Chin, 2010). This lack of communication skills may cause them to have fewer high-paying employment opportunities regardless of where they live. Thus, we are more likely to find that ethnic segregation negatively correlates with immigrants' wages.

To address these issues, we conduct several balancing tests to determine if immigrants' characteristics can predict their own ethnic segregation intensity in the triple differences model. If the isolation index is unrelated to immigrants' observed characteristics, then our estimated results should not be influenced by immigrants' sorting behaviors based on observed characteristics. Also, as Altonji et al. (2005) and Bifulco et al. (2011) have pointed out, the extent of individuals' sorting behaviors based on observables can be used as a guideline for the extent of their sorting behaviors based on unobserved variables. Thus, a lack of correlation between immigrants' observed characteristics and the isolation index suggests that the triple differences model minimizes the likelihood that the estimated results are driven by immigrants' selections into ethnic enclaves based on unobserved characteristics.

We also examine whether the interaction terms between immigrants' individual education levels and other characteristics are correlated with the isolation index. If the interaction terms between individual education levels and other individual characteristics fail the balancing tests, then the estimated differential segregation effects associated with immigrants' individual education levels might be biased. For example, if we observe that the coefficient of the interaction between age and education is negative in the balancing test, then the implication is that immigrants who are older and have higher education levels are less likely to sort into their ethnic enclaves. If they are more likely to be harmed by ethnic segregation, then the estimated results of how immigrant education influences segregation effects might be underestimated.

In Table 2, we present balancing test results using the triple differences model.[12] In Column 1, the coefficients of individual characteristics are insignificant. The value of the *F*-test

Table 2 Balancing Tests, Triple Differences Model

Dependent Variable: Isolation Index × 1000		
	(1)	**(2)**
Education	−0.007	0.042
	(0.007)	(0.060)
Age	−0.006	0.014
	(0.006)	(0.021)
Male	0.053	0.144
	(0.042)	(0.127)
Black	−0.285	−0.000
	(0.313)	(0.870)
Other Nonwhite	0.088	−0.026
	(0.215)	(0.395)
Age × Education	-	−0.002
		(0.002)
Male × Education	-	−0.008
		(0.011)
Black × Education	-	−0.022
		(0.060)
Other Nonwhite × Education	-	0.009
		(0.027)
F-test	1.05	0.45
P-value	0.385	0.771
N	102,475	

Notes: The isolation index is inflated by 1000. We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level. The values of F-statistic in column 1 are for the joint effect of all the individual characteristics that are included in the regressions. The values of F-statistic in column 2 are for the joint effect of all the interaction terms. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

12 In order to show the magnitudes of the coefficients, we multiply the isolation index by 1000 in all balancing tests. We also use the earnings regressions sample for all balancing tests. When we use the sample for employment regressions, individual characteristics and their interactions with individual education are still insignificant as a group.

shows that individual characteristics cannot explain the isolation index as a group. In Column 2, we add interaction terms between individual education and other characteristics into the regression. The joint effects of all the interaction terms are also insignificant. Results of the balancing tests show that the triple differences model is sufficient to remove immigrants' systematic selections into ethnic enclaves based on their observables. Therefore, the concern that immigrants' sorting behaviors might lead to biased estimates is minimal.[13]

6 Main Results

In this section, we begin by discussing the estimated relationship between the isolation index and log earnings and how ethnic group average education affects the segregation effects. In Table 3, the first three columns show the empirical results of Model (2) with an increasing number of control variables in individual characteristics and education levels. The estimated coefficients of the isolation index in the first three columns are consistently statistically insignificant and negative. This suggests that there is no significant difference between earnings of immigrants living in various degrees of ethnic segregation.

Table 3 Triple differences, *ISO* Model and $ISO \times \overline{S}$ Model, Earnings Results

Dependent variable: log earning						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
ISO	−0.217	−0.233	−0.156	0.438	−0.498	−0.582
	(0.469)	(0.436)	(0.417)	(1.721)	(1.599)	(1.556)
$ISO \times \overline{S}$	-	-	-	−0.039	0.059	0.061
				(0.187)	(0.175)	(0.170)
$\overline{S}$	-	-	-	0.055***	0.049***	0.023*
				(0.013)	(0.012)	(0.012)
S	-	-	0.028***	-	-	0.028***
			(0.002)			(0.002)
+characteristic control	No	Yes	Yes	No	Yes	Yes
F-test	-	-	-	0.04	0.06	0.07
P-value	-	-	-	0.960	0.943	0.933
N	102,475					

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. We control for MSA-specific time trends, MSA-specific group effects, and group-specific time trends in all regressions. Characteristic controls include age, gender, marital status, children in the household, black, Asian, and Hispanic dummy variables. The values of F-statistic are for the joint effect of the isolation index and the interaction term between isolation index and group average education. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

13 A commonly used method to test the validity of double differences models is to examine the parallel trends, i.e., without any treatment, the treatment group and the control group should have similar changes in outcomes over time/across different locations. Olden and Møen (2022) show that the triple differences model is essentially the difference between two sets of double differences, and it does not require both double differences to have parallel trends at the same time. Instead, the two double differences just need to have biases in similar direction and magnitude. By following their method, we present evidence of parallel trend for our model in Appendix B.

This insignificant segregation effect, however, could be because of a lack of control variables for group education levels. For example, if the segregation effects are opposite for different group education levels for reasons suggested in Section 2, then the average segregation effect might be statistically insignificant. We therefore present the estimated results of Model (3), which is Model (2) with group average education level and its interaction term with the isolation index, in the last three columns of Table 3. The consistently statistical insignificant coefficients of the isolation index and its interaction term suggest that the effect of isolation on log earnings is not influenced by group average educations.[14]

Moreover, the group average education variable becomes less significant after individual education level is controlled for. Therefore, we further investigate whether individual education level could alter the effects of segregation on immigrants' labor market outcomes, both individually and collectively with group average education. We show the results of regressions with all two-way and three-way interaction terms between the isolation index, individual education, and group average education level in Table 4. To properly interpret the coefficients of these interaction terms, we center all the involving variables at their mean level. Specifically, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively.

We begin with the regression containing only the isolation index and its interaction with immigrants' individual education levels. The main effect of segregation is negative but statistically insignificant, but the coefficient of the interaction term is negative and statistically significant. This suggests that highly educated immigrants who are living in ethnic enclaves have lower earnings than they would in diverse or native communities. Specifically, one percent increase in segregation reduces highly educated immigrants' earnings by 0.174%. This is consistent with the hypothesis in Section 2.2., in that higher opportunity costs penalizes highly educated immigrants who live in ethnic enclaves, while lower educated immigrant can benefit from local labor market networks and less reliance on the host country linguistic skills. Additionally, higher individual education level also raises earnings, as per conventional wisdom.

In Column 2 of Table 4, we show the estimated results for the regression that only includes double interaction terms between isolation index, group average education and individual education. The estimated coefficient for the interaction between isolation index and individual education is consistently negative but has become statistically insignificant. The coefficients of the isolation index and its interaction with group average education are both statistically insignificant. Lastly, the positive estimated coefficient of the interaction term between group average education level and individual education level suggests that generally, immigrants tend to earn higher wages if they are similarly educated as their ethnic groups, irrespective of where they reside. One possible explanation is that ethnic groups with a certain level of group average education may concentrate in and have labor market advantages in certain industries. When immigrants' own education level matches their ethnic group average education, they could

14 Our findings in Table 3 are inconsistent with the OLS model findings in Table 7 of Cutler et al. (2008). There are two potential reasons behind the difference. First, we use a triple differences identification strategy. As shown in Appendix Table A.2, estimations without the triple differences controls do not entirely take immigrants' endogenous sorting into account, which could bias results. Second, we also conjecture that the reason could be stemming from differences in samples. Our sample contains a larger number of ethnicity-city-year cells than Cutler et al. (2008).We also use yearly data (the 2000 and 2010 samples) to replicate findings in Cutler et al. (2008).The results are shown in Appendix Table A.1.

Table 4 Triple Differences and Triple Interactions Model (Centered), Earnings Results

Dependent variable: log earning				
	(1)	**(2)**	**(3)**	**(4)**
ISO	−0.136	0.375	0.415	0.327
	(0.414)	(0.846)	(0.847)	(0.837)
$ISO \times S$	−0.174***	−0.029	−0.160*	−0.216**
	(0.021)	(0.023)	(0.083)	(0.096)
$ISO \times \overline{S}$	-	0.109	0.117	0.098
		(0.169)	(0.169)	(0.168)
$ISO \times \overline{S} \times S$	-	-	−0.036*	-
			(0.184)	
$ISO \times \overline{SH} \times S$	-	-	-	0.505***
				(0.138)
$ISO \times \overline{SL} \times S$	-	-	-	−0.048**
				(0.021)
$\overline{S} \times S$	-	0.010***	0.011***	0.011***
		(0.001)	(0.001)	(0.001)
$\overline{S}$	-	0.007	0.006	0.007
		(0.012)	(0.012)	(0.012)
S	0.040***	0.063***	0.066***	0.065***
	(0.002)	(0.002)	(0.002)	(0.002)
F-test	33.97	0.69	1.10	3.09
P-value	0.000	0.559	0.354	0.009
N		102,475		

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. To center the variables, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively. $ISO \times \overline{SH} \times S$ is equal to $ISO \times \overline{S} \times S$ if $\overline{S}$ is greater than its MSA-year mean, and 0 otherwise. Similarly, $ISO \times \overline{SL} \times S$ is equal to $ISO \times \overline{S} \times S$ if $\overline{S}$ is lower than its MSA-year mean, and 0 otherwise. The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for $\overline{S} \times S$. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

benefit either from the ethnic neighborhoods' group network effect or from the overarching shared expertise in a specific industry even if they do not live with their co-ethnics.

As mentioned in Section 2.4, there could be a more intricate effect of isolation that depends on both individual education and ethnic group average education, so in Column 3 of Table 4, we include the triple interaction term in the regression. The estimated coefficient of this triple interaction term is negative and statistically significant at 10% level, suggesting that match or mismatch between an immigrant's own education level and their group average education level play an important role in the effect of isolation. Even though the model specification in Column (3) points to the importance of the triple interaction term, the interpretation of its coefficient is still unclear. As discussed in Section 2.4, the negative coefficient of the triple interaction term might have different implications for immigrants with different levels of individual education.

As a result, instead of using one triple interaction term, in Column 4 of Table 4, we parse out the effects of segregation by considering highly educated and lower educated ethnic groups separately.[15] The estimated coefficient of the triple interaction term is 0.505 for the highly educated ethnic group and -0.048 for the lower educated ethnic groups. Since all the variables within the interaction terms are centered, the interpretation is that for lower educated immigrants, the positive segregation effect on earnings decreases if they live amongst lower education co-ethnics but living with highly educated co-ethnics reduces the positive segregation effects more. For highly educated immigrants, residing in ethnic enclaves with high average education can improve the negative segregation effects, while living with lower educated ethnic peers increases segregation effects by a smaller magnitude.

When considering all statistically significant terms involving the isolation index, the model shows that for lower educated immigrants who live in ethnic enclaves with low average education levels, a one percent increase in the isolation index increases earnings by 0.168 percent.[16] On the other hand, if the same lower educated immigrant resides with many highly educated co-ethnics, a one degree increase in the isolation index would decrease their earnings by 0.289 percent. For a highly educated immigrant who live in ethnic enclaves with low average education, if the isolation index increases by one degree, they experience a 0.168 percent decrease in earnings, and if they live in a highly educated enclave, they will experience a 0.289 percent increase in their earnings.[17]

These results are notable in two important ways. First, while the average effect of isolation is positive for lower educated immigrants and negative for highly educated immigrants, this effect can be significantly influenced by whether the immigrants' individual education matches the group average education. For lower educated immigrants, the reduction in segregation effect is significantly larger if they are isolated with highly educated co-ethnics than if they are isolated with lower educated co-ethnics. This means while the competition with similarly educated co-ethnics could hurt lower educated immigrants' wages, removing them from their similarly educated co-ethnics, and therefore removing their access to education-specific networks, would reduce their wages so much that it reverses the general positive effect of isolation. On the other hand, while isolation causes an average negative effect for highly educated immigrants, being isolated with other highly educated co-ethnics could provide sufficiently high-quality networks to counteract that effect, reversing the isolation effect to positive. Both scenarios point to the same crucial role education-specific networks play.

Second, there is a seemingly contradiction between the average isolation effect found in Column 1 and the net effects of isolation found in Column 4. This contradiction can be reconciled by further investigation into the composition of immigrant groups. Because a significant

15 Specifically, we generate two new variables: $ISO \times \bar{SH} \times S$ and $ISO \times \bar{SL} \times S$. The former one is equal to the value of the triple interaction term if the ethnic group average education level is higher than the average level. Otherwise, it is equal to zero. The latter one is equal to the value of the triple interaction term if the ethnic group average education level is lower than the average level. All three parts of the interaction terms are centered as before.

16 This is the net effect of isolation when the general effect of 0.216 is reduced by the triple interaction effect of -0.048. The net effects in the rest of this section are all calculated accordingly.

17 We have also conducted the experiment using categorical education variables instead of years of schooling. The model specification is more complex due to the nature of necessary interaction terms in the model design. The results are presented in the tables in Appendix C. Using the categorical education variables makes the interpretation of the estimated coefficient of Isolation Index itself more complex, but the directions of the effects are similar to those found in the preferred model, i.e. for lower educated immigrants who isolate, the benefits are larger if they live amongst lower educated co-ethnics; and for highly educated immigrants, being isolated with co-ethnics of high average education could reduce the negative effects of isolation.

portion of immigrant population are lower educated, lower educated immigrants are 3.5 times more likely to be isolated with other lower educated co-ethnics than with highly educated co-ethnics. Similarly, highly educated immigrants are 2.87 times more likely to live in ethnic enclaves with low average education than in ethnic enclaves with high average education. Therefore, the positive segregation effects for lower educated immigrants and the negative segregation effects for highly educated immigrants dominates.

In summary, our findings suggest a rather clear story: the net effects of ethnic segregation on earnings depends on the match or mismatch between immigrants' own education levels and their group average education levels. While on average, ethnic segregation hurts highly educated immigrants and benefits lower educated immigrants, the opposite is true in the case of being segregated with co-ethnics of high average education levels.

7 Robustness Tests

7.1 Robustness to Residential Sorting by Gender

As shown in Table 2, the estimated coefficients of male dummy and the interaction term between male dummy and individual education are both statistically insignificant in the balancing test regression. This suggests that any potential residential sorting according to gender should not affect our main empirical results. However, since 68% of the sample are male immigrants, we nevertheless test whether our model is robust when males and females are estimated separately. The results in the first two columns of Table 5 show a rather similar picture for male and female immigrants, except that the coefficients for female immigrants are less precisely estimated than for male immigrants.

7.2 Robustness to Sampling

As stated in Section 3, we have restricted our sample to immigrants who migrated to the U.S. after age 17 to avoid endogenous choices in education post arrival in the host country. However, their education decisions may still be partially influenced by their experience in the U.S. Therefore, we further restrict our sample to immigrants who migrated after age 21 and age 23 to test the robustness of our findings, because these immigrants' education levels are more likely to have been determined before arrival in the U.S.

In Columns 3 and 4 of Table 5, we re-estimate the model using only immigrants who migrated after age of 21 and immigrants who migrated after age of 23, respectively. The estimated coefficients are largely robust in terms of signs, but with lower level of statistical significance for some variables. This is not surprising because the sample size has decreased by almost 50% and 70% after the sample restrictions. Given the number of variables in the triple differences regressions, we expect the estimation to be less precise.

7.3 Robustness to Additional Controls

The literature of immigration has suggested that immigrants' English language skills (Bleakley and Chin, 2004) and their unobserved degrees of social assimilation (Furtado and Song, 2015)

Table 5 Robustness Tests, Triple Differences and Triple Interactions Model (Centered)

Dependent variable: log earning				
	(1) Male	**(2) Female**	**(3) Arrived at the U.S. after age 21 or later**	**(4) Arrived at the U.S. after age 23 or later**
ISO	−0.945	1.962	−0.572	0.294
	(1.146)	(1.618)	(1.001)	(1.262)
$ISO \times S$	−0.281***	−0.108	−0.179	−0.158
	(0.095)	(0.082)	(0.115)	(0.129)
$ISO \times \overline{S}$	−0.124	0.449	−0.187	0.115
	(0.221)	(0.329)	(0.212)	(0.281)
$ISO \times \overline{S}H \times S$	0.554***	0.439*	0.513***	0.368*
	(0.156)	(0.233)	(0.183)	(0.222)
$ISO \times \overline{S}L \times S$	−0.060***	−0.035*	−0.038	−0.033
	(0.021)	(0.020)	(0.025)	(0.029)
$\overline{S} \times S$	0.010***	0.012***	0.010***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
$\overline{S}$	−0.011	0.014	−0.001	−0.010
	(0.015)	(0.025)	(0.019)	(0.025)
S	0.060***	0.071***	0.065***	0.063***
	(0.003)	(0.003)	(0.003)	(0.003)
F-test	3.68	1.88	1.73	0.73
P-value	0.003	0.095	0.123	0.600
N	69,166	31,718	51,828	32,188

Notes: To center the variables, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for $\overline{S} \times S$. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level. Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10

could play an important role in labor market outcomes. Since it is likely that immigrants' education is correlated with these factors, we further test whether the estimated segregation effect according to education is robust to additional controls of linguistic and social skills.

We start with conducting a second series of balancing tests with two additional variables: an indicator of whether the immigrant has insufficient English-speaking skill and the number of years the immigrant has lived in the U.S.[18] The results presented in Table 6 suggests that immigrants who speak English poorly tend to live in less segregated areas. Since this group of immigrants are more likely to have lower levels of education and are more likely to benefit from segregation, it is likely that we have underestimated the positive segregation effects for lower educated immigrants and overestimated the negative segregation effects for highly educated immigrants.

Therefore, in Table 7, we test the robustness of our model by including the indicator of insufficient English-speaking skill, number of years in the U.S., and their interactions with the

18 We categorize immigrants who self-report as speaking English not well or not at all as having insufficient English-speaking skill.

Table 6 Balancing Tests with Assimilation Measures, Triple Differences Model

Dependent Variable: Isolation Index × 1000		
	(1)	**(2)**
Education	−0.011	0.017
	(0.007)	(0.061)
Age	−0.002	0.008
	(0.008)	(0.023)
Male	0.051	0.142
	(0.042)	(0.129)
Black	−0.278	−0.051
	(0.313)	(0.867)
Other Nonwhite	0.097	0.023
	(0.215)	(0.409)
English Insufficiency	−0.132***	−0.255*
	(0.050)	(0.154)
Years in the U.S.	−0.012	0.007
	(0.008)	(0.019)
Age × Education	-	−0.001
		(0.002)
Male × Education	-	−0.008
		(0.011)
Black × Education	-	−0.017
		(0.060)
Other Nonwhite × Education	-	0.006
		(0.028)
English Insufficiency × Education	-	0.012
		(0.014)
Years in the U.S. × Education	-	−0.002
		(0.002)
F-test	1.73	0.064
P-value	0.097	0.695
N	102,475	

Notes: The isolation index is inflated by 1000. We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level. The values of F-statistic in column 1 are for the joint effect of all the individual characteristics that are included in the regressions. The values of F-statistic in column 2 are for the joint effect of all the interaction terms. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

isolation index.[19] In Column 2, the estimated coefficient of the English insufficiency dummy is negative and statistically significant, but the interaction term between the English insufficiency dummy and isolation index has a positive and statistically insignificant coefficient. This

19 Even though in our data we already exclude immigrants who have not completed their secondary education, we argue there is value in identifying the heterogeneous effects for immigrants with different degrees of English language skills. First, different home countries offer varying degrees of English language education and therefore there is still sufficient variation in English language skills in the data. Second, given the variation in the English language skills for the immigrants in the sample, we could also use English language skill as a proxy for social skills. Third, the estimated coefficients we gather in Column 2 in Table 6 can be considered a lower-bound effect of English language's effect as immigrants who have not completed secondary education upon arriving in the U.S. would tend to have worse English language skills and labor market outlooks. Additionally, in Appendix A Table A.4, we also show the results by using the sample excluding immigrants who identified themselves as English speakers. There aren't significant changes in the results either.

Table 7 Mechanisms and Heterogeneity Tests, Earning Results

Dependent variable: log earning			
	Baseline	**English insufficiency**	**Years in the U.S.**
ISO	0.327	0.320	0.673
	(0.837)	(0.843)	(0.809)
$ISO \times S$	−0.216**	−0.247***	−0.244***
	(0.096)	(0.090)	(0.094)
$ISO \times \overline{S}$	0.098	0.132	0.115
	(0.168)	(0.171)	(0.164)
$ISO \times \overline{S}H \times S$	0.505***	0.516***	0.544***
	(0.138)	(0.139)	(0.141)
$ISO \times \overline{S}L \times S$	−0.048**	−0.056***	−0.057***
	(0.021)	(0.020)	(0.021)
insufEng		−0.183***	-
		(0.012)	
$ISO \times insufEng$	-	0.046	-
		(0.131)	
Years in US		-	0.037***
			(0.002)
$ISO \times Years\ in\ US$	-	-	−0.041**
			(0.017)
F-test	3.09	3.69	3.85
P-value	0.009	0.003	0.002
N		102,475	

Notes: To center the variables, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively. The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for $\overline{S} \times S$. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Therefore, singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

suggests that immigrants who have limited English proficiency experience an earnings penalty, which is consistent with previous literature (Bleakley and Chin, 2004). However, segregation effects do not vary by insufficient English-speaking skills. In addition, the estimated coefficients for $ISO \times S$ and the two triple interaction terms remain statistically significant at 1% level and their magnitudes only increase slightly.

The results in Column 3 of Table 7 suggest that additional years of living in the U.S. increase immigrants' earnings, consistent with the literature that ceteris paribus, an increase in social assimilation induces bettering labor market outcomes (Furtado and Song 2015). Moreover, immigrants who have stayed in the U.S. for longer time would suffer wage loss if they live in ethnic enclaves, possibly due to the lack of access to a wider labor market as suggested in Section 2. Similar to the results in Column 2, the estimated coefficients of $ISO \times S$ and the two triple interaction terms do not have significant changes. This implies that the higher returns to host-country-social-skills are not the reason for the differential segregation effects based on education.

To summarize, results in this sub-section show that the baseline results are robust to linguistic skills or social assimilation, which suggests that the differential segregation effects are not likely caused by immigrants' socio-communication skills. However, we can only test a limited number of variables here. The baseline results could be caused by other unobserved factors that are correlated with education levels, ethnic enclave residence, and earnings. For example, immigrants from families that are heavily influenced by traditional culture and values may be more likely to live in ethnic enclaves. If these immigrants are from wealthy families, they are more likely to get higher levels of education but may give up job opportunities that are far away from their ethnic enclaves. On the contrary, if they are from poor families, they would have lower educational attainment and are less capable in getting higher quality jobs.

8 Employment Outcomes

In this section, we test whether ethnic segregation has differential segregation effects on immigrants' employment outcomes. Similar to the earnings results, in Table 8, we did not find differential segregation effects on employment for immigrants from ethnic groups with different average education levels. In Table 9, only the coefficient of the double interaction between isolation index and individual education in Column 1 is statistically significant and negative, suggesting that on average, ethnic segregation reduces the likelihood of being employed for

Table 8 Triple differences, *ISO* Model and $ISO \times \overline{S}$ Model, Employment Results

Dependent variable: if employed, =1						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
ISO	−0.106	−0.109	−0.092	0.014	−0.061	−0.068
	(0.151)	(0.152)	(0.153)	(0.514)	(0.512)	(0.512)
$ISO \times \overline{S}$	-	-	-	−0.014	−0.006	−0.005
				(0.050)	(0.050)	(0.050)
$\overline{S}$	-	-	-	−0.001	−0.001	−0.004
				(0.005)	(0.005)	(0.005)
S	-	-	0.003***	-	-	0.003***
			(0.000)			(0.000)
+characteristic control	No	Yes	Yes	No	Yes	Yes
+individual education control	No	No	Yes	No	No	Yes
F-test	-	-	-	0.37	0.31	0.31
P-value	-	-	-	0.690	0.736	0.736
N	127,769					

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. We control for MSA-specific time trends, MSA-specific group effects, and group-specific time trends in all regressions. Characteristic controls include age, gender, marital status, children in the household, black, Asian, and Hispanic dummy variables. The values of the F-statistic are for the joint effect of the isolation index and the interaction term between isolation index and group average education. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Therefore, singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

Table 9 Triple Differences and Triple Interactions Model (Centered), Employment Results

Dependent variable: if employed, =1				
	(1)	**(2)**	**(3)**	**(4)**
ISO	−0.089	−0.009	−0.012	−0.018
	(0.153)	(0.249)	(0.249)	(0.249)
$ISO \times S$	−0.013**	−0.005	0.005	0.001
	(0.005)	(0.006)	(0.017)	(0.018)
$ISO \times \overline{S}$	-	0.027	0.026	0.025
		(0.056)	(0.056)	(0.056)
$ISO \times \overline{S} \times S$	-	-	0.003	-
			(0.004)	
$ISO \times \overline{S}H \times S$	-	-	-	0.042
				(0.037)
$ISO \times \overline{S}L \times S$	-	-	-	0.002
				(0.004)
$\overline{S} \times S$	-	0.001***	0.000**	0.000**
		(0.000)	(0.000)	(0.000)
$\overline{S}$	-	−0.005	−0.005	−0.005
		(0.004)	(0.004)	(0.004)
S	0.004***	0.005***	0.005***	0.005***
	(0.000)	(0.001)	(0.001)	(0.001)
F-test	2.97	0.49	0.57	0.68
P-value	0.052	0.687	0.686	0.635
N		127,769		

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. To center the variables, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively. $ISO \times \overline{S}H \times S$ is equal to $ISO \times \overline{S} \times S$ if $\overline{S}$ is greater than its MSA-year mean, and 0 otherwise. Similarly, $ISO \times \overline{S}L \times S$ is equal to $ISO \times \overline{S} \times S$ if $\overline{S}$ is lower than its MSA-year mean, and 0 otherwise. The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for $\overline{S} \times S$. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Therefore, singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.
Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

highly educated immigrants but improve lower educated immigrants' employment. However, this coefficient becomes statistically insignificant in the other model specifications. Moreover, the coefficients of the other double interaction and triple interaction terms show no discernable segregation effect on employment.

Table 10 presents a similar picture for male and female immigrants as well as immigrants who migrated after age of 23. The model becomes slightly more significant for the sample of immigrants who arrived after 21. Given the results are robust when using other three samples, this result is likely caused by random sample variations as opposed to a different labor market mechanism for immigrants who migrated when they were 21 and 22. Lastly, Table 11 suggests that controlling for English language skills and length of stay in the U.S. does not affect the statistical significance of the isolation index and its interaction terms with education variables.

Table 10 Robustness Test, Triple Differences and Triple Interactions Model (Centered), Employment Results

Dependent variable: if employed, =1				
	(1) Male	**(2) Female**	**(3) Arrived at the U.S. after age 21 or later**	**(4) Arrived at the U.S. after age 23 or later**
ISO	0.134	−0.346	−0.119	−0.133
	(0.304)	(0.464)	(0.280)	(0.332)
$ISO \times S$	−0.004	−0.001	0.039**	0.030
	(0.016)	(0.029)	(0.018)	(0.019)
$ISO \times \overline{S}$	0.038	0.006	0.013	0.115
	(0.065)	(0.113)	(0.069)	(0.077)
$ISO \times \overline{S}H \times S$	0.034	0.068	0.036	0.029
	(0.046)	(0.057)	(0.051)	(0.065)
$ISO \times \overline{S}L \times S$	0.001	−0.001	0.009**	0.007
	(0.004)	(0.007)	(0.004)	(0.005)
$\overline{S} \times S$	0.001***	−0.001	0.001**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$\overline{S}$	−0.005	−0.007	−0.009	−0.011
	(0.005)	(0.009)	(0.006)	(0.007)
S	0.005***	0.004***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
F-test	0.67	0.62	1.41	2.09
P-value	0.647	0.686	0.219	0.063
N	83,164	42,838	65,149	40,642

Notes: The sample used for robustness test is restricted to immigrants who arrived at the U.S. after age 21 or 23. To center the variables, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for $\overline{S} \times S$. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Therefore, singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

9 Conclusion

Residential segregation may have positive or negative impacts on immigrants' labor market outcomes. On the one hand, segregation might provide social networks where immigrants can exchange job information in their own languages. On the other hand, segregation might also slow immigrants' assimilation process by isolating immigrants from natives. In the current literature, there is no consensus as to whether ethnic segregation is beneficial or harmful to immigrants. In this paper, we attempt to use young immigrants' individual education levels to gain further insights on segregation effects and to measure the impacts of ethnic group average education on segregation effects.

Our estimation shows that on average, ethnic segregation reduces the earnings of highly educated young immigrants and boosts the earnings of lower educated young immigrants, but it does not have differing impacts on employment. Moreover, the education compositions

Table 11 Mechanisms and Heterogeneity Tests, Employment Results

Dependent variable: if employed, =1			
	Baseline	**English insufficiency**	**Years in the U.S.**
ISO	−0.018	−0.028	0.024
	(0.249)	(0.249)	(0.251)
$ISO \times S$	0.001	−0.000	−0.001
	(0.018)	(0.018)	(0.018)
$ISO \times \overline{S}$	0.025	0.031	0.027
	(0.056)	(0.056)	(0.056)
$ISO \times \overline{S}H \times S$	0.042	0.042	0.043
	(0.037)	(0.037)	(0.038)
$ISO \times \overline{S}L \times S$	0.002	0.001	0.001
	(0.004)	(0.004)	(0.004)
$ISO \times insufEng$	-	0.045	-
		(0.043)	
insufEng	-	−0.021***	-
		(0.004)	
$ISO \times Years\ in\ US$	-	-	−0.006
			(0.007)
Years in US	-	-	0.003***
			(0.001)
F-test	0.68	0.65	0.57
P-value	0.635	0.664	0.721
N		127,769	

Notes: To center the variables, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively. The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for $\overline{S} \times S$. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Therefore, singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.01.

of ethnic enclaves lead to different net segregation effects for immigrants with different education levels. Specifically, the net segregation effects are only positive when immigrants' own education levels match the average education levels of their ethnic enclaves. This is consistent with the labor market network theory: removing the lower educated co-ethnic networks could reduce lower educated immigrants' earnings, even when they live in ethnic enclaves. On the other hand, providing a high-quality network of similarly educated co-ethnics for highly educated immigrants would reverse the average negative effect of isolation and provide a net benefit in their earnings.

From a policy standpoint, this paper suggests that when policy makers promote diversity in ethnic neighborhoods, they should take into consideration the valuable benefits of ethnic segregation for lower educated immigrants, especially for those living with many other lower educated co-ethnics. While diversified neighborhoods have their own advantages, how to compensate for the potential reduction in earnings due to a loss of access to the education-specific co-ethnic labor market networks should be considered in the process of integration promotion.

List of Abbreviations

MSA, Metropolitan Statistical Area; ACS, American Community Survey; ISO, Isolation index.

Declarations

Availability of Data and Materials

The datasets generated and/or analyzed during the current study are available in the Integrated Public Use Microdata Series (IPUMS) repository, http://doi.org/10.18128/D010.V6.0, and U.S. Census Data repository, https://data.census.gov/cedsci/.

Competing Interest

The authors declare that they have no competing interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions

TL and TS developed the research questions, preliminary research methods and preliminary data analyses together. TL finalized the research methods, conducted the triple interactions and triple differences analyses, robustness tests and heterogeneous tests, and did proofs for the parallel trends for the triple differences model. TL was a major contributor in writing the manuscript. TL and TS both contributed to revise the manuscript. All authors read and approved the manuscript.

Acknowledgements

We are grateful to Dr. Stephen Ross and Dr. Delia Furtado for providing advice and guidance on this project. We also thank the participants of University of Connecticut Economics Department Labor, Development, and Health Economics seminar, 2017 Eastern Economics Association Annual Conference, and 2017 Western Economics Association Annual Conference for helpful feedbacks and comments.

References

Altonji, Joseph; Todd Elder; Christopher R. Taber (2005): Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. Journal of political economy 113(1), 151–184. doi: 10.1086/426036

Angrist, Joshua; Jörn-Steffen Pischke (2008): Mostly harmless econometrics. Princeton University Press.

Bayer, Patrick; Stephen Ross; Giorgio Topa (2008): Place of work and place of residence: Informal hiring networks and labor market outcomes. Journal of political Economy 116(6), 1150–1196. doi: 10.1086/595975

Beaman, Lori (2012): Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the U.S. The Review of Economic Studies 79(1), 128–161. doi:10.1093/restud/rdr017

Bell, Wendell (1954): A Probability Model for the Measurement of Ecological Segregation. Social Forces 32(4), 357–64. doi: 10.2307/2574118

Bifulco, Robert; Jason Fletcher; Stephen Ross (2011): The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. American Economic Journal: Economic Policy 3(1), 25–53. doi: 10.1257/pol.3.1.25

Bleakley, Hoyt; Aimee Chin (2004): Language Skills and Earnings: Evidence from Childhood Immigrants. The Review of Economics and Statistics 86(2), 481–96. doi: 10.1162/003465304323031067

Bleakley, Hoyt; Aimee Chin (2010):Age at arrival, English proficiency, and social assimilation among US immigrants. American Economic Journal: Applied Economics 2(1), 165–192. doi: 10.1257/app.2.1.165

Borjas, George (2000): Ethnic enclaves and assimilation. Swedish Economic Policy Review 7, 89–122.

Borjas, George (1998): To ghetto or not to ghetto: Ethnicity and residential segregation. Journal of Urban Economics 44, 228–253. doi: 10.1006/juec.1997.2068

Carliner, Geoffrey (2000): The language ability of US immigrants: Assimilation and cohort effects. The International Migration Review 34, 158–182. doi: 10.2307/2676016

Chiswick, Barry; Sarinda Taengnoi (2007): Occupational Choice of High Skilled Immigrants in the United States. International Migration 45(5), 3–34. doi: 10.1111/j.1468-2435.2007.00425.x

Cortes, Patricia (2008): The effect of lower educated immigration on US prices : Evidence from CPI data. Journal of Political Economy 116(3), 381–422. doi: 10.1086/589756

Cutler, David; Edward Glaeser (1997): Are ghettos good or bad? The Quaterly Journal of Economics 112(3), 827–872. doi: 10.1162/003355397555361

Cutler, David; Edward Glaeser; Jacob Vigdor (1999): The Rise and Decline of the American Ghetto. Journal of Political Economy 107(3), 455–506. doi:10.1086/250069

Cutler, David; Edward Glaeser; Jacob Vigdor (2008): When Are Ghettos Bad? Lessons from Immigrant Segregation in the United States. Journal of Urban Economics 63(3), 759–74. doi: 10.1016/j.jue.2007.08.003

Damm, Anna (2009): Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. Journal of Labor Economics 27(2), 281–314. doi: 10.1086/599336

Danzer, Alexander; Firat Yaman (2016): Ethnic Concentration and Language Fluency of Immigrants: Evidence from the Guest-Worker Placement in Germany. Journal of Economic Behavior and Organization 131, 151–65. doi: 10.1016/j.jebo.2016.08.012

Edin, Per-Anders; Peter Fredriksson,; Olof Åslund (2003): Ethnic enclaves and the economic success of immigrants : Evidence from a natural experiment. The Quarterly Journal of Economics 118(1), 329–357. doi: 10.1162/00335530360535225

Evans, William; Wallace Oates; Robert Schwab (1992): Measuring peer group effects : A study of teenage behavior. Journal of Political Economy 100(5), 966–991.

Furtado, Delia; Tao Song (2015): Intermarriage and Socioeconomic Integration: Trends in Earnings Premiums among US Immigrants Who Marry Natives. The Annals of the American Academy of Political and Social Science 662(1), 207–22. doi: 10.1177/0002716215594629

Glitz, Albrecht (2014): Ethnic segregation in Germany. Labour Economics 29, 28–40. doi: 10.1016/j.labeco.2014.04.012

Grönqvist, Hans (2006): Ethnic enclaves and the attainments of immigrant children. European Sociological Review 22(4), 369–382. doi: 10.1093/esr/jcl002

Ihlanfeldt, Keith; David Sjoquist (1998): The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform. Housing Policy Debate 9(4), 849–892. doi: 10.1080/10511482.1998.9521321

Ioannides, Yannis; Linda Loury (2004): Job information networks, neighborhood effects, and inequality. Journal of Economic Literature 42(4), 1056–1093. doi: 10.1257/0022051043004595

Munshi, Kaivan (2003): Networks in the modern economy: Mexican migrants in the US labor market. The Quarterly Journal of Economics 118(2), 549–599.

Olden, Andreas; Jarle Møen (2022): The triple difference estimator. The Econometrics Journal 25(3), 531–553.

Patel, Krishna;Francis Vella (2013): Immigrant Networks and Their Implications for Occupational Choice and Wages. Review of Economics and Statistics 95(4) 1249–77. doi: 10.1162/REST_a_00327

Peri, Giovanni; Chad Sparber (2011): Highly Educated Immigrants and Native Occupational Choice. Industrial Relations 50(3) 385–411. doi: 10.1111/j.1468-232X.2011.00643.x

Ravallion, Martin; Emanuela Galasso; Teodoro Lazo; Ernesto Philipp (2005): What Can Ex-Participants Reveal about a Program's Impact? Journal of Human Resources 40(1) 208–30. doi: 10.3368/jhr.XL.1.208

Ruggles Steven; Katie Genadek; Ronald Goeken; Josiah Grover; Matthew Sobek (2015): Integrated Public Use Microdata Series. Minneapolis: University of Minnesota. http://doi.org/10.18128/D010.V6.0

Song, T., & Xu, H. (2020). Anywhere They Go, We Go: Immigration Inflow's Impact on Co-ethnic Natives in the U.S. Southern Economic Journal, 87(1), 191–215.

U.S. Census. (2022): American Community Survey 5-year Data (2009 - 2021). Last modified: December 08, 2022. https://www.census.gov/data/developers/data-sets/acs-5year.html

Wahba, Jackline; Yeves Zenou (2005): Density, social networks and job search methods: Theory and application to Egypt. Journal of Development Economics 78(2), 443–473. doi: 10.1016/j.jdeveco.2004.11.006

Yelowitz, Aaron (1995): The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions. The Quarterly Journal of Economics 110(4) 909–39. doi: 10.2307/2946644

Zong, Jie; Jeanne Batalova (2016): College-educated immigrants in the United State, in: Migration Policy. https://www.migrationpolicy.org/article/college-educated-immigrants-united-states-2014. Accessed 15 December 2016

Appendix A

In Appendix A, we use 2000 and 2010 data to replicate the main model results in Cutler et al. (2008). We also test the validity of Cutler et al.'s (2008) instrumental variable.

$$y_{igm} = \alpha + \beta ISO_{gm} + \phi\left(ISO_{gm} \cdot \overline{S}_{gm}\right) + \theta share_{gm} + X_{igm}\Gamma + Edu_{igm}\Lambda + \mu_g + \delta_m + u_{igm} \quad \text{(A.1)}$$

Equation (A.1) is the model used in Cutler et al. (2008). All notations are the same as the models in Section 4. Cutler et al. (2008) included MSA and group fixed effects (μ_g and δ_m) to eliminate unobserved differences across MSAs and ethnic groups. In addition, they also used mean years of ethnic group *g* in MSA *s* since immigration (referred to as mean years since immigration in the rest of the appendix) to correct for the potential endogeneity of the isolation index.

In Column 1 of Table A.1, when we only control for MSA and group fixed effects, the coefficient of the isolation index is insignificant, although its sign changes as we vary data years. After using mean years since immigration to instrument the isolation index, the estimated segregation effects are all statistically insignificant.

In the $ISO \times \overline{S}$ model, Cutler et al. (2008) found that, if they only controlled for MSA and group fixed effects, the main effect of segregation on immigrants' wages is positive and

Table A.1 Cutler et al. (2008) Model, Earning Results

Dependent Var: log earning	***ISO* Model**		***ISO*, *ISO*×$\overline{S}$ Model**			
	(1)	**(2)**	**(3)**		**(4)**	
	OLS+FE	**IV**	**OLS+FE**		**IV**	
	ISO	**ISO**	**ISO**	**ISO×$\overline{S}$**	**ISO**	**ISO×$\overline{S}$**
Cutler et al. (2007)	0.250	2.963**	2.965*	−0.331*	−14.021***	1.637***
	(0.512)	(1.525)	(1.600)	(0.183)	(5.212)	(0.568)
N	22,564	22,564	22,564		22,564	
2000, 60 groups (age 20-30)	−0.280**	1.557	−2.034	0.229	−3.977	0.361
	(0.142)	(0.971)	(5.157)	(0.356)	(3.118)	(0.221)
N			97,861			
2010, 74 groups (age 20-30)	0.256*	1.048	−0.425	0.062	5.618	−0.319
	(0.143)	(0.703)	(0.547)	(0.049)	(8.350)	(0.596)
N			98,672			
2000, 60 groups (age 20-30, arrived after age 17)	−0.361*	−0.308	1.300*	−0.162**	132.097	−8.553
	(0.187)	(1.667)	(0.723)	(0.067)	(4536.585)	(293.454)
N			53,690			
2010, 74 groups (age 20-30, arrived after age 17)	0.304	−0.277	0.051	0.023	36.310	−2.641
	(0.203)	(0.988)	(0.767)	(0.070)	(105.392)	(7.558)
N			51,994			

Notes: We control for MSA and country of origin fixed effects in the OLS model. All the control variables are the same as models in Cutler et al. (2008). The instrumental variable used for IV estimation is mean years of ethnic group *g* in MSA *s* since immigration. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

an increase in group average education decreases the positive segregation effects. After they instrumented the isolation index by mean years since immigration, the signs of the coefficients are all reversed. However, we cannot replicate these results by using our data. This may imply that the finding that high group average education can moderate the negative segregation effects on immigrants' wages might be specific to the 20-30 years old cohort group in 1990. Otherwise, there may exist unobserved time trends that influence the estimated segregation effects.

Next, we show the results of the balancing tests by using the Cutler et al. (2008) model. In addition to the isolation index, we also run balancing tests for Cutler et al. (2008)'s instrumental variable. Since it is used for correcting the endogeneity of the isolation index, the instrumental variable should satisfy the exclusion restriction—only influencing immigrants' labor market outcomes through ethnic segregation. However, if the instrument is correlated with immigrants' individual characteristics, it implies that selection bias may still present. For example, if we find that males are negatively correlated with mean years since immigration, then it indirectly suggests that males are less likely to sort into highly-segregated ethnic enclaves (since isolation intensity is positively correlated with mean years since immigration (Cutler et al., 2008)). Consequently, we are more likely to observe negative segregation effects.

We first run MSA and country of origin fixed effects regressions by year. The results are shown in Table A.2. We find that there is a steady negative relationship between individual education and the isolation index in both years' results. This implies that immigrants with

Table A.2 Balancing Tests for Isolation Index and Cutler et al. (2007) Instrument Variable

Dependent Variable:	**Isolation Index × 1000**		**Mean Years of Immigrant Group *g* in MSA *s* since immigration**	
	(1)	**(2)**	**(3)**	**(4)**
	2000	**2010**	**2000**	**2010**
Education	−0.663***	−0.466***	0.002	0.007
	(0.137)	(0.110)	(0.005)	(0.005)
Age	−0.037	0.072	0.020***	0.021***
	(0.056)	(0.046)	(0.005)	(0.004)
Male	−0.518	−0.315	−0.103***	−0.157***
	(0.343)	(0.357)	(0.029)	(0.031)
Black	1.573	−1.077	−0.113	−0.209
	(2.385)	(2.571)	(0.270)	(0.355)
Other Nonwhite	−1.177	1.382	−0.370**	−0.530**
	(1.325)	(1.497)	(0.157)	(0.197)
F-test	6.48	4.07	6.71	10.53
P-value	0.000	0.001	0.000	0.000
N	53,690	51,994	53,690	51,994

Notes: Sample is restricted to immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. MSA and country of origin fixed effects are included in all the regressions. The isolation index is inflated by 1000. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level. The F-statistic is for the joint effect of all the individual characteristics that are included in the regressions.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

higher levels of education are less likely to reside in their own ethnic enclaves. The values of the *F*-test indicate that individual characteristics can explain the isolation index as a group. The results of the mean years since immigration balancing tests are shown in Columns 3 and 4. The correlations between individual characteristics and mean years since immigration are even stronger. This evidence suggests that there are still selection biases associated with immigrants' sorting behaviors even after we control for differences across MSAs and ethnic groups.

We also compare the results of the MSA, group fixed effects model, and triple differences model by using multiple cross-sections of data. In Table A.3, individual characteristics and the interactions between immigrants' education levels and other individual characteristics have strong correlations with both isolation index and mean years since immigration in the fixed effects model. When we switch to the triple differences model, the coefficients of individual characteristics are all insignificant. The values of the *F*-test show that individual characteristics cannot explain the isolation index as a group. The joint effects of all the interaction terms are also insignificant. These results indicate that, compared with the MSA and group fixed effects model, the triple differences model can better explain immigrants' systematic selections into ethnic enclaves.

Regarding the results for mean years since immigration, although the triple differences model cannot explain all the correlations between individual characteristics (or interaction

Table A.3 Balancing Tests, MSA and Group Fixed Effects vs. Triple Differences Model

Panel 1. Dependent Variable: Isolation Index × 1000				
	MSA FE, Group FE		**Triple Differences**	
Education	−0.607***	−0.171	−0.007	0.042
	(0.108)	(0.449)	(0.007)	(0.060)
Age	0.006	0.402**	−0.006	0.014
	(0.061)	(0.196)	(0.006)	(0.021)
Male	−0.637**	−3.671***	0.053	0.144
	(0.290)	(1.239)	(0.042)	(0.127)
Black	1.682	1.804	−0.285	−0.000
	(1.793)	(5.049)	(0.313)	(0.870)
Other Nonwhite	0.506	-4.596	0.088	−0.026
	(1.047)	(3.064)	(0.215)	(0.395)
Age × Education	-	−0.038**	-	−0.002
		(0.017)		(0.002)
Male × Education	-	0.262**	-	−0.008
		(0.109)		(0.011)
Black × Education		−0.028	-	−0.022
		(0.342)		(0.060)
Other Nonwhite × Education	-	0.392*	-	0.009
		(0.210)		(0.027)
F-test	7.76	6.26	1.05	0.45
P-value	0.000	0.000	0.385	0.771
N	104,555		102,475	

(*Continued*)

Table A.3 Continued

Panel 2. Dependent Variable: Mean Years of Immigrant Group *g* in MSA *s* since immigration				
	MSA FE, Group FE		**Triple Differences**	
Education	0.040***	0.206***	0.002**	−0.010
	(0.006)	(0.037)	(0.001)	(0.009)
Age	0.052***	0.100***	0.002**	−0.002
	(0.005)	(0.010)	(0.001)	(0.002)
Male	−0.155***	0.074	−0.006	0.006
	(0.028)	(0.064)	(0.005)	(0.013)
Black	−0.022	−0.534	−0.002	−0.107
	(0.264)	(0.609)	(0.057)	(0.126)
Other Nonwhite	−0.178	0.228	−0.012	−0.028
	(0.141)	(0.395)	(0.039)	(0.086)
Age × Education	-	−0.005***	-	0.000**
		(0.001)		(0.000)
Male × Education	-	−0.020***	-	−0.001
		(0.006)		(0.001)
Black × Education	-	0.040	-	0.008
		(0.043)		(0.008)
Other Nonwhite × Education	-	−0.031	-	0.001
		(0.028)		(0.006)
F-test	39.86	11.77	2.80	1.58
P-value	0.000	0.000	0.016	0.178
N	104,555		102,475	

Notes: The isolation index is inflated by 1000. The sample is restricted to immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level. The values of F-statistic in column 1 and 3 are for the joint effect of all the individual characteristics that are included in the regressions. The values of F-statistic in column 2 and 4 are for the joint effect of all the interaction terms. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Therefore, singleton observations are dropped in the triple differences model.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

terms between individual education and other characteristics) and mean years since immigration, it largely decreases the value of the *F*-test. Because mean years since immigration fails the balancing tests, we did not use it to instrument the isolation index in our main model.

Table A.4 Additional Robustness Tests Results

Dependent variable: log earning						
	Baseline		**Drop blacks**		**Drop English speakers**	
ISO	0.415	0.327	0.656	0.558	1.401	1.324
	(0.847)	(0.837)	(0.863)	(0.852)	(0.904)	(0.893)
$ISO \times S$	−0.160*	−0.216**	−0.174**	−0.235**	−0.178**	−0.227**
	(0.083)	(0.096)	(0.087)	(0.103)	(0.086)	(0.098)
$ISO \times \overline{S}$	0.117	0.098	0.147	0.127	0.359**	0.344**
	(0.169)	(0.168)	(0.171)	(0.170)	(0.176)	(0.174)
$ISO \times \overline{S} \times S$	−0.036*	-	−0.039**	-	−0.041**	-
	(0.018)		(0.019)		(0.019)	
$ISO \times \overline{SH} \times S$	-	0.505***	-	0.546***	-	0.511***
		(0.138)		(0.142)		(0.143)
$ISO \times \overline{SL} \times S$	-	−0.048**	-	−0.052**	-	−0.051**
		(0.021)		(0.022)		(0.022)
F-test	1.10	3.09	1.24	3.39	2.32	4.03
P-value	0.354	0.009	0.293	0.005	0.055	0.001
N	102,475		99,939		93,251	
Dependent variable: if employed, =1						
	Baseline		**Drop blacks**		**Drop English speakers**	
ISO	−0.012	−0.018	−0.097	−0.104	0.023	0.020
	(0.249)	(0.249)	(0.247)	(0.247)	(0.261)	(0.261)
$ISO \times S$	0.005	−0.001	0.002	−0.003	0.007	0.005
	(0.017)	(0.018)	(0.017)	(0.018)	(0.017)	(0.017)
$ISO \times \overline{S}$	0.026	0.025	0.015	0.014	0.039	0.038
	(0.056)	(0.056)	(0.055)	(0.055)	(0.059)	(0.059)
$ISO \times \overline{S} \times S$	0.003	-	0.002	-	0.004	-
	(0.004)		(0.004)		(0.004)	
$ISO \times \overline{SH} \times S$	-	0.042	-	0.047	-	0.024
		(0.037)		(0.038)		(0.034)
$ISO \times \overline{SL} \times S$	-	0.002	-	0.001	-	0.003
		(0.004)		(0.004)		(0.004)
F-test	0.57	0.68	0.59	0.74	0.76	0.69
P-value	0.686	0.635	0.672	0.590	0.552	0.633
N	127,769		124,480		116,493	

Notes: To center the variables, we replace ISO_{gmt}, $\overline{S}_{gmt}$, and S_{igmt} by $ISO_{gmt} - \overline{ISO}_{mt}$, $\overline{S}_{gmt} - \overline{\overline{S}}_{mt}$, and $S_{igmt} - \overline{S}_{gt}$, respectively. The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for $\overline{S} \times S$. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. Therefore, singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

Appendix B. Parallel Trends of the Triple Differences Model

We test the validity of the triple differences model by conducting balancing tests in Section 5. The results show that after controlling for MSA-specific time trends, group-specific time trends, and time-invariant MSA-specific ethnic group effects, isolation intensity does not vary significantly by immigrants' demographic characteristics, years of schoolings, and their interaction terms. This suggests that our model minimizes immigrants' sorting into ethnic enclaves based on observed characteristics.

To provide additional evidence that our model estimates causal effects, in this appendix, we test if the parallel trend assumption holds in our triple differences model. We focus on the parallel trends in log earnings because our main findings and conclusions come from earnings regressions. Since our sample (immigrants aged 20-30 years old and moved to the U.S. after age 17) includes 59 ethnic groups and 267 MSAs, we start with a simplified model setup to illustrate the concept of parallel trend. Given earnings of two ethnic groups in MSA *A* and MSA *B* in 2010, and the treatment being ethnic segregation, we use G_H to denote the ethnic group with high level of segregation and G_L to denote the ethnic group with low level of segregation. Similar to Equation (2) in Section 4, earnings of immigrants from group *g* in MSA *m* in year *t* is represented by y_{gmt}. We then use y^0_{gmt} and y^1_{gmt} to distinguish the earnings of immigrants who are being exposed to high level of segregation and the earnings of immigrants who are being exposed to low level of segregation, respectively. Equation (B.1) shows the parallel trend assumption for a difference-in-differences model:

$$E\left(y^0_{gmt}|MSA_A=1,G_H=1,t_{2010}=1\right)-E\left(y^0_{gmt}|MSA_A=0,G_H=1,t_{2010}=1\right)=$$
$$E\left(y^0_{gmt}|MSA_A=1,G_H=0,t_{2010}=1\right)-E\left(y^0_{gmt}|MSA_A=0,G_H=0,t_{2010}=1\right) \quad \text{(B.1)}$$

The Equation shows that the double differences model estimates causal effects if the two ethnic groups have the same earnings differences across MSAs when being exposed to low level of ethnic segregation. However, this assumption is rarely true, because even when both groups were not being exposed to ethnic segregation, earnings differences across MSAs for the two ethnic groups may still be dissimilar due to their underlying differences in unobserved abilities, skill sets, and access to labor market resources. However, if we know the earnings of the two ethnic groups in 2000, and the intensity of ethnic segregation for both groups is low in 2000, we can use a triple differences model to remove the differences between the two groups' earnings differences across MSAs. Equation (B.2) shows the parallel trend assumption for the triple differences model:

$$\left[E\left(y^0_{gmt}|MSA_A=1,G_H=1,t_{2010}=1\right)-E\left(y^0_{gmt}|MSA_A=0,G_H=1,t_{2010}=1\right)\right]-$$
$$\left[E\left(y^0_{gmt}|MSA_A=1,G_H=0,t_{2010}=1\right)-E\left(y^0_{gmt}|MSA_A=0,G_H=0,t_{2010}=1\right)\right]=$$
$$\left[E\left(y^0_{gmt}|MSA_A=1,G_H=1,t_{2010}=0\right)-E\left(y^0_{gmt}|MSA_A=0,G_H=1,t_{2010}=0\right)\right]-$$
$$\left[E\left(y^0_{gmt}|MSA_A=1,G_H=0,t_{2010}=0\right)-E\left(y^0_{gmt}|MSA_A=0,G_H=0,t_{2010}=0\right)\right] \quad \text{(B.2)}$$

The expression on the left-hand side of Equation (B.2) is the same as the double differences in Equation (B.1). The expression on the right-hand side is the double differences for year 2000. According to Olden and Møen (2022), the triple differences model does not require the parallel trend assumption to hold for both double differences at the same time, i.e., the expressions on both sides do not need to be zero. Instead, we only need the biases in the two sets of double differences to be the same.

To test the condition in Equation (B.2), we can transform it to the double differences format. If we use Δy_{gt}^{0} to represent earnings differences across MSAs for group *g* in year *t*, then we get:

$$\left[E\left(\Delta y_{gt}^{0}|G_H=1,t_{2010}=1\right)-E\left(\Delta y_{gt}^{0}|G_H=0,t_{2010}=1\right)\right]=$$

$$\left[E\left(\Delta y_{gt}^{0}|G_H=1,t_{2010}=0\right)-E\left(\Delta y_{gt}^{0}|G_H=0,t_{2010}=0\right)\right] \tag{B.3}$$

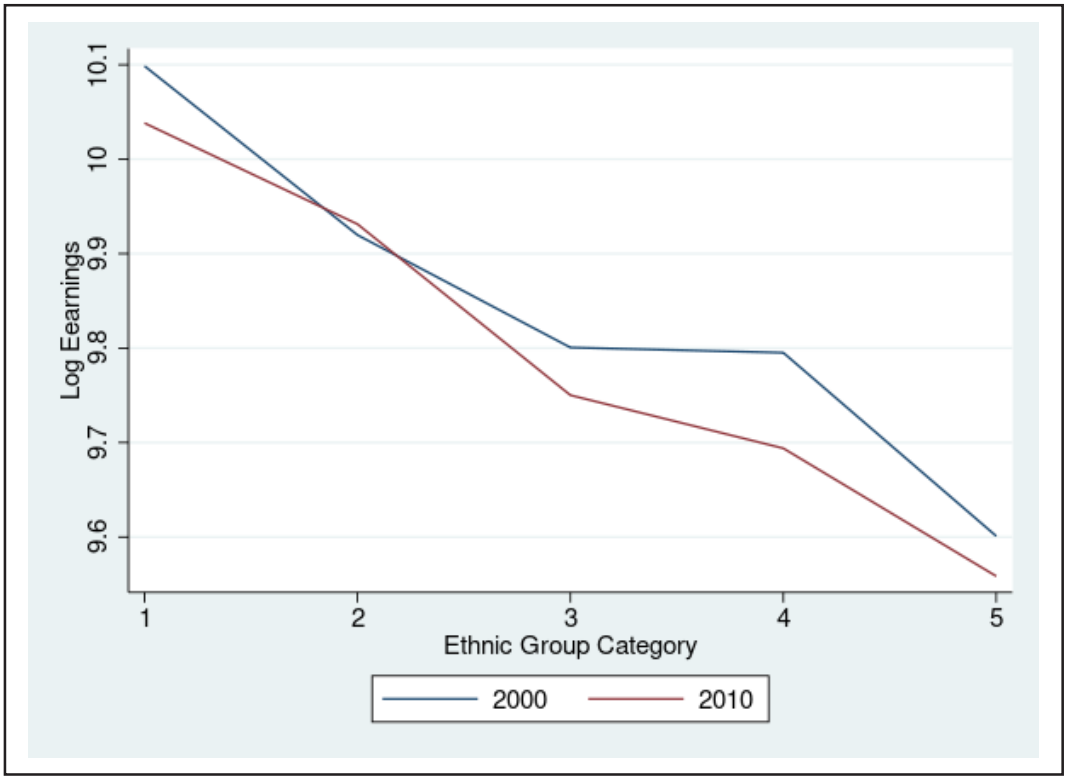
Thus, we can prove the parallel trend in a triple differences model by using the methods that are used to prove the parallel trends in double differences models. In practice, previous literature often includes multiple pre-treatment time periods to check if the trends in outcomes are similar for the treatment and control groups before the treatment becomes effective. Alternatively, some estimate the difference-in-differences regression model with pre-treatment data, and if the resulting treatment effect is insignificant, then the parallel trend condition holds.

It is, however, difficult to adopt these methods in this paper because of the following reasons. First, our sample only includes two time periods, so we are not able to observe trends in earnings during pre-treatment time periods or estimate pre-treatment effects. Second, even if we add earlier data into the sample, our treatment, ethnic segregation intensity, is a continuous variable. One could argue ethnic concentration is a reality as long as there are immigrants, so it is difficult to pinpoint a "pre-treatment" period. Instead, our model compares outcomes of immigrants residing in more segregated areas with those who are living in less segregated areas. Finally, as mentioned above, there are 59 groups and 267 MSAs in our sample. It is more complicated and visually cumbersome to present the parallel trends in our paper in ways similar to those done in previous research. Therefore, we need to modify previous approaches to present a practical visual representation of the parallel trend. Specifically, we need to reduce the number of ethnic groups and MSAs.

Here, we present two methods for testing the parallel trend assumption visually for our triple differences model. In method 1, we divide ethnic groups into five equally sized categories based on their segregation intensities. The ethnic groups in each category have similar isolation indices, except for the top outlier group. We then use ethnic group categories as the x-axis. The labels on the x-axis only represent ranks of these categories, i.e., the ethnic groups on the left have lower levels of segregation and the ethnic groups on the right have higher levels of segregation.

In Figure B.1., we show the trends in average log earnings at ethnic group level in 2000 and 2010 separately. This is equivalent to a difference-in-differences model. The figure shows that the overall average log earnings are higher in 2000. As segregation intensity increases, average log earnings decrease. However, the trends are somewhat different across years. This implies the double differences model is not sufficient in removing selection bias, necessitating a third dimension of differences.

Figure B.1 Differences-in-Differences Trends, Method 1.



To present the triple differences in a two-dimensional graph, we first estimate the triple differences model and then subtract the ethnic group-specific time trends from each person's log earnings. This serves the same function as calculating earnings differences across MSAs within ethnic groups. To see this, consider the specification in Equation (B.4). It is the triple differences regression in Equation (2) in Section 4, but without control variables for brevity.

$$y_{igmt} = \alpha + \beta ISO_{gmt} + \mu_{gt} + \delta_{mt} + \tau_{gm} + u_{imgt} \tag{B.4}$$

When calculating delta log earnings between MSA_A and MSA_B for ethnic group G_H in 2010 (the first line of Equation (B.2)), we can plug in the expression in Equation (B.4) for each $E[y_{gmt}]$. As shown in Equation (B.5), the common factor between the two expressions in the second line, $\mu_{H,2010}$, is cancelled out. The rest MSA-specific time trends and ethnic group-specific MSA factors will be removed through later differences as shown in Equation (B.2). Similarly, the other three terms in Equation (B.2) remove $\mu_{L,2010}$, $\mu_{H,2000}$, and $\mu_{L,2000}$.

$$\left[E\left(y_{gmt}|MSA_A=1,G_H=1,t_{2010}=1\right)-E\left(y_{gmt}|MSA_A=0,G_H=1,t_{2010}=1\right)\right]=$$

$$\left(\alpha+\beta ISO_{AH,2010}+\mu_{H,2010}+\delta_{A,2010}+\tau_{AH}\right)-\left(\alpha+\beta ISO_{BH,2010}+\mu_{H,2010}+\delta_{B,2010}+\tau_{BH}\right)=$$

$$\left(\beta ISO_{AH,2010}-\beta ISO_{BH,2010}\right)+\left(\delta_{A,2010}-\delta_{B,2010}\right)+\left(\tau_{AH}-\tau_{BH}\right) \tag{B.5}$$

We then calculate the average of log earnings that exclude μ_{gt} at ethnic group level and use them as the y-axis in Figure B.2. We can see that the trends now become relatively more parallel than the trends in Figure B.1. Technically, if the treatment is effective, the trends should become less parallel as immigrants being exposed to higher levels of ethnic segregation. Since

Figure B.2 Triple Differences Trends, Method 1.

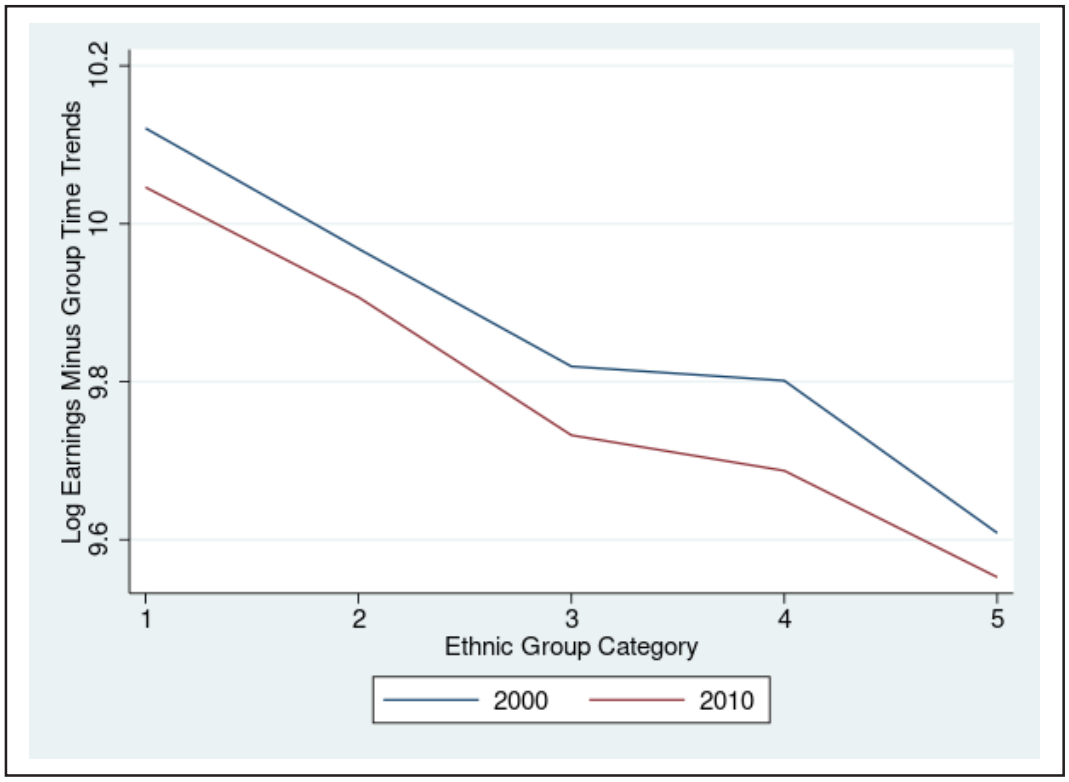
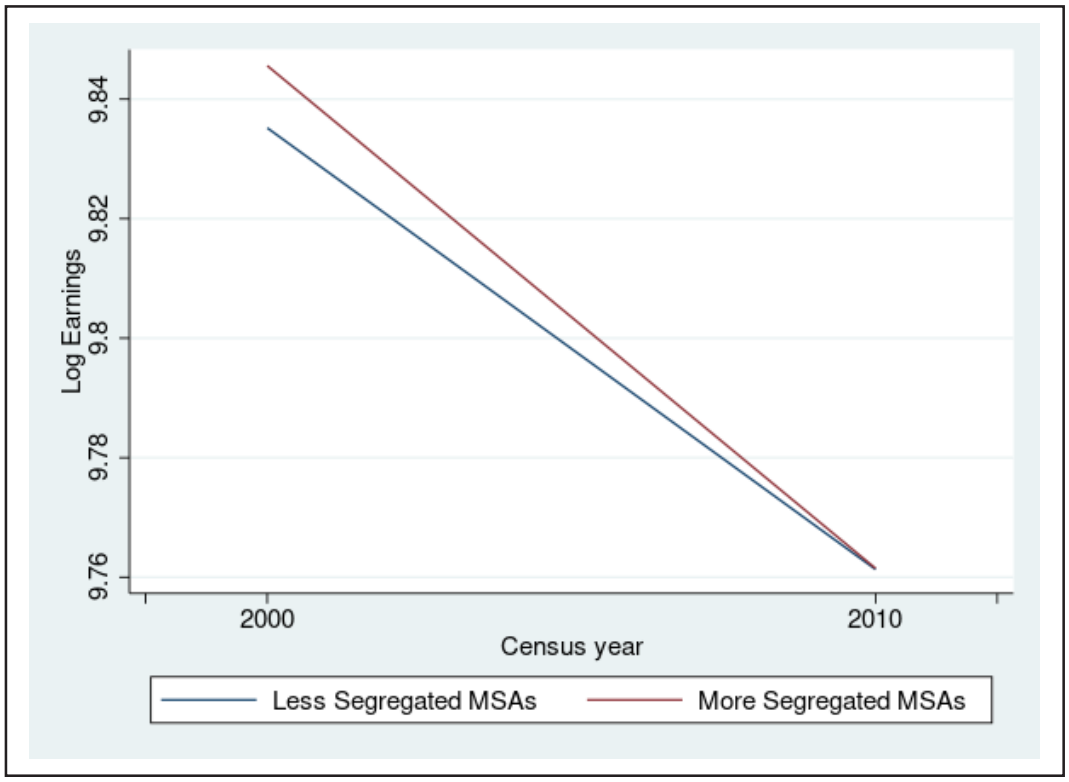


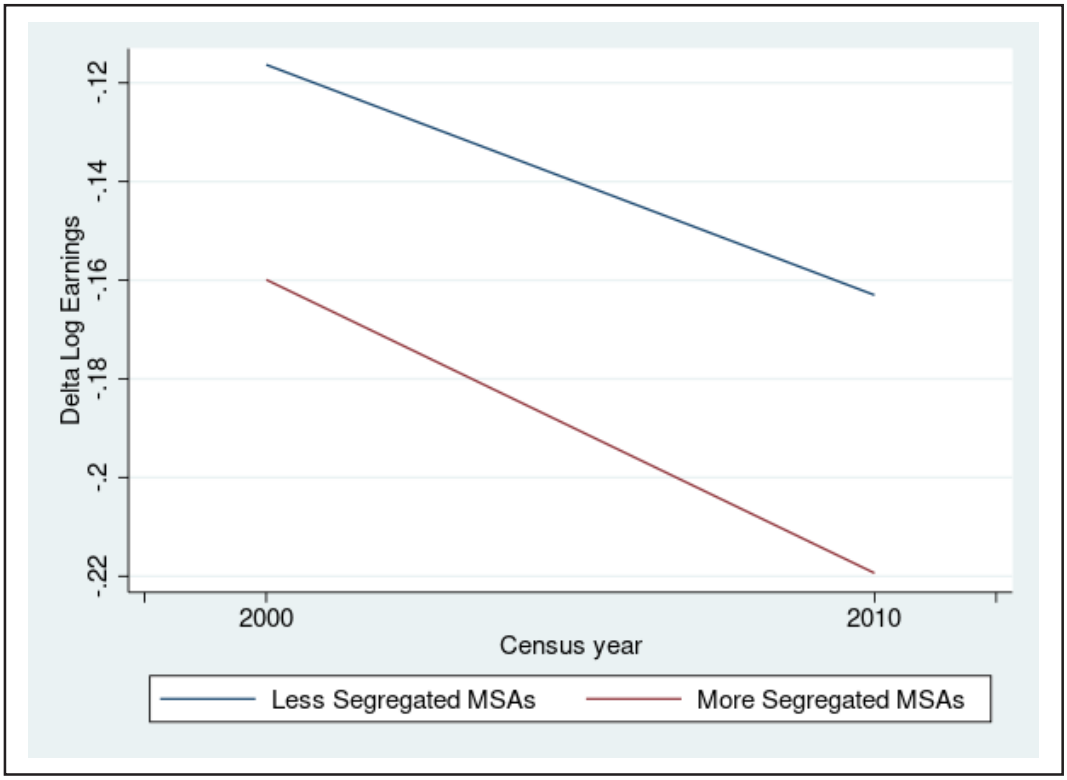
Figure B.3 Differences-in-Differences Trends, Method 2.



our estimated results in Table 3 show that the coefficients of isolation index are all statistically insignificant, the empirical results are consistent with what we observe in Figure B.2.

In method 2, we first divide all MSAs into two broad groups based on whether their isolation indices are greater or lower than MSA-level average. In Figure B.3, we show average

Figure B.4 Triple Differences Trends, Method 2.



log earnings by broad MSA groups. This figure again shows that the parallel trend assumption does not hold for the double differences model. The average earnings in highly segregated MSAs reduce in a faster pace than the average earnings in less segregated MSAs. Therefore, within each MSA, we further divide ethnic groups into two broad categories, i.e., whether having higher or lower than the group-level average segregation intensity. Then within each MSA, we calculate log earnings differences between highly segregated ethnic groups and less segregated ethnic groups and aggregate them to the broad MSA group level. This step removes some MSA-specific time trends. Figure B.4 again shows that the triple differences trends are relatively more parallel than the double differences trends in Figure B.3.

In summary, both methods show that the parallel trend assumption holds for our triple differences model and our model is more effective in minimizing selection bias than a double differences model. However, the graphs we present in this appendix are based on largely simplified scenarios. In reality, the trends and comparisons between ethnic groups and MSAs should be more complicated. Therefore, we believe that the evidence in this appendix only supplement but cannot replace the balancing tests results.

Appendix C. Discrete Education Measures

Table C.1 Discrete Education Measure, Triple differences, *ISO* Model and $ISO \times \overline{S}$ Model

Dependent variable: log earning						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
ISO	−0.217	−0.233	−0.136	−1.074	−1.044	−1.142
	(0.469)	(0.436)	(0.416)	(1.047)	(0.963)	(0.934)
ISO × *%HS*	-	-	-	3.608	3.122	3.181
				(3.381)	(3.084)	(3.030)
ISO × *%Some_Coll*	-	-	-	2.358	−0.283	−0.780
				(5.551)	(5.207)	(4.869)
ISO × *%College*	-	-	-	−0.040	1.427	2.001
				(2.317)	(2.190)	(2.075)
%High School	-	-	-	−0.158	−0.145	−0.210*
				(0.130)	(0.124)	(0.122)
%Some College	-	-	-	0.042	0.079	−0.068
				(0.167)	(0.158)	(0.154)
%College	-	-	-	0.651***	0.559***	0.139
				(0.142)	(0.134)	(0.131)
High School	-	-	0.090***	-	-	0.090***
			(0.006)			(0.006)
Some College	-	-	0.185***	-	-	0.185***
			(0.010)			(0.010)
College	-	-	0.465***	-	-	0.463***
			(0.016)			(0.016)
+characteristic control	No	Yes	Yes	No	Yes	Yes
F-test	-	-	-	0.40	0.32	0.44
P-value	-	-	-	0.805	0.868	0.781
N	102,475					

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. The omitted category is high school dropouts. We control for MSA-specific time trends, MSA-specific group effects, and group-specific time trends in all regressions. Characteristic controls include age, gender, marital status, children in the household, black, Asian, and Hispanic dummy variables. The values of F-statistic are for the joint effect of the isolation index and the interaction term between isolation index and group average education. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

Table C.2 Discrete Education Measure, Triple Differences and Triple Interactions Model (Centered)

Dependent variable: log earning			
	(1)	**(2)**	**(3)**
ISO	0.119	−1.398	−2.487**
	(0.414)	(0.936)	(0.978)
ISO × *HS*	−0.113	−0.160	1.632***
	(0.104)	(0.106)	(0.526)
ISO × *Some Coll*	−0.041	−0.025	1.721*
	(0.158)	(0.180)	(0.962)
ISO × *College*	−1.420***	−0.281	2.123*
	(0.419)	(0.499)	(1.197)
ISO × *%HS*	-	2.900	6.802**
		(3.158)	(3.385)
ISO × *%Some Coll*	-	−0.059	4.090
		(5.118)	(5.570)
ISO × *%Coll*	-	3.395	-2.486
		(2.209)	(3.139)
ISO × *%HS* × *HS*	-	-	-4.871***
			(1.698)
ISO × *%HS* × *Some Coll*	-	-	-6.146**
			(3.012)
ISO × *%HS* × *Coll*	-	-	-7.871**
			(3.901)
ISO × *%Some Coll* × *HS*	-	-	-1.560
			(2.872)
ISO × *%Some Coll* × *Some Coll*	-	-	2.251
			(4.501)
ISO × *%Some Coll* × *Coll*	-	-	-8.266
			(6.364)
ISO × *%Coll* × *HS*	-	-	−0.408
			(1.730)
ISO × *%Coll* × *Some Coll*	-	-	−0.426
			(2.611)
ISO × *%Coll* × *Coll*	-	-	6.065*
			(3.231)
%HS × *HS*	-	−0.203*	0.061
		(0.104)	(0.143)
%HS × *Some Coll*	-	−0.175	0.059
		(0.148)	(0.180)
%HS × *Coll*	-	−0.266	−0.003
		(0.192)	(0.203)
%Some Coll × *HS*	-	0.251*	0.364**
		(0.138)	(0.163)
%Some Coll × *Some Coll*	-	0.334**	0.374**
		(0.164)	(0.186)
%Some Coll × *Coll*	-	0.016	0.154
		(0.183)	(0.200)
%Coll × *HS*	-	−0.002	0.025
		(0.076)	(0.088)

(*Continued*)

Table C.2 Continued

Dependent variable: log earning			
	(1)	**(2)**	**(3)**
%Coll × Some Coll	-	0.034	0.043
		(0.090)	(0.097)
%Coll × Coll	-	0.559***	0.447***
		(0.122)	(0.109)
%HS	-	−0.021	−0.204
		(0.126)	(0.141)
%Some Coll	-	−0.250	−0.302
		(0.178)	(0.187)
%Coll	-	−0.218	−0.119
		(0.146)	(0.148)
High School	0.102***	0.151***	0.046
	(0.011)	(0.039)	(0.053)
Some College	0.198***	0.221***	0.143**
	(0.015)	(0.057)	(0.072)
College	0.523***	0.396***	0.340***
	(0.017)	(0.096)	(0.095)
F-test	2.92	0.81	2.72
P-value	0.020	0.576	0.000
N		102,475	

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. The omitted category is high school dropouts. The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for the interactions between individual education levels and average education levels. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

Table C.3 Discrete Education Measure, Triple differences, ISO Model and $ISO \times \bar{S}$ Model

Dependent variable: if employed, =1						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
ISO	−0.106	−0.109	−0.093	0.300	0.293	0.279
	(0.151)	(0.152)	(0.151)	(0.321)	(0.320)	(0.319)
ISO × *%HS*	-	-	-	−1.550*	−1.540*	−1.521*
				(0.906)	(0.912)	(0.909)
ISO × *%Some_Coll*	-	-	-	1.136	0.798	0.856
				(1.578)	(1.621)	(1.607)
ISO × *%College*	-	-	-	−0.136	0.030	0.064
				(0.606)	(0.607)	(0.604)
%High School	-	-	-	0.108**	0.110**	0.096**
				(0.049)	(0.049)	(0.049)
%Some College	-	-	-	0.124**	0.135**	0.103**
				(0.052)	(0.052)	(0.052)
%College	-	-	-	0.036	0.035	−0.002
				(0.046)	(0.046)	(0.047)
High School	-	-	0.014***	-	-	0.014***
			(0.003)			(0.003)
Some College	-	-	0.033***	-	-	0.033***
			(0.004)			(0.004)
College	-	-	0.039***	-	-	0.039***
			(0.004)			(0.004)
+characteristic control	No	Yes	Yes	No	Yes	Yes
F-test	-	-	-	1.01	1.01	1.03
P-value	-	-	-	0.402	0.400	0.388
N				127,769		

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. The omitted category is high school dropouts. We control for MSA-specific time trends, MSA-specific group effects, and group-specific time trends in all regressions. Characteristic controls include age, gender, marital status, children in the household, black, Asian, and Hispanic dummy variables. The values of F-statistic are for the joint effect of the isolation index and the interaction term between isolation index and group average education. We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.

Table C.4 Discrete Education Measure, Triple Differences and Triple Interactions Model (Centered)

Dependent variable: if employed, =1			
	(1)	**(2)**	**(3)**
ISO	−0.039	0.218	0.379
	(0.153)	(0.324)	(0.337)
ISO × *HS*	−0.082*	−0.050	−0.366*
	(0.048)	(0.051)	(0.221)
ISO × *Some Coll*	−0.083	−0.032	−0.374
	(0.073)	(0.079)	(0.367)
ISO × *College*	−0.190***	−0.153**	−0.549*
	(0.061)	(0.073)	(0.311)
ISO × *%HS*	-	−1.309	−1.949*
		(0.950)	(1.027)
ISO × *%Some Coll*	-	0.928	1.610
		(1.680)	(1.985)
ISO × *%Coll*	-	0.248	−0.256
		(0.602)	(1.018)
ISO × *%HS* × *HS*	-	-	1.328*
			(0.717)
ISO × *%HS* × *Some Coll*	-	-	0.762
			(1.146)
ISO × *%HS* × *Coll*	-	-	0.850
			(0.947)
ISO × *%Some Coll* × *HS*	-	-	-1.097
			(1.066)
ISO × *%Some Coll* × *Some Coll*	-	-	0.534
			(1.742)
ISO × *%Some Coll* × *Coll*	-	-	−0.437
			(1.572)
ISO × *%Coll* × *HS*	-	-	0.132
			(0.695)
ISO × *%Coll* × *Some Coll*	-	-	0.104
			(1.119)
ISO × *%Coll* × *Coll*	-	-	0.104
			(1.119)
%HS × *HS*	-	−0.010	−0.060
		(0.039)	(0.050)
%HS × *Some Coll*	-	−0.014	−0.067
		(0.050)	(0.059)
%HS × *Coll*	-	0.020	−0.040
		(0.048)	(0.058)
%Some Coll × *HS*	-	0.018	0.049
		(0.051)	(0.062)
%Some Coll × *Some Coll*	-	0.035	0.028
		(0.056)	(0.067)
%Some Coll × *Coll*	-	0.044	0.043
		(0.053)	(0.063)
%Coll × *HS*	-	0.046*	0.036
		(0.027)	(0.031)

(*Continued*)

Table C.4 Continued

Dependent variable: if employed, =1			
	(1)	**(2)**	**(3)**
%Coll × Some Coll	-	0.045	0.027
		(0.032)	(0.034)
%Coll × Coll	-	0.052*	0.017
		(0.027)	(0.034)
%HS	-	0.073	0.106*
		(0.052)	(0.056)
%Some Coll	-	0.085	0.068
		(0.066)	(0.072)
%Coll	-	−0.050	−0.040
		(0.051)	(0.052)
High School	0.021***	0.016	0.029
	(0.004)	(0.015)	(0.019)
Some College	0.040***	0.032	0.054**
	(0.005)	(0.021)	(0.026)
College	0.049***	0.028	0.058**
	(0.005)	(0.021)	(0.026)
F-test	3.09	1.36	1.44
P-value	0.015	0.219	0.115
N		127,769	

Notes: We use the sample of immigrants who are from age group 20-30 and arrived at the U.S. after age 17 or later. The omitted category is high school dropouts. The values of the F-statistic are for the joint effect of the isolation index and all the interaction terms, except for the interactions between individual education levels and average education levels. Control variables include age, male dummy, whether married, whether have children in the household, black, Asian, Hispanic, group average education (centered), and immigrants' own education (centered). We use STATA command reghdfe (high dimensional fixed effects) to estimate the parameters. The sample size becomes slightly smaller since singleton observations are dropped. Figures in parentheses are standard errors robust to clustering at the MSA by country of origin level.

Significance is defined as follows: ***p<0.01; **p<0.05; *p<0.10.