Working Paper Series

Occupational segregation by race and ethnicity in the US: Differences across states

Carlos Gradín
Coral del Río
Olga Alonso-Villar

ECINEQ WP 2011 – 190

ECINEQ 2011-190
January 2011

www.ecineq.org

Occupational segregation by race and ethnicity in the US: Differences across states*

Carlos Gradín†
Coral del Río
Olga Alonso-Villar
Universidade de Vigo

Abstract

Using the 2005–2007 American Community Survey, we analyze the occupational segregation of workers by race and ethnicity across states. Although the unconditional analysis shows great geographical variation in segregation, with the largest levels in the Southwest, the analysis of segregation conditioned on the distribution of characteristics reveals that segregation of workers with similar characteristics is generally greater in the East Central region. To quantify conditional segregation, we adapt a propensity score technique that simultaneously controls for several characteristics, allowing the identification of the factors that explain the geographical variation of unconditional segregation.

Keywords: occupational segregation, race, ethnicity, states, United States.
JEL Classification: J15, J71, D63.

* Financial support from the Ministerio de Ciencia e Innovación (grants ECO2008-03484-C02-01/ECON and ECO2010-21668-C03-03) and from FEDER is gratefully acknowledged.
† **Addresses of correspondence**: Universidade de Vigo; Facultade de CC. Económicas; Departamento de Economía Aplicada; Campus Lagoas-Marcosende s/n; 36310 Vigo; Spain. Tel.: +34 986813527; fax: +34 986812401; e-mail: cgradin@uvigo.es.

1. Introduction

The United States (US) has an outstanding and ever-increasing racially and ethnically diverse population. The proportion of non-Hispanic whites decreased from 76% in 1990 to 65% in 2009.[1] Minority groups, though, are not evenly distributed across states: Hispanics are more concentrated in California, Texas, and Florida; African Americans in Southeastern states; Asians in California, Hawaii, and New York; and Native Americans (including American Indian, Alaskan, Hawaiian, and Pacific Islander natives) in Alaska, Arizona, New Mexico, Oklahoma, and Hawaii. In this multiracial society, many studies show that minority groups are not equally distributed across neighborhoods in metropolitan areas (Farley, 1977; Massey and Denton, 1987; Cutler et al., 1999; Reardon and Yun, 2001; Iceland, 2004) or educational centers in school districts (Theil and Finizza, 1971; James and Taueber, 1985; Clotfelter, 1999; Frankel and Volij, 2010). This uneven distribution of racial/ethnic groups across neighborhoods (schools) is referred to in the literature as residential (school) segregation by race and ethnicity.

Although residential and especially school segregation by race in the US has been extensively documented, evidence on racial/ethnic inequality in the labor market is scarcer. Some papers have addressed the analysis of segregation across occupations (Albelda, 1986; King, 1992; Watts, 1995; Queneau, 2009; Alonso-Villar et al., 2010; Gradín, 2010) and workplaces (Tomaskovic-Devey et al., 2006; Hellerstein and Neumark, 2008), but only at the national level. To our knowledge, there is no evidence of disparities among states.

To close this gap, this paper aims to analyze the geographical dimension of occupational segregation by race/ethnicity in the US. Thus, we analyze whether occupational disparities among workers self-reporting different race/ethnicity share a pattern common across the country or instead vary state by state. For that purpose, we show differences in occupational segregation by race/ethnicity among states using various multigroup segregation indexes.

[1] See the Census Bureau population estimates by race and ethnicity (http://www.census.gov/).

However, observed differences in segregation levels among states do not necessarily reflect disparities in the performance of minorities. When comparing two states, it could be plausible that the level of observed segregation of one was substantially higher than that of the other due to a compositional effect. This would happen if the state with the highest segregation level had a larger proportion of groups of workers who typically faced stronger segregation. From previous research (Hellerstein and Neumark, 2008; Alonso-Villar et al., 2010), we know that workers with certain characteristics are more likely to be more segregated in the labor market than others: minorities are more segregated than non-Hispanic whites; high and low-educated workers are more segregated than workers with intermediate grades; and recent immigrants, especially if they lack English proficiency, are excluded from a large number of jobs. Thus, states with larger proportions of highly segregated groups are more likely to show higher segregation even if the probability of a worker with certain attributes of over/underrepresentation in some jobs is essentially the same.

It is crucial to separate segregation that can be explained by the specific distribution of characteristics in the state from unexplained segregation. The first part represents the explained segregation, but the second is the conditional segregation and could be entirely attributed to state-specific segregation patterns. Thus, we measure not only unconditional segregation but also conditional segregation of each state based on an estimated counterfactual distribution in which the state is given the relevant characteristics of a state of reference. Relevant characteristics include the racial/ethnic composition, education, immigration profile (including English proficiency), and industrial structure.

To implement this, we borrow and adapt to our context the methodology initially proposed by Di Nardo et al. (1996) to analyze wage disparities, and later adapted by Gradín (2010) for the analysis of conditional segregation of each non-white group with respect to non-Hispanic whites at the national level. According to this propensity score procedure, we construct the counterfactual occupational distribution of each target state by reweighting its original observations of workers by their predicted probability of belonging to the state of reference based on their own characteristics. These probabilities are estimated using a simple logit model. As a result of this procedure, we can compare segregation across states considering a common distribution of relevant characteristics shared by all of them. Segregation disparities among states will arise from differences in the occupational distributions of the groups into which the entire population has been partitioned (i.e, from differences in the occupational distributions of the groups defined according to the above characteristics) and not from discrepancies in their demographic or industrial structures.

In quantifying segregation, most measures only account for disparities between the distributions of two population groups (blacks and whites, for example), so segregation arises when these distributions depart from each other, as happens when using the index of dissimilarity proposed by Duncan and Duncan (1955). These measures are weak in that when more than two groups are involved (not only blacks and whites, but also Hispanics, Asians, and Native Americans), binary comparisons among groups become cumbersome, and no summary segregation can be obtained. Fortunately, in recent years, new measures have been proposed that allow the analysis of segregation in a multigroup context by simultaneously quantifying the disparities among all groups. This is the case of the mutual information index, recently characterized by Frankel and Volij (2010) in terms of basic axioms; the *IP* index proposed by Silber (1992); and the Gini index proposed by Reardon and Firebaugh (2002), which will be used in this paper.[2] These multigroup segregation indices provide a methodological advantage over most segregation analyses (including studies not only of occupational but also residential and school segregation) because they mainly use dichotomous measures.[3]

The paper is structured as follows. Section 2 introduces several segregation measures and shows occupational segregation by race/ethnicity in the US at the state level. Section 3 displays conditional segregation according to the main characteristics of states. This allows us to explain spatial disparities across states. Section 4 summarizes the main conclusions.

[2] The Gini index we refer to is the unbounded version.
[3] Exceptions are Watts (1995) and Iceland (2004). The former analyzes multigroup segregation by using the *IP* index, and the latter uses the entropy index proposed by Reardon and Firebaugh (2002).

2. Measuring segregation across states

2.1 Segregation indexes

Following Reardon and O'Sullivan (2004, p. 122), "[S]egregation can be thought of as the extent to which individuals of different groups occupy and experience different social environments." Different dimensions of the problem have been described in the literature (Massey and Denton, 1988), with evenness the most popular (Duncan and Duncan, 1955; James and Taueber, 1985; Reardon and Firebaugh, 2002; Hutchens, 2004). According to this perception (followed in this paper), segregation exists if the population subgroups into which the economy can be partitioned in a mutually exclusive manner are unevenly distributed among organizational units (neighborhoods, schools, occupations).

In the case of occupational segregation by race and ethnicity, differences among the distributions of these groups across jobs may arise from several sources. Thus, they could be the result of discriminatory employers' views or attitudes toward some demographic groups. In addition, educational disparities may shape the type of jobs to which these groups can apply. Language and cultural differences also affect the range of jobs that workers coming from other countries are offered (Maxwell, 2010), especially if the number of years of residence in the US is low. Moreover, the job opportunities of newly-arrived immigrants are likely to depend on migrant networks (Hellerstein et al., 2010), which may reinforce the concentration of immigrants of a race/ethnic group in occupations/establishments with a high presence for that group (Patel and Vella, 2007).

Given that we are interested in measuring segregation in a multigroup context, we use multigroup indexes rather than the popular indexes employed in binary comparisons. In what follows, we present the three indexes used in our empirical analysis: *M*, *IP*, and *G*.

The mutual information index (*M*), borrowed from the information theory, measures the reduction in the uncertainty of the distribution of employment among occupations due to the knowledge of the distribution of population among racial/ethnic groups. It can be written as

$$M = \sum_g \frac{C^g}{T} \log\left(\frac{T}{C^g}\right) - \sum_j \frac{t_j}{T}\left[\sum_g \frac{c_j^g}{t_j} \log\left(\frac{t_j}{c_j^g}\right)\right],$$

where C^g is the size of racial/ethnic group *g*; *T* represents total population; c_j^g is the number of individuals of group *g* in occupation *j*; and t_j is the size of occupation *j*. This index has been recently characterized in terms of basic axioms and used to quantify school segregation by race in the US (Frankel and Volij, 2010).

The *IP* index proposed by Silber (1992),

$$IP = \frac{1}{2}\sum_g \sum_j \left| \frac{c_j^g}{T} - \frac{C^g}{T}\frac{t_j}{T} \right|,$$

is the generalization of the popular index of dissimilarity to the multigroup case according to the proposal previously made by Karmel and MacLachlan (1988) in the dichotomous case.

Finally, our Gini index corresponds to the unbounded version of the measure proposed by Reardon and Firebaugh (2002):

$$G = \frac{1}{2}\sum_g \sum_{i,j} \frac{t_i}{T}\frac{t_j}{T} \left| \frac{c_i^g}{t_i} - \frac{c_j^g}{t_j} \right|.$$

As shown by Alonso-Villar and Del Río (2010), multigroup segregation indexes *M*, *IP*, and *G* can be written as the sum of the segregation level of each group into which the economy is partitioned, weighted by the demographic weight of the group. For example, the mutual information index can be written as $M = \sum_g \frac{C^g}{T} M^g$, where $M^g = \sum_j \frac{c_j^g}{C^g} \ln\left(\frac{c_j^g / C^g}{t_j / T} \right)$ represents the segregation of group *g* (according to the Theil index that results from comparing the distribution of group *g* with the distribution of total jobs across occupations). Consequently, the contribution of race/ethnicity to the overall segregation of the state depends on both its demographic weight and the disparities between the employment distribution of that group and the occupational structure of the state. This is important for an adequate interpretation of variation in segregation across states in a country like the US, with such racial/ethnic diversity. Indeed, a low level of overall segregation in a state does not rule out the case that some minorities are highly segregated when their relative sizes are small.

2.2 Data

The data used in this section come from the 2005–2007 Public Use Microdata Sample (PUMS) files of the American Community Survey (ACS) conducted by the US Census Bureau. This survey was conducted throughout the US using a series of monthly samples jointly accounting for 3% of the overall population living in housing units during the period (and 2% of those living in-group quarters during 2006 and 2007). This survey provides a variety of information on demographic and labor-related characteristics reflecting the labor market performance right before the 2008 economic recession.

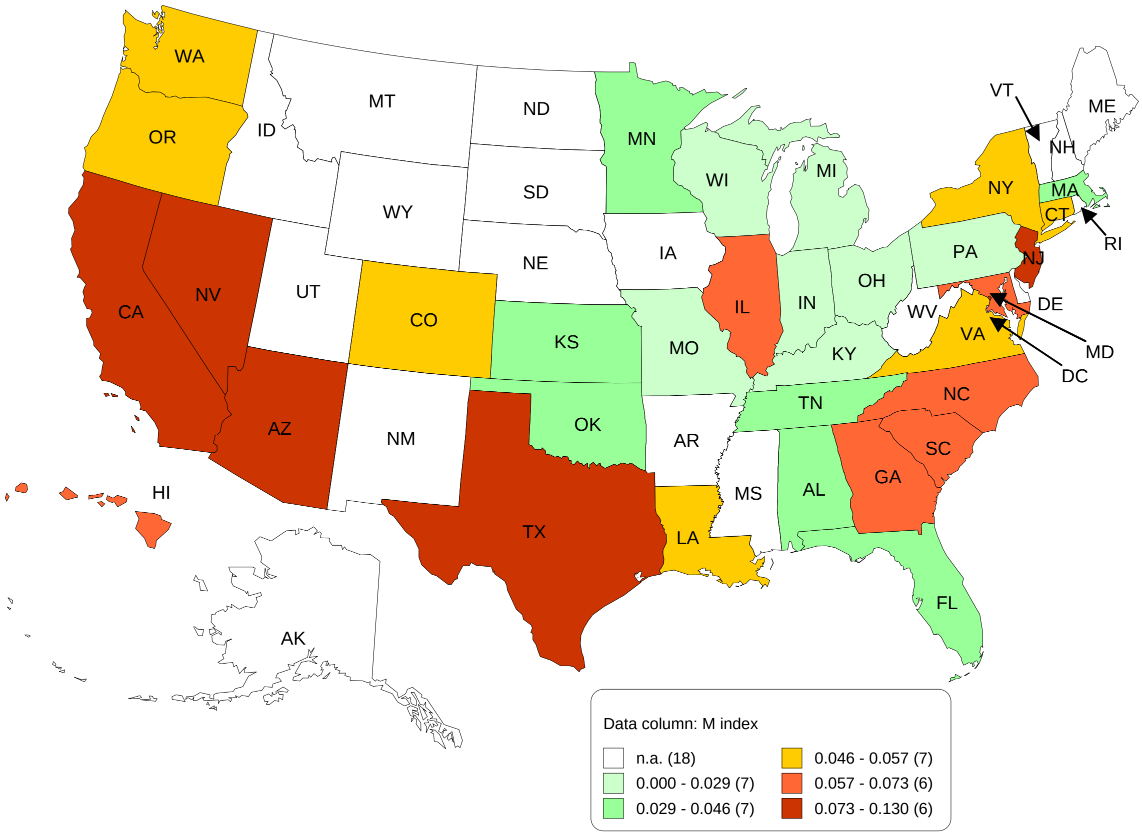
Regarding race and ethnicity, people are asked to choose the race or races with which they most closely identify and answer whether they have or not Spanish/Hispanic/Latino origin. Based on this self-reported identity, we produce six mutually exclusive groups of workers composed by the four major single race groups that do not have a Hispanic origin, plus Hispanics of any race, and others.[4] In other words, we consider whites; African Americans or blacks; Asians; American Indians, Alaskan, Hawaiian, and Pacific Islander natives (referred here for simplicity as Native Americans); Hispanics; and other races (those non-Hispanics reporting some other race or more than one race). Occupations are considered at a 2-digit level of the Standard Occupational Classification (SOC) System, which includes 52 occupations.[5] Multigroup segregation measurement requires focusing the analysis on 32 states (out of 50), together with the District of Columbia, with a significant sample for most demographic groups. Dropped states are mainly those with smaller and less demographically diverse populations, mostly included in the Central and Northwest areas of the country.[6] The final sample used in our analysis includes a total of 3,747,905 employed workers (from a minimum of 18,692 observations in Hawaii to a maximum of 467,119 in California; see Table A5 in the Appendix).

2.3 Segregation at state level

In this section, we are interested in measuring occupational segregation by race and ethnicity across states. We use a multigroup approach so that we can assess how uniform the

[4] In what follows, we omit the “non-Hispanic” origin of these groups for the sake of simplicity.
[5] We discard the use of the 3-digit level SOC because the analysis would be problematic in most states due to the relatively small number of observations for various demographic groups.
[6] The 18 states dropped are the following: Alaska, Arkansas, Delaware, Idaho, Iowa, Maine, Mississippi, Montana, Nebraska, New Hampshire, New Mexico, North Dakota, Rhode Island, South Dakota, Utah, Vermont, West Virginia, and Wyoming. These states represent 9% of workers in our survey.

distribution of this phenomenon is across US. Using the *M* index, Map 1 shows the segregation levels of states classified in five groups, each including six or seven states (the corresponding values for indexes *M*, *IP*, and *G* are given in the Appendix, columns 1 to 3 of Table A4). The map shows the existence of a great geographical variation in segregation across the US, the coefficient of variation of segregation being equal to 0.482. The highest level of segregation is found in the District of Columbia, which more than doubles the average segregation (0.052), distantly followed by several Southwestern states (such as California, Nevada, and Arizona) and Texas. In the East, only New Jersey joins the District of Columbia in this highly segregated group. The lowest segregation levels can be found in Northeastern states such as Ohio, Wisconsin, Missouri, Kentucky, Indiana, Michigan, and Pennsylvania (which barely reach half of the average segregation). The *G* and *IP* indexes produce a similar ranking of states, except that they also include Hawaii in the former group and Minnesota in the latter (their respective coefficients of variation are 0.407 and 0.408).



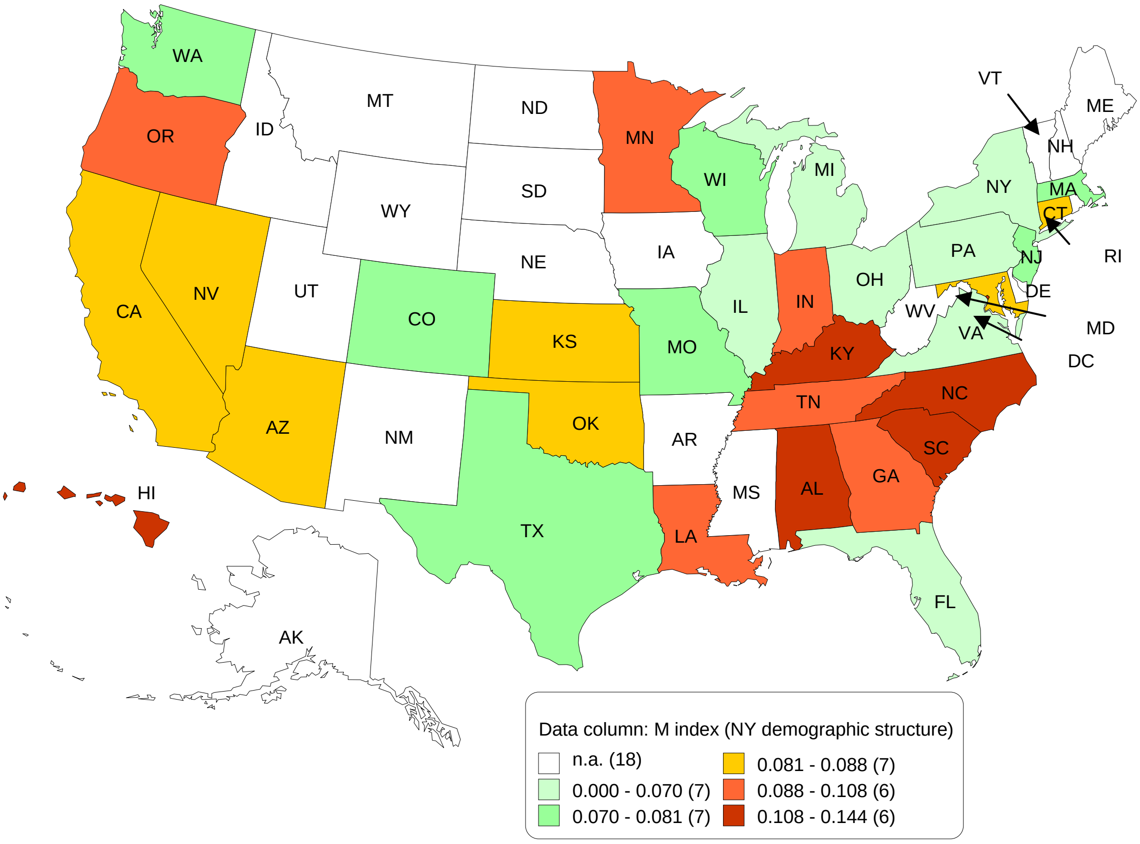
Map 1. Occupational segregation by race/ethnicity in selected states (*M* index).

Note: White states have not been assigned a value due to lack of data in the survey.

There seems to be a clear link between the level of segregation of a state and its demographic composition. Highly segregated states share a relatively low presence of whites; some states have a large proportion of Hispanics while others show remarkable racial diversity (see columns 1 to 6 of Table A1 in the Appendix). On the contrary, low-segregated states are more likely to have higher proportions of whites. This demographic pattern suggests that the greater the segregation in a state, the greater the degree of racial/ethnic diversity. This relationship is in line with Iceland's (2004) results, which found that residential segregation by race/ethnicity increased during the 1980s and 1990s as a result of the country becoming more multiethnic. But diversity alone does not explain segregation as states sharing similar demographic structures have different segregation levels. Indeed, on the one hand, Tennessee, Alabama, Louisiana, Georgia, North and South Carolina, Virginia, and Maryland do not experience a similar segregation level (segregation is remarkably lower in Tennessee, Alabama, and Virginia) despite the large proportion of African-Americans and low proportion of other minorities. On the other hand, Florida has much lower segregation than California, Arizona, Texas, and Nevada, notwithstanding the similar large presence of Hispanics.

To measure the importance of the geographical distribution of minorities in explaining spatial disparities in occupational segregation levels in the US, we can undertake a simple shift-share analysis. Thus, we recalculate segregation in every state keeping the original occupational structure of all racial/ethnic groups while assigning to each group the weight it has in a state of reference. We chose New York as a reference because it has a racial/ethnic composition close to that of the entire country. For this purpose, we take advantage of the decomposability property of the *M* index, according to which this measure can be written as the weighted sum of the segregation of racial/ethnic groups across occupations, with weights being their relative population sizes.[7] For example, the shift-share analysis for the *M* index in the case of Alabama is $M_{AL}^{\text{shift-share}} = \sum_{g} \frac{C_{NY}^{g}}{T_{NY}} M_{AL}^{g}$. This exercise is shown in Map 2 (and in Table A4 in the Appendix, columns 4–6).

[7] The segregation of a racial/ethnic group quantifies the discrepancy between its distribution across occupations and that of the state. In our shift-share analysis, we keep the segregation of each group unaltered, which implies that the original occupational structure of the state is taken as the benchmark distribution. Note that changing the weights of groups necessarily leads to a new occupational structure in each state, which could be considered an alternative benchmark.



Map 2. Occupational segregation by race/ethnicity in selected states using the demographic structure of New York (*M* index).

Note: Note: White states have not been assigned a value due to the small sample size for some demographic groups in the survey.

The analysis reveals that segregation dispersion across states is substantially reduced after controlling for disparities in demographic structures across states (the coefficient of variation, equal to 0.232, is more than halved).[8] In other words, by imposing a demographic structure common to all states, we eliminate an important source of variability in segregation. Further, the highest segregation remains in the South but shifts to the East. On the one hand, several states in the Southwest and Central areas that previously had relatively high segregation—such as California, Nevada, Arizona, and Texas—now show relatively lower levels. In fact, among the most segregated group of states in Map 1, only the District of Columbia remains in a similar position in Map 2.[9] On the other hand, segregation rises remarkably in Hawaii and Southeastern states such as Kentucky, Alabama, Tennessee, and both Carolinas.[10] A

[8] The coefficient of variation is even smaller with the other two indices: *IP* (0.186) and *G* (0.183).

[9] With the *IP* and *G* indexes, reductions in the level of segregation (and changes in the ranking) are of a lower magnitude compared with the *M* index.

[10] Hawaii has a very specific racial structure compared with other states (see Table A1 in the Appendix). Asians represent almost 40% of the population while whites are a minority (28%). In addtion, Hispanics and Asians in Hawaii experience the lowest segregation of the country, but African Americans and whites experience the

noteworthy exception to this geographical pattern is Florida, which is slightly affected by the change in demographic weights and now shows lower segregation than any other state (even though it is surrounded by states with the highest levels of segregation in the country). Other states in the Southeast improving their relative positions in terms of segregation are Virginia and Maryland. Minnesota and Oregon have experienced a relative worsening while New York, New Jersey, and Illinois have improved. On the contrary, Pennsylvania, Michigan, and Ohio do not vary, keeping their relatively low initial segregation levels.

As a conclusion of this shift-share analysis, the map of (unconditional) occupational segregation by race/ethnicity in the US is strongly driven by the heterogeneity of demographic structures of states. Conditioning on this single factor dramatically reduces the dispersion of segregation across the country while producing a substantial re-ordering of states according to their relative segregation levels. Segregation tends to rise in states with low numbers of Hispanics and Asians because these minorities generally face the highest segregation among all racial/ethnic groups. But the demographic structure may not be the only relevant factor in explaining segregation disparities. The literature has also pointed out that the risk of segregation may differ across population groups according to other worker characteristics such as education and immigration status (aggravated by short periods of residence in the US and lack of English proficiency).[11] Therefore, there is room for an analysis of segregation conditioned on the distribution of all these characteristics in a more general way. We do this in the next section.

3. Conditional segregation: Differences across states

In this section, we extend the previous shift-share analysis to assess more rigorously the importance of the spatial distribution of various characteristics, jointly considered, in explaining disparities in occupational segregation by race/ethnicity. In what follows, we present a propensity score methodology initially proposed by Di Nardo et al. (1996) for the decomposition of wage differentials and later adapted by Gradín (2010) to measure conditional occupational segregation of nonwhites versus whites in the US at the national level.

---

highest level. These facts explain why segregation increases in Hawaii using the racial structure of New York (the effect on segregation of the decrease of Asians and the increase of African Americans and whites dominates the effect of the larger number of Hispanics).

[11] See Hellerstein and Neumark (2008), Alonso-Villar et al. (2010), and Gradín (2010).

3.1 Measuring conditional segregation

Let us denote by $z \equiv (z_1,...,z_k)$ a vector of k covariates describing the main attributes of workers, which can be grouped in four factors: race/ethnicity, education, immigration profile (including English proficiency), and industry. By using z, we can partition workers in each state into several mutually exclusive subgroups or "cells," with each being a specific combination of attributes. An example would be Hispanic immigrants who have lived up to five years in the US, have a university degree, and work in the manufacturing sector. If we represent by $F(o, z|D = s)$ the joint distribution of occupations and attributes in state s (where D is the categorical variable representing state membership), its discrete density function across occupations can be written as follows:

$$f(o|D = s) = \int_z dF(o, z|D = s)dz = \int_z f(o|z, D = s) f(z|D = s)dz ,$$

where $f(o|z, D = s)$ is the distribution across occupations of individuals in s having attributes z, and $f(z|D = s)$ is the attribute density in state s.

If we assume that the distribution across occupations of individuals in each cell does not depend on the distribution of attributes (i.e., if $f(o|z, D = s)$ and $f(z|D = s)$ are independent), then we can define the counterfactual density function of state s, $\tilde{f}_s(o)$, as the density function across occupations that the state would have were it given the same distribution of attributes of a state of reference (in our case, New York: $f(z|D = \text{New York})$) while keeping unchanged the distribution of every subgroup across occupations (i.e., $f(o|z, D = s)$). In other words, the counterfactual distribution for state s

$$\tilde{f}_s(o) = \int_z f(o|z, D = s).f(z|D = \text{New York})dz$$

represents the occupational distribution that would prevail in that state if each subgroup of individuals (defined by the cross of the main characteristics defined above) kept their own conditional probability of being in a given occupation, but the state had the same characteristics of New York in terms of racial/ethnic composition, immigration profile, and educational and industrial structures. One could proxy $f(z|D = \text{New York})$ by the frequency distribution of attributes empirically observed in New York provided that all covariates in z

are dummies (as in our analysis). However, this process is cumbersome when many categories are involved, and it could be problematic if some cells are empty.[12] Furthermore, it would be difficult to separate the individual effects of each covariate on segregation. To overcome these problems, we follow Di Nardo et al. (1996) and re-formulate the counterfactual density,

$$\tilde{f}_s(o) = \int_z f(o|z, D = s).\Psi_z f(z|D = s)dz = \int_z \Psi_z f(o, z|D = s)dz \,,$$

such that it can be simply obtained from reweighting the original observations in the target state by $\Psi_z = \frac{f(z|D = \text{New York})}{f(z|D = s)}$. By using the Bayes's theorem, weights can be rewritten as the product of two ratios that can be easily estimated from the data:

$$\Psi_z = \frac{\dfrac{\Pr(D = \text{NewYork}|\mathrm{z})}{\Pr(D = \text{NewYork})}}{\dfrac{\Pr(D = s|\mathrm{z})}{\Pr(D = s)}} = \frac{\Pr(D = s)}{\Pr(D = \text{NewYork})} \frac{\Pr(D = \text{NewYork}|\mathrm{z})}{\Pr(D = s|\mathrm{z})} \,.$$

The first component can be directly approximated by the ratio between the population samples in both states. The second component can be obtained by estimating the probability of an individual with attributes *z* to belong to New York (rather than to its own state *s*) using a binary probability model. Thus, we estimate the following logit model,

$$\Pr(D = \text{New York}|z) = \frac{\exp(z\hat{\beta})}{1 + \exp(z\hat{\beta})} \,,$$

over the pool sample with observations from both states, where $\hat{\beta}$ is the associated vector of estimated coefficients.

The segregation level obtained in the counterfactual distribution reflects the amount of unexplained segregation in state *s* that remains after controlling for state characteristics.[13] After completing the same exercise for every state of interest, we can compare segregation across states in the US under a similar distribution of characteristics, and then its variability only reflects geographical differences in the conditional distributions across occupations.

[12] In our empirical analysis, we use a total of 34 categories (or dummies) to account for all four factors.

[13] Unlike the exercise in the previous section, when measuring the segregation of each racial/ethnic group, we now consider the occupational structure of the counterfactual distribution as the benchmark, rather than that of the original distribution. This implies that this method assumes that the occupational structure of each state is endogenous because it depends on the distribution of characteristics across the population.

Moreover, the difference between unconditional and conditional segregation provides a measure of the segregation that is actually explained by our covariates z.[14] This explained term can be additionally disaggregated into the detailed contribution of each factor (a subset of covariates) to identify which are more explicative (see Gradín, 2010). These contributions are obtained by using the Shapley decomposition (Chantreuil and Trannoy, 1999; Shorrocks, 1999).

Thus, to obtain the contribution of education, for example, we use the logit coefficients obtained above as follows. First, we calculate the prediction of $\Pr(D = \text{New York} \mid z)$ by assuming that all coefficients except those of education dummies are zero, and we then compare the conditional segregation resulting from this new counterfactual with the unconditional segregation of the state. It would appear that this is the contribution of education to the explained segregation; however, the (marginal) contribution of education would be different if we had previously modified the coefficients of one or more other factors (this is the well-known path-dependency problem found in income inequality decompositions). Furthermore, marginal contributions of all variables would not sum up the total effect. Therefore, the above contribution is actually the one corresponding to education when this is the first explaining factor that changes; thus, we need to consider the remaining cases. We must calculate the prediction of the aforementioned probability assuming that the coefficients of all covariates—except immigration and education covariates—are zero. The resulting counterfactuals are then compared to obtain the marginal contribution of education when immigration has been taken into account.

Similarly, the analysis should be repeated when race/ethnicity or industry, rather than immigration, is the first factor that changes. By following the same procedure, we have to consider all possible sequences where education is the third and forth factor to change. Averaging over all possible marginal contributions of education, we compute the contribution of this covariate to explained segregation.[15]

The main advantage of the Shapley decompositions, widely used in income distribution analysis, is that the contributions of covariates are path independent and sum up the overall explained segregation. Moreover, we can use the same technique to decompose the explained

[14] This is in line with the conventional wage gap decomposition in the explained and unexplained effects (characteristics and coefficients, respectively).

[15] See, e.g, Sastre and Trannoy (2002) for a formalization of the procedure to compute the Shapley decomposition.

level in any statistic computed over segregation measures across states such as the mean, the standard deviation, or the coefficient of variation.

3.2 Conditional segregation at state level

We first pool the sample of each target state and that of New York, and estimate a logit regression of the probability of a worker belonging to the latter. These auxiliary regressions are shown in the Appendix (Tables A5 and A6).[16] The distribution of characteristics across states is also reported in Tables A1 through A3 in the Appendix. By reweighting the original distribution using the predicted probabilities, we then construct the counterfactual density of each state as if it had the same distribution of characteristics as New York. We use this density function to measure conditional segregation, which will be compared with the unconditional segregation.

The change in segregation experienced by each state after conditioning on characteristics is reported in Figure 1. This figure shows the contribution to the overall change (estimated using Shapley decomposition) of the set of explanatory factors: race/ethnicity composition, educational level, immigration profile, and industry structure. Positive (negative) values indicate that segregation increases (decreases) after conditioning, meaning that the distribution of characteristics in the target state is less (more) segregating than that in New York. Obviously, segregation in the state of reference, New York, does not change at all by construction. The sum of contributions by the four factors (either positive or negative) represents the net overall change in segregation. The exact values of the conditional segregation and factor contributions can be found in Tables A4 (columns 7–9) and A7 in the Appendix, respectively.

[16] The explicative variable is a dummy that has a value of 1 if the worker belongs to the New York sample and 0 if she/he belongs to the target state. Explaining variables are an array of dummies accounting for four factors: race and ethnicity (six groups omitting whites); attained education (less than high school (omitted), high school diploma, some college, and bachelor or higher); immigration (born in the US (omitted), immigrant with up to 5 years of residence, between 6 and 10, between 11 and 15, or more than 15) and English proficiency (speaking only English (omitted), speaking English very well, well, not well, not at all); and industry (NAICS at one digit, 14 groups (omitting group 10)).

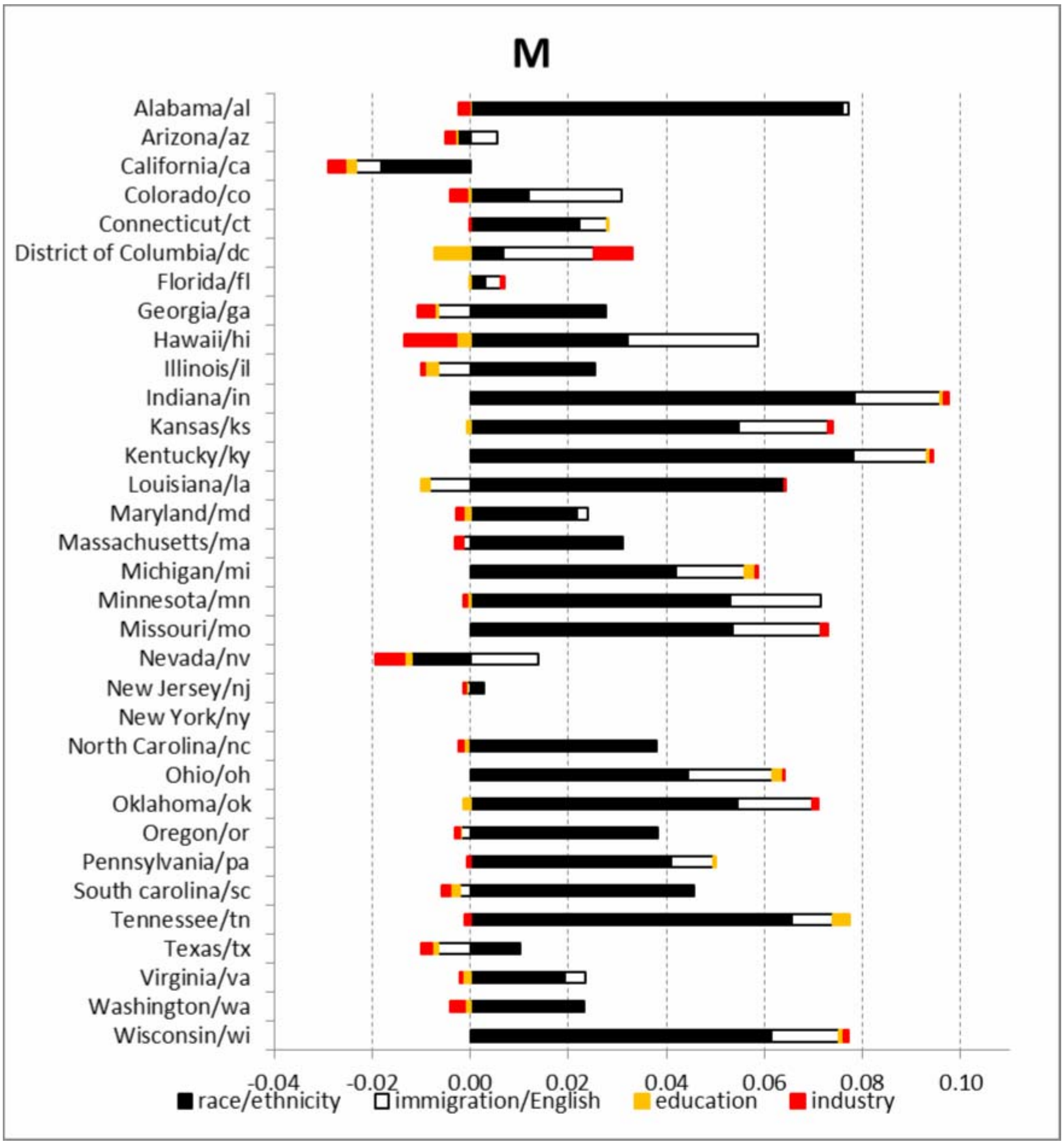


Figure 1. Conditional-unconditional segregation gap in selected states (*M* index): Factors' contributions using the Shapley decomposition.

We find that most states experience a net increase in segregation after conditioning on characteristics, the largest being in Alabama (from 0.042 to 0.117, with *M* index), Indiana (from 0.025 to 0.123), and Kentucky (from 0.023 to 0.117), indicating that their distributions of characteristics, compared with that of New York, partially offset the underlying level of segregation faced by their minorities in the labor market. The main exceptions are Western states such as California and Nevada, where the net effect is negative; their distributions of characteristics produce more segregation than that of New York. Other states such as Arizona, Texas, and New Jersey experience virtually no net change because positive and negative effects cancel each other. Similarly, Florida also shows a very small variation.

Table 1 reveals the impact that conditioning has on the mean and dispersion of segregation across states. On average, segregation increases by 73% with the *M* index (around 46–47% with the other two measures). However, the geographical dispersion of conditional segregation, measured by the coefficient of variation, is much lower than that in the unconditional case: it is reduced by 50% (*M*) or more (64% and 67% for *IP* and *G*, respectively). This means that at least half of the relative variability observed among the unconditional segregation levels of states can be explained by differences in the distributions of variables included in the conditional analysis.

	Unconditional segregation	**Conditional segregation (ref. New York)**									
M		**All**	**Δ%**	**Race/ ethnicity**	**Δ%**	**Immigration**	**Δ%**	**Education**	**Δ%**	**Industry**	**Δ%**
Mean	0.052	0.091	73.4	0.085	63.4	0.059	13.1	0.052	-1.2	0.051	-1.9
Standard deviation	0.025	0.022	-12.4	0.021	-14.8	0.026	4.8	0.024	-3.8	0.026	1.4
Coef. of Variation (St. dev / mean)	0.482	0.243	-49.5	0.264	-45.3	0.463	-3.9	0.471	-2.3	0.492	2.1
IP											
Mean	0.096	0.142	47.3	0.137	41.9	0.105	8.8	0.095	-1.7	0.095	-1.7
Standard deviation	0.039	0.021	-46.4	0.023	-42.7	0.038	-2.1	0.039	-2.1	0.040	0.5
Coef. of Variation (St. dev / mean)	0.408	0.149	-63.6	0.174	-57.5	0.382	-6.5	0.405	-0.7	0.413	1.2
Gini											
Mean	0.132	0.193	45.7	0.185	39.5	0.144	8.8	0.131	-1.1	0.130	-1.5
Standard deviation	0.054	0.026	-51.4	0.028	-48.5	0.053	-1.9	0.053	-1.7	0.054	0.7
Coef. of Variation (St. dev / mean)	0.407	0.136	-66.7	0.159	-60.9	0.383	-6.1	0.404	-0.9	0.412	1.2

Table 1. Summary of statistics for segregation indexes across states.

The racial/ethnic structure is clearly the most important factor. It explains an increase of 64% in the average segregation and a reduction of about 45% in geographical variation (coefficient of variation). This is not a surprise considering that it is consistent with the strong demographic effect described in the previous section. The main difference is that now we incorporate other factors as well, making the demographic effect cleaner (because we control for the potential correlation of race/ethnicity with the other relevant characteristics). Indeed, several states experience large increases in segregation mainly driven by the contribution of race/ethnicity, but now fueled by the immigration profile that stands out as the second most important factor (explaining 13% of the increase in the average level of segregation and a reduction of 4% in the coefficient of variation). This is the case of most of the states in the Southeast, such as Alabama, Kentucky, Louisiana, and Tennessee. After controlling for race/ethnicity, most more than double the initial level of segregation while immigration has an effect that varies between almost zero in the case of Alabama and a 50% increase in the case of Kentucky. A similar pattern is found in Oklahoma and in most of the Midwestern states (Indiana, Kansas, Minnesota, Missouri, Ohio, Wisconsin and, to a lesser extent, Michigan). In both regions, the listed states' demographic profiles differ from those of the rest of the country due to less immigration and smaller populations of Hispanics and Asians. Given that these groups are in general highly segregated, increasing their weight to make states share the same structure raises segregation.

The opposite occurs in those states with an overall negative effect on segregation after conditioning on characteristics. This is the case of California and Nevada. California is characterized by strong and recent immigration flows (35% of its workers are born outside the US and 12% speak English not well or not at all; in New York these figures are 29% and 6%, respectively). Moreover, this state has larger shares of Hispanic (33%) and Asian (13%) populations than the state of reference (15% and 7%, respectively). In the case of Nevada, where Hispanics also represent a large minority (22%), the share of immigrants is lower than in New York, which explains the positive impact of this factor.

Regarding the other explicative factors used in the analysis, none seems to be crucial to explain the segregation of any state (their impacts on the mean and dispersion of segregation are also small). Some facts are noteworthy, however. Education and industry do play significant roles in explaining segregation in the District of Columbia, even though they are of opposite sign, thus canceling each other. The effect of education on segregation could be the result of a higher level of education in this district, which explains why segregation decreases when controlling for this factor. Washington, D.C. has the largest relative concentration of workers with a university degree in the country (59% of the labor force versus 37% in New York, which is also one of the largest shares) and individuals with either high or low education tend to be more segregated at the national level than people with intermediate grades (Alonso-Villar et al., 2010). The District of Columbia also stands out for having the largest public administration (26% of the work force compared with around 5% in New York and in others states); it is expected that this industry has less segregation by race and ethnicity than does the private sector.[17] Therefore, segregation increases as we reduce its weight.

[17] Indeed, according to our calculations, occupational segregation by race and ethnicity in the public administration at the national level is half the level in the remaining sectors (0.018 compared to 0.043, *M* index).

With respect to the rest of the country, education is also of some relevance in states such as California and Tennessee. In California, controlling for education has the same effect as in the District of Columbia but for a different reason: The population with only primary education is around 50% higher than in New York. In Tennessee, there is an opposite impact of education because Tennessee has a higher population of intermediately educated workers than New York. Industry is also an important factor in Nevada and Hawaii, where high segregation appears to be partially connected to industrial structures. The former state places much weight on construction, nearly twice that of New York; the latter places an emphasis on active duty military (6% of the work force). Both share important entertainment-related activities, 24% and 14% of employment, respectively (compared with just 8% in New York).

As a result, we can derive conclusions about the overall effect of accounting for characteristics on the relative position of states in the unconditional segregation ranking. Figure 2 reports the segregation of states, relative to the average segregation, both before and after conditioning for the four factors. Because most states experience increments after conditioning, a state is expected to raise its level of segregation when it increases more than the average. It is unsurprising that this is the case of most states in the East Central region that have strong race/ethnicity and/or immigration effects. Similarly, the relative level of a state decreases when segregation either decreases or increases less than the average. The most significant reductions in relative segregation occur in some states on the east coast (New York, New Jersey, Georgia, Florida, Maryland, and the District of Columbia) and in most Southwestern states. Illinois also presents a remarkable reduction.

Map 3 shows the resulting geographical distribution of conditional segregation. It identifies the area with the highest segregation around the vertical line in the East Central region running from Indiana down to Alabama, passing through Kentucky and Tennessee. This is in addition to the particular cases of Hawaii and the District of Columbia. Moving to the center of the country, we find states with intermediate-high levels of segregation, both in the North (Minnesota and Wisconsin) and South (Kansas, Oklahoma, and Louisiana). A similar segregation level is found in South Carolina and North Carolina. The group with the lowest conditional segregation comprises states on the east coast (Florida, Virginia, New York, and Massachusetts), the Pacific coast (especially Washington and California), and Illinois.[18]

[18] This map is similar to Map 2, where only race/ethnicity was controlled through the shift-share analysis. The map is different in that California joins the group with the lowest segregation while Tennessee and Indiana join the group with the largest segregation levels.

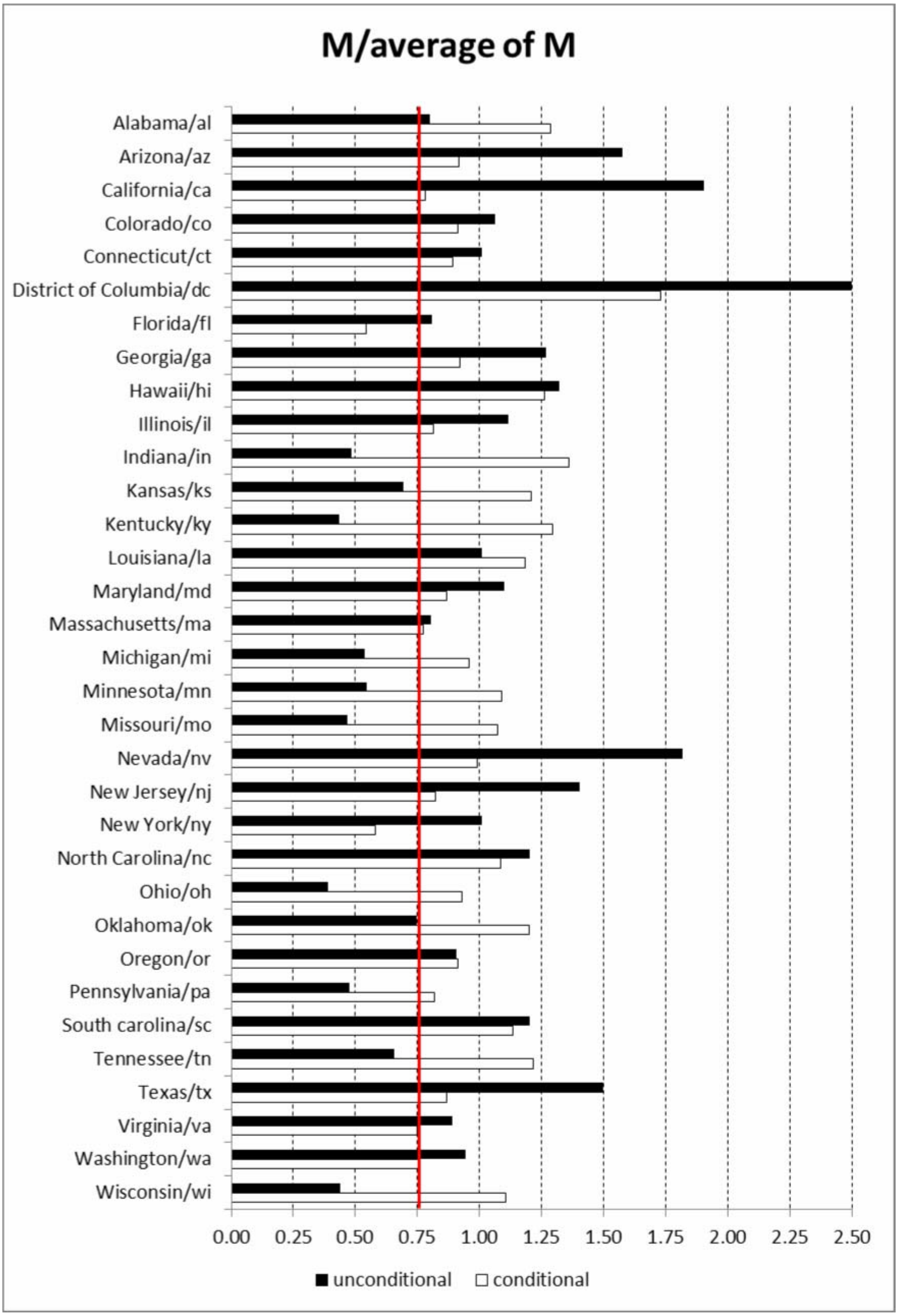
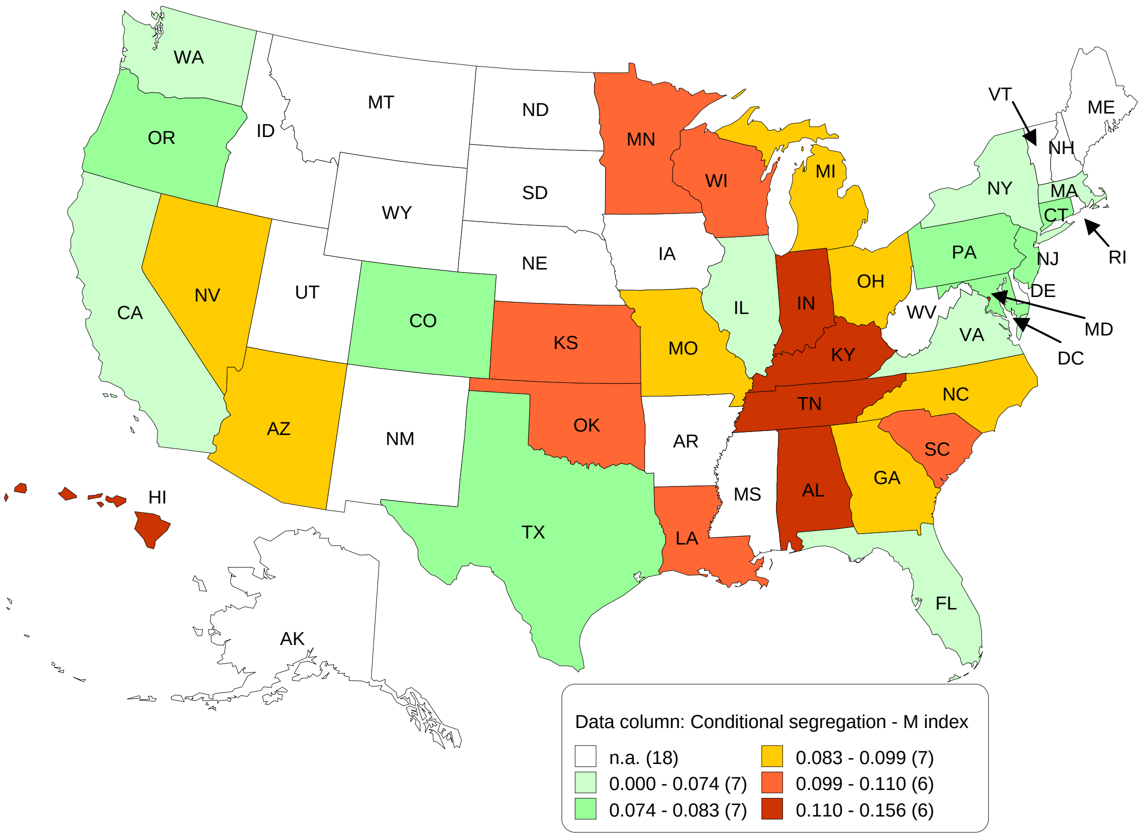


Figure 2. Conditional and unconditional occupational segregation in selected states (*M* index expressed relative to the average of states).



Map 3. Conditional occupational segregation by race/ethnicity in selected states (*M* index).

Note: White states have not been assigned a value due to the small sample size for some demographic groups in the survey.

To summarize, our analysis shows that the low (unconditional) segregation level found in some states is a consequence of their low racial diversity and immigration profile and not the result of a better integration of minorities in these states. This is the case in Indiana, Kentucky, Tennessee, and Alabama. On the contrary, states such as Florida, Pennsylvania, Massachusets, and Washington show low segregation levels even after controlling for their attributes, suggesting that minorities are more widely integrated into their labor markets. We also find that the high segregation level observed in the District of Columbia and Hawaii (and to a lesser extent, South and North Carolina) is not just the consequence of their demographic and industrial structures. Their higher levels of segregation could be the result of either unobservable characteristics or more segregative labor markets. Finally, a considerable part of the segregation found in California, Texas, New Jersey, and Illinois appears to be a consequence of their higher racial diversity and not the result of a lower integration of their racial/ethnic minorities.[19]

Conclusions

This paper has analyzed the extent of geographical disparities in occupational segregation by race and ethnicity across the United States. The unconditional analysis has resulted in great spatial discrepancies, with segregation being highly concentrated in the District of Columbia, New Jersey, Hawaii, and Southwestern states.

Because these disparities may arise from an uneven distribution of workers' characteristics across states, this paper has estimated conditional segregation by using a distribution of the relevant attributes of individuals which is similar across states. We initially undertook a simple shift-share analysis that controlled for the racial/ethnic structure of population, and later we used a propensity score technique that allows the consideration of various factors simultaneously—not only the racial/ethnic composition, but also the immigration profile, the educational achievement, and the industrial structure of each state. In addition, we determined the contribution of each factor to the differential between conditional and unconditional segregation.

The study has revealed that the geographical dispersion of segregation is significantly reduced after conditioning for the characteristics of states, where the racial/ethnic composition appears as the most relevant factor. The segregation map dramatically changes considering conditional segregation, with higher segregation moving toward the East. Thus, apart from the District of Columbia and Hawaii, which retained their high segregation levels, the highest conditional segregation was found in the East Central region, mostly in Alabama, Kentucky, Tennessee, and Indiana. Our analysis suggests that the low levels of unconditional segregation in this region arise from its low racial diversity rather than from a wider integration of minorities into their labor markets. On the contrary, Pennsylvania, Massachusets, and Florida

[19] The results we have highlighted do not depend on the index used because the *IP* and *G* indexes produce very similar results. We have also checked the robustness of these results by using California as the state of reference, which has a different distribution of demographic characteristics than the national average. The results summarized earlier still remained unchanged (except in the case of Hawaii). In fact, the Spearman rank correlation coefficient and the Pearson correlation coefficient between states in the New York and California benchmarks are 0.88 and 0.92, respectively, when using the *M* index. Discrepancies between both benchmarks are mainly due to the race/ethnic factor. In the case of Hawaii, when using California as the state of reference, the performance of this state improves substantially (with a change of 17 positions in the ranking with respect to New York). The remarkably low segregation of Hispanics in Hawaii (who compose 7% of workers), makes conditional segregation decrease notably when using California as the state of reference because in California, Hispanics compose 33% of the work force yet only 15% in New York.

have low segregation levels even when controlling for their attributes, indicating that this is not the result of a compositional effect.

These discrepancies among states show that even though the factors considered in this analysis explain a half or more of the segregation disparities across the states, a substantial portion of the segregation remains unexplained. Therefore, the probability of US minorities being confined to certain occupations is higher in some areas of the country. Understanding the reasons for that goes beyond the scope of this paper but could inspire future research.

References

Albelda, R. (1986): “Occupational segregation by race and gender, 1958-1981,” *Industrial and Labor Relations Review* 39(3), 404-411.

Alonso-Villar, O. and Del Río, C. (2010): “Local versus overall segregation measures,” *Mathematical Social Sciences* 60, 30-38.

Alonso-Villar, O., Del Río, C., and Gradín, C. (2010): “The extent of occupational segregation in the US: Differences by race, ethnicity, and gender,” ECINEQ WP2010-180, August.

Chantreuil, F. and Trannoy, A. (1999): “Inequality decomposition values: The trade-off between marginality and consistency,” THEMA Working Papers 99-24, Univesité de Cergy-Pontoise.

Clotfelter, C. (1999): “Public school segregation in metropolitan areas,” *Land Economics* 75, 487-504.

Cutler, D., Glaeser, E., and Vigdor, J. (1999): “The rise and decline of the American ghetto,” *Journal of Political Economy* 107, 455–506.

DiNardo, J., Fortin, N., and Lemieux, T. (1996): “Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach,” *Econometrica* 64, 1001-1044.

Duncan, O. and Duncan B. (1955): “A methodological analysis of segregation indexes,” *American Sociological Review* 20(2), 210-217.

Farley, R. (1977): “Residential segregation in urbanized areas of the United States in 1970: An analysis of social class and racial differences,” *Demography* 14(4), 497-518

Frankel, David M. and Oscar Volij. 2010. “Measuring School Segregation.” *Journal of Economic Theory*, forthcoming (doi: 10.1016/j.physletb.2003.10.071).

Gradín, C. (2010): “Conditional occupational segregation of minorities in the US,” ECINEQ WP2010-185, October.

Hellerstein, J. and Neumark D. (2008): “Workplace segregation in the United States: race, ethnicity, and skill,” *Review of Economics and Statistics* 90(3), 459-477.

Hellerstein, J., McInerney, M., and Neumark D. (2010): “Spatial mismatch, immigrant networks, and Hispanic employment in The United States,” The College of William and Mary, Department of Economics, Working Paper 100, July.

Hutchens, R. (2004): “One measure of segregation,” *International Economic Review* 45(2), 555-578.

Iceland, J. (2004): “Beyond black and white. Metropolitan residential segregation in multi-ethnic America,” *Social Science Research* 33, 248-271.

James, D. and Taeuber, K. (1985): “Measures of segregation,” *Sociological Methodology* 15, 1-32.

Karmel, T. and MacLachlan, M. (1988): “Occupational sex segregation—Increasing or decreasing?,” *The Economic Record* 64, 187-195.

King, M. 1992: “Occupational segregation by race and sex, 1940-88,” *Monthly Labor Review* 115, 30-37.

Massey, D. and Denton, N. (1987): “Trends in the residential segregation of blacks, Hispanics, and Asians: 1970-1980,” *American Sociological Review* 52(6), 802-825.

Massey, D. and Denton, N. (1988): “The dimensions of residential segregation,” *Social Forces* 67(2), 281-315.

Maxwell, N. (2010): “English language and low-skilled jobs: The structure of employment,” *Industrial Relations* 49(3), 457-465.

Patel, K. and Vella, F. (2007): “Immigrant networks and their implications for occupational choice and wages,” IZA Discussion Paper No. 3217.

Queneau, H. (2009): “Trends in occupational segregation by race and ethnicity in the USA: Evidence from detailed data,” *Applied Economics Letters* 16, 1347-1350.

Reardon, S. and Firebaugh, G. (2002): “Measures of multigroup segregation,” *Sociological Methodology* 32, 33-76.

Reardon, S.F. and O’Sullivan, D. (2004): “Measures of spatial segregation,” *Sociological Methodology* 34, 121-162.

Reardon, S. and Yun, J. (2001): “Suburban racial change and suburban school segregation, 1987-95.” *Sociology of Education* 74, 79-101.

Sastre, M. and Trannoy. A. (2002): “Shapley inequality decomposition by factor components: Some methodological issues,” *Journal of Economics* 9, 51-89.

Shorrocks, A. (1999): “Decomposition procedures for distributional analysis: A unified framework based on the Shapley value,” Mimeo, University of Essex.

Silber, J. (1992): “Occupational segregation indices in the multidimensional case: A note,” *The Economic Record* 68, 276-277.

Theil, H. and Finizza, A.J. (1971): “A note on the measurement of racial integration of schools by means of informational concepts,” *Journal of Mathematical Sociology* 1, 187-194.

Tomaskovic-Devey, D., Zimmer, C., Stainback, K., Robinson, C., Taylor, T., and McTague, T. (2006): “Documenting desegregation: Segregation in American workplaces by race, ethnicity, and sex, 1966-2003,” *American Sociological Review* 71, 565-588.

Watts, M. (1995): “Trends in Occupational segregation by race and gender in the U.S.A., 1983-92: A multidimensional approach,” *Review of Radical Political Economics* 27(4), 1-36.

Appendix

States	Race/ethnicity						Education			
	Whites	African Americans	Asians	Native Americans	Hispanics	Other Races	Less than High School	High School	Some College	Bachelor
Alabama/al	71.9	23.2	1.1	0.5	2.5	0.9	12.9	31.1	31.7	24.4
Alaska/ak	73.2	3.5	5.0	9.0	5.3	4.1	8.9	28.7	36.5	25.9
Arizona/az	63.4	3.0	2.5	3.0	27.0	1.1	14.5	25.3	34.3	25.9
Arkansas/ar	79.5	12.7	1.1	0.8	4.8	1.1	13.3	34.9	30.8	21.0
California/ca	46.8	5.4	12.6	0.7	32.7	1.8	15.9	22.8	30.7	30.7
Colorado/co	75.5	3.2	2.5	0.7	16.7	1.3	10.4	23.4	31.3	34.9
Connecticut/ct	76.7	8.2	3.5	0.2	10.0	1.4	8.7	27.9	27.6	35.9
Delaware/de	74.4	16.9	2.9	0.3	4.6	0.9	10.8	30.5	28.6	30.2
District of Columbia/dc	49.2	33.6	6.9	0.2	8.5	1.6	6.1	15.1	20.2	58.7
Florida/fl	61.5	13.8	2.3	0.3	20.9	1.2	12.2	29.8	31.5	26.5
Georgia/ga	62.0	26.6	2.8	0.2	7.3	1.0	12.8	29.5	28.8	28.9
Hawaii/hi	27.5	2.6	39.6	7.6	7.2	15.4	6.9	29.5	33.7	29.9
Idaho/id	87.8	0.4	1.0	1.2	8.4	1.2	11.6	28.3	35.9	24.2
Illinois/il	69.7	11.2	4.4	0.2	13.6	0.9	10.6	25.9	31.3	32.3
Indiana/in	86.5	7.0	1.3	0.2	4.2	0.9	11.2	34.3	30.8	23.7
Iowa/ia	92.7	1.9	1.5	0.2	3.2	0.5	9.2	31.0	34.4	25.4
Kansas/ks	83.5	5.2	2.0	0.7	7.3	1.3	9.7	27.6	34.2	28.5
Kentucky/ky	88.9	7.0	1.1	0.2	2.1	0.7	11.4	34.1	31.0	23.4
Louisiana/la	66.8	26.9	1.5	0.5	3.5	0.9	13.6	34.2	29.5	22.7
Maine/me	95.7	0.8	1.0	0.5	1.0	1.0	7.9	33.9	31.0	27.3
Maryland/md	64.3	23.5	4.9	0.3	5.8	1.2	9.9	26.9	27.9	35.2
Massachusetts/ma	82.1	4.9	4.6	0.2	6.6	1.7	8.5	25.4	25.9	40.2
Michigan/mi	81.9	10.6	2.5	0.4	3.5	1.1	8.5	28.3	35.3	28.0
Minnesota/mn	88.8	3.2	3.0	0.7	3.3	1.0	7.8	25.3	35.4	31.5
Mississippi/ms	63.5	32.3	1.0	0.4	2.3	0.5	14.3	31.5	32.6	21.6
Missouri/mo	85.2	9.1	1.5	0.4	2.8	1.0	10.3	30.4	31.7	27.6
Montana/mt	91.0	0.5	0.9	3.7	2.4	1.6	8.2	31.2	33.5	27.1
Nebraska/ne	88.0	3.2	1.5	0.5	5.9	0.9	9.2	27.7	36.0	27.1
Nevada/nv	62.3	6.2	6.4	1.3	22.1	1.8	15.2	30.8	32.7	21.3
New Hampshire/nh	93.8	0.9	1.9	0.3	2.3	0.7	9.2	29.8	30.4	30.6
New Jersey/nj	64.6	11.8	6.8	0.2	15.5	1.1	10.2	28.8	26.1	34.9
New Mexico/nm	46.9	1.9	1.5	6.9	41.7	1.1	13.7	27.6	32.8	26.0
New York/ny	63.6	12.9	7.2	0.3	14.7	1.4	10.4	26.1	27.0	36.6
North Carolina/nc	71.2	18.6	1.8	1.0	6.5	0.9	12.6	28.8	31.2	27.3
North Dakota/nd	92.3	0.7	0.7	3.9	1.8	0.7	7.7	26.8	38.6	26.9
Ohio/oh	85.9	9.3	1.6	0.2	2.1	0.9	9.1	33.7	30.9	26.3
Oklahoma/ok	75.2	6.4	1.9	5.8	6.6	4.1	12.1	31.1	32.6	24.2
Oregon/or	82.8	1.4	3.7	1.1	9.2	1.9	10.9	25.5	35.2	28.4
Pennsylvania/pa	85.4	7.9	2.5	0.1	3.4	0.7	8.9	34.9	26.9	29.4
Rhode Island/ri	83.8	3.7	2.6	0.3	8.1	1.6	11.2	26.9	29.9	32.0
South Carolina/sc	68.4	25.3	1.3	0.3	3.9	0.8	12.4	32.1	30.4	25.1
South Dakota/sd	91.3	0.8	0.7	4.4	2.0	0.9	9.8	31.6	33.2	25.4
Tennessee/tn	79.8	14.7	1.4	0.3	3.1	0.8	11.9	33.6	29.6	24.9
Texas/tx	52.7	10.3	3.5	0.4	32.2	1.0	16.9	26.5	30.4	26.3
Utah/ut	84.5	1.0	2.0	1.6	10.1	0.8	10.8	25.7	38.0	25.5
Vermont/vt	96.3	0.6	1.0	0.3	1.1	0.8	7.4	30.8	28.9	32.8
Virginia/va	68.7	19.2	4.5	0.3	6.2	1.2	10.7	27.6	28.8	32.9
Washington/wa	78.5	3.0	6.8	1.4	8.1	2.2	9.5	24.1	35.1	31.2
West Virginia/wv	94.7	2.9	0.8	0.1	0.9	0.6	9.8	39.3	29.1	21.8
Wisconsin/wi	89.1	3.9	1.6	0.7	3.9	0.7	9.1	31.2	33.0	26.6
Wyoming/wy	88.8	0.7	0.7	1.5	6.9	1.4	9.1	30.7	37.2	22.9

Table A1. Demographic structure and educational level (all rows in percentage).

States	Years of residence					English proficiency				
	Born in the US	0-5 years	6-10 years	11-15 years	16+ years	only English	very well	well	not well	not at all
Alabama/al	95.26	0.94	0.98	0.45	2.37	95.17	2.53	0.81	0.98	0.5
Alaska/ak	89.47	1.33	1.34	1.6	6.27	86.01	8.22	3.65	1.93	0.18
Arizona/az	79.82	3.61	3.96	2.93	9.68	72.51	14.5	4.59	4.98	3.42
Arkansas/ar	93.83	1.1	1.24	0.8	3.03	93.32	3.01	1.36	1.63	0.68
California/ca	64.52	3.52	5.14	4.59	22.23	58.32	21.25	8.9	7.69	3.83
Colorado/co	86.75	1.95	2.9	2	6.4	83.99	8.27	3.26	3.16	1.33
Connecticut/ct	81.5	2.41	3.55	2.77	9.77	80.51	11.38	4.51	2.82	0.78
Delaware/de	89.84	1.42	2.45	1.28	5.02	88.79	6.78	2.02	1.61	0.79
District of Columbia/dc	77.44	3.53	4	2.73	12.3	78.55	13.92	4.22	2.57	0.74
Florida/fl	73.41	3.72	5.13	3.81	13.93	72.71	14.58	5.73	4.64	2.34
Georgia/ga	86.61	2.54	3.21	1.98	5.66	87.11	6.1	2.76	2.77	1.27
Hawaii/hi	77.34	2.48	2.81	3	14.38	75.85	13.83	6.53	3.42	0.36
Idaho/id	92.22	0.99	1.26	1.12	4.4	89.72	5.76	1.94	1.91	0.67
Illinois/il	81.64	2.13	3.58	2.95	9.7	77.76	11.59	5.15	4.06	1.43
Indiana/in	94.43	0.94	1.32	0.8	2.51	92.7	4.17	1.4	1.33	0.39
Iowa/ia	95.06	0.82	1.17	0.71	2.24	93.84	3.2	1.41	1.19	0.36
Kansas/ks	92.31	1.25	1.65	0.93	3.86	90.71	4.94	1.81	1.75	0.79
Kentucky/ky	95.56	0.91	1.04	0.6	1.88	95.22	2.64	1.08	0.77	0.29
Louisiana/la	94.94	0.88	0.81	0.5	2.86	91.44	5.69	1.51	1.01	0.35
Maine/me	95.55	0.47	0.61	0.4	2.98	93.3	5.08	0.93	0.57	0.1
Maryland/md	84.39	2.59	3.08	2.07	7.87	85.24	8.6	3.27	2.18	0.72
Massachusetts/ma	81.22	2.71	3.68	2.6	9.78	80.48	11.05	4.57	2.87	1.03
Michigan/mi	92.28	1.07	1.66	1.17	3.83	91.16	5.51	1.93	1.09	0.3
Minnesota/mn	92.04	1.37	1.81	1.27	3.5	91.25	5.06	1.94	1.38	0.37
Mississippi/ms	96.48	0.85	0.62	0.33	1.72	95.73	2.39	0.65	0.81	0.41
Missouri/mo	95.07	0.81	0.98	0.76	2.38	94.13	3.63	1.14	0.83	0.27
Montana/mt	97.23	0.3	0.28	0.3	1.89	96.04	3.16	0.49	0.25	0.06
Nebraska/ne	92.91	1.22	1.36	1.16	3.35	91.81	3.72	1.75	1.78	0.94
Nevada/nv	74.89	3.66	4.29	3.72	13.45	73.01	12.79	6.4	5.51	2.28
New Hampshire/nh	92.78	0.96	1.28	0.88	4.11	91.72	5.69	1.35	1.17	0.07
New Jersey/nj	73.18	3.4	4.83	4.12	14.46	71.83	15.59	6.21	4.8	1.57
New Mexico/nm	87.28	1.91	2.02	1.53	7.25	65.07	25.14	4.46	3.45	1.88
New York/ny	71.3	2.98	4.61	4.55	16.57	71.95	15.54	6.37	4.64	1.49
North Carolina/nc	90.28	1.89	2.63	1.46	3.73	89.95	4.76	1.89	2.34	1.07
North Dakota/nd	96.7	0.44	0.61	0.3	1.95	95.16	3.62	0.84	0.3	0.08
Ohio/oh	95.11	0.74	1.02	0.59	2.54	93.93	4.02	1.19	0.71	0.14
Oklahoma/ok	92.35	1.3	1.66	1.07	3.63	91.2	4.26	2.03	1.87	0.64
Oregon/or	87.19	1.73	2.45	1.99	6.64	85.98	6.9	2.91	2.75	1.45
Pennsylvania/pa	92.61	1.09	1.32	1.01	3.97	91.28	5.36	1.84	1.17	0.36
Rhode island/ri	84.4	1.76	2.58	1.93	9.33	81.8	10.3	3.71	2.95	1.24
South Carolina/sc	93.14	1.51	1.54	0.81	3.02	93.13	3.39	1.32	1.51	0.64
South Dakota/sd	96.9	0.89	0.45	0.42	1.34	94.62	3.25	1.12	0.78	0.24
Tennessee/tn	94.02	1.21	1.29	0.86	2.62	93.88	3.19	1.26	1.23	0.44
Texas/tx	78.72	2.77	4.08	3.22	11.21	66.74	18.67	5.76	5.53	3.29
Utah/ut	88.78	1.75	2.59	1.71	5.17	85.35	8.31	2.78	2.61	0.95
Vermont/vt	95.66	0.57	0.58	0.53	2.66	95.29	3.44	0.76	0.43	0.08
Virginia/va	86.07	1.98	2.89	1.81	7.25	86.59	7.5	3	2.2	0.71
Washington/wa	83.99	2.14	3.11	2.35	8.42	83.73	8.47	3.92	2.73	1.14
West Virginia/wv	97.69	0.38	0.28	0.28	1.37	97.37	1.71	0.45	0.4	0.07
Wisconsin/wi	94.79	0.73	1.16	0.7	2.63	92.8	4.33	1.4	1.18	0.29
Wyoming/wy	95.89	0.71	0.66	0.23	2.51	93.75	4.3	0.94	0.78	0.23

Table A2. Immigration status: years of residence in the US and English proficiency (all rows in percentage.)

States	Agric. (1)	Construct. (2)	Manufact. (3)	Wholesale (4)	Retail (5)	Transport. (6)	Information (7)	Finance (8)	Profess. (9)	Education, Health (10)	Arts, Entertain. (11)	Other Services (12)	Public Admin. (13)	Active Duty Military (14)
Alabama/al	1.8	8.1	15.0	3.6	12.3	5.4	1.9	5.9	8.0	20.1	7.2	5.1	5.1	0.6
Alaska/ak	4.9	8.0	4.0	1.9	11.0	8.1	2.1	4.4	8.2	20.3	7.9	5.3	10.0	3.9
Arizona/az	1.3	10.7	7.9	3.1	12.0	4.8	1.8	8.4	11.0	18.4	9.8	4.7	5.3	0.7
Arkansas/ar	3.4	7.6	15.9	3.2	13.2	5.8	1.9	5.0	6.3	21.4	7.1	4.7	4.2	0.5
California/ca	1.9	7.9	10.6	3.7	11.1	4.6	3.0	7.4	11.8	18.7	8.9	5.2	4.4	0.8
Colorado/co	2.1	9.7	7.3	3.2	11.4	4.6	3.5	8.0	12.3	17.7	10.0	5.0	4.4	1.0
Connecticut/ct	0.4	6.7	12.5	3.0	11.4	3.7	2.6	9.6	10.3	23.1	7.9	4.5	3.7	0.5
Delaware/de	1.0	8.2	10.2	2.9	12.2	4.1	1.7	11.8	9.9	20.5	8.0	4.3	4.7	0.5
District of Columbia/dc	0.2	3.7	1.4	0.6	2.7	3.3	4.5	5.9	20.3	15.4	7.0	7.4	25.8	1.9
Florida/fl	1.1	10.3	5.9	3.5	12.8	5.0	2.4	8.5	11.4	18.4	10.3	5.1	4.7	0.7
Georgia/ga	1.2	8.7	11.5	3.6	11.8	6.1	2.8	7.0	10.1	18.4	7.8	4.8	5.1	1.2
Hawaii/hi	1.4	7.7	3.3	2.5	11.3	5.4	1.9	6.3	9.4	18.4	13.9	4.4	8.1	6.0
Idaho/id	5.4	10.1	10.4	2.9	12.2	4.4	2.2	5.9	8.9	19.6	8.1	4.2	5.1	0.7
Illinois/il	1.0	6.6	13.5	3.9	10.8	6.0	2.4	8.0	10.3	20.6	8.2	4.7	3.8	0.3
Indiana/in	1.3	6.8	20.7	3.2	11.3	5.2	1.9	5.5	7.0	20.8	8.2	4.7	3.4	0.1
Iowa/ia	4.0	6.7	15.6	3.5	11.9	4.9	2.2	7.3	6.2	23.2	7.3	4.3	3.0	0.1
Kansas/ks	3.7	6.3	14.0	3.4	11.3	5.2	3.1	6.1	7.7	22.3	7.3	4.3	4.6	0.8
Kentucky/ky	3.1	7.1	14.8	3.3	11.6	6.0	1.7	5.7	7.0	21.6	7.6	4.6	4.7	1.1
Louisiana/la	4.1	9.2	8.4	3.2	12.0	5.3	1.7	5.7	8.3	21.5	9.1	5.2	5.5	0.9
Maine/me	2.5	8.2	10.3	2.9	13.8	3.7	2.0	6.1	7.2	25.6	8.4	4.5	4.4	0.6
Maryland/md	0.6	8.1	6.1	2.7	11.1	4.5	2.5	7.2	13.3	21.9	7.4	5.1	8.7	0.9
Massachusetts/ma	0.4	6.7	10.6	3.2	10.8	3.8	2.9	8.1	12.3	25.1	7.9	4.4	3.9	0.2
Michigan/mi	1.1	5.9	19.1	3.1	11.4	4.1	1.9	5.9	8.6	22.0	8.8	4.6	3.5	0.1
Minnesota/mn	2.3	6.8	14.5	3.5	11.3	4.6	2.3	7.6	9.1	22.5	7.7	4.5	3.2	0.1
Mississippi/ms	2.7	7.6	15.4	3.1	11.7	4.8	1.7	5.1	6.0	22.2	8.8	4.7	5.3	1.1
Missouri/mo	1.8	7.2	12.1	3.2	11.8	5.5	2.4	7.4	8.5	21.5	8.6	4.9	4.5	0.6
Montana/mt	7.4	9.4	5.0	2.8	12.7	4.9	2.0	5.9	7.2	21.5	10.1	4.9	5.7	0.6
Nebraska/ne	5.0	6.7	11.0	3.5	11.4	6.6	2.0	7.7	7.6	21.6	7.8	4.4	4.0	0.7
Nevada/nv	1.5	11.3	4.6	2.8	10.5	4.9	1.7	7.1	10.0	13.2	24.0	3.9	4.2	0.6
New Hampshire/nh	0.8	7.8	12.7	3.5	14.9	3.6	2.3	6.7	8.8	22.5	8.3	4.5	3.5	0.1
New Jersey/nj	0.4	6.4	10.4	4.1	11.5	6.1	3.0	8.3	11.2	21.7	7.6	4.5	4.7	0.2
New Mexico/nm	3.9	9.0	5.4	2.5	11.7	4.4	1.9	5.3	10.7	22.4	10.3	4.4	7.3	0.9
New York/ny	0.6	6.1	7.6	3.1	10.5	5.4	3.5	9.2	10.9	25.2	8.2	4.8	4.7	0.3
North Carolina/nc	1.5	8.8	14.0	3.2	11.4	4.4	2.1	6.6	8.7	21.2	8.0	4.4	4.1	1.7
North Dakota/nd	8.0	6.8	8.4	3.4	12.0	5.1	2.1	5.6	6.1	24.3	8.4	3.9	4.6	1.4

Table A3. Industrial structure (all rows in percentage).

Ohio/oh	1.1	6.0	16.9	3.4	11.5	5.0	2.0	6.8	8.5	22.0	8.4	4.4	3.8	0.2
Oklahoma/ok	4.4	7.1	10.2	3.5	11.4	5.2	2.4	6.1	7.7	21.2	8.7	5.1	5.8	1.1
Oregon/or	3.4	7.6	12.6	3.6	12.5	4.5	2.1	6.5	9.8	19.5	8.8	4.5	4.4	0.1
Pennsylvania/pa	1.3	6.4	13.2	3.3	11.8	5.2	2.1	6.6	9.4	24.1	7.8	4.7	4.0	0.1
Rhode island/ri	0.4	7.0	11.7	2.6	11.2	3.6	2.1	8.0	8.7	25.3	9.5	4.8	4.5	0.6
South Carolina/sc	1.0	8.8	14.9	3.1	11.8	4.8	1.8	5.9	8.3	19.5	9.0	4.8	4.8	1.7
South Dakota/sd	7.1	6.4	10.2	3.1	11.3	4.2	2.0	8.8	6.4	21.9	8.6	4.8	4.9	0.6
Tennessee/tn	1.1	7.7	15.4	3.6	12.0	6.7	2.1	6.2	8.3	19.8	8.1	5.0	3.9	0.3
Texas/tx	2.8	9.1	10.0	3.6	11.6	5.6	2.3	6.9	10.1	19.8	8.1	5.3	4.2	0.8
Utah/ut	1.8	8.8	11.0	3.2	12.5	5.0	2.6	7.0	10.2	19.8	8.2	4.3	5.4	0.3
Vermont/vt	2.3	8.2	12.3	2.8	11.6	3.2	2.3	5.0	7.2	26.0	9.2	4.6	5.1	0.2
Virginia/va	1.2	8.2	8.7	2.4	11.1	4.3	2.7	6.7	13.1	18.9	7.5	4.9	7.5	2.8
Washington/wa	2.5	7.7	11.0	3.4	11.3	4.8	2.9	6.3	10.7	19.7	8.5	4.5	5.1	1.5
West Virginia/wv	4.7	7.6	9.2	2.6	12.5	5.6	1.7	4.8	7.4	24.4	8.9	4.5	6.1	0.1
Wisconsin/wi	2.6	6.4	19.1	3.3	11.7	4.5	2.0	6.3	7.3	21.4	8.2	4.0	3.2	0.1
Wyoming/wy	11.7	9.0	4.6	2.5	11.5	6.0	1.5	4.3	6.4	21.6	9.6	4.4	6.0	1.0

Table A3 (Cont.). Industrial structure (all rows in percentage).

Industry codes (NAICS):

1 Agriculture, Forestry, Fishing and Hunting, and Mining
2 Construction
3 Manufacturing
4 Wholesale Trade
5 Retail Trade
6 Transportation and Warehousing, and Utilities
7 Information
8 Finance and Insurance, and Real Estate and Rental and Leasing
9 Professional, Scientific, and Management, and Administrative and Waste Management services
10 Educational Services, and Health Care and Social Assistance
11 Arts, Entertainment, and Recreation, and Accommodation and Food Services
12 Other Services (except Public Administration)
13 Public Administration
14 Active Duty Military

States	M	IP	Gini	M shift-share*	IP shift-share*	Gini shift-share*	M conditional**	IP conditional**	Gini conditional**
Alabama	0.042	0.091	0.126	0.109	0.138	0.185	0.117	0.169	0.228
Arizona	0.082	0.146	0.200	0.083	0.143	0.197	0.083	0.129	0.179
California	0.099	0.166	0.236	0.081	0.150	0.213	0.071	0.127	0.176
Colorado	0.056	0.106	0.143	0.070	0.123	0.168	0.083	0.137	0.187
Connecticut	0.053	0.093	0.132	0.081	0.120	0.169	0.081	0.134	0.188
District of Columbia	0.130	0.194	0.260	0.144	0.197	0.264	0.156	0.211	0.282
Florida	0.042	0.098	0.139	0.049	0.102	0.145	0.049	0.101	0.144
Georgia	0.066	0.122	0.164	0.101	0.143	0.191	0.083	0.139	0.183
Hawaii	0.069	0.133	0.184	0.118	0.173	0.239	0.114	0.151	0.203
Illinois	0.058	0.107	0.148	0.069	0.117	0.162	0.074	0.122	0.169
Indiana	0.025	0.048	0.064	0.088	0.107	0.141	0.123	0.167	0.218
Kansas	0.036	0.070	0.094	0.082	0.117	0.155	0.110	0.162	0.215
Kentucky	0.023	0.040	0.055	0.108	0.109	0.148	0.117	0.159	0.217
Louisiana	0.053	0.107	0.151	0.092	0.132	0.181	0.107	0.161	0.218
Maryland	0.057	0.105	0.148	0.085	0.122	0.169	0.079	0.127	0.176
Massachusetts	0.042	0.077	0.104	0.078	0.115	0.157	0.070	0.126	0.175
Michigan	0.028	0.058	0.079	0.068	0.098	0.132	0.087	0.134	0.179
Minnesota	0.029	0.051	0.069	0.091	0.114	0.155	0.099	0.156	0.212
Missouri	0.024	0.050	0.070	0.072	0.099	0.136	0.097	0.147	0.198
Nevada	0.095	0.157	0.210	0.085	0.148	0.199	0.090	0.136	0.188
New Jersey	0.073	0.127	0.176	0.074	0.128	0.177	0.075	0.128	0.177
New York	0.053	0.110	0.156	0.053	0.110	0.156	0.053	0.110	0.156
North Carolina	0.063	0.113	0.149	0.110	0.146	0.190	0.098	0.151	0.199
Ohio	0.020	0.044	0.062	0.065	0.091	0.125	0.084	0.135	0.183
Oklahoma	0.039	0.080	0.109	0.082	0.120	0.162	0.109	0.163	0.220
Oregon	0.047	0.076	0.102	0.088	0.117	0.157	0.083	0.128	0.177
Pennsylvania	0.025	0.050	0.071	0.065	0.096	0.132	0.074	0.131	0.176
South Carolina	0.063	0.123	0.166	0.114	0.155	0.207	0.103	0.155	0.208
Tennessee	0.034	0.068	0.093	0.104	0.122	0.164	0.110	0.161	0.218
Texas	0.078	0.150	0.208	0.076	0.143	0.199	0.078	0.130	0.179
Virginia	0.046	0.095	0.132	0.070	0.114	0.157	0.068	0.123	0.168
Washington	0.049	0.080	0.110	0.081	0.110	0.151	0.068	0.124	0.169
Wisconsin	0.023	0.044	0.059	0.078	0.101	0.139	0.100	0.150	0.200

Table A4. Segregation indexes.
* Shift-share analysis, reweighting each index by New York's race/ethnicity distribution
** Conditional analysis reweighting observations in each state using New York's distribution by race/ethnicity, education, immigration profile, and industry.

State	N. observations	Wald Chi2 (29)	p-value	Pseudo R2
Alabama	59,611	12,629	0	0.121
Arizona	79,257	13,510	0	0.079
California	467,119	33,779	0	0.066
Colorado	71,691	9,968	0	0.064
Connecticut	51,063	3,941	0	0.022
District of Columbia	21,061	13,701	0	0.168
Florida	247,129	12,481	0	0.031
Georgia	126,423	17,046	0	0.074
Hawaii	18,692	19,588	0	0.343
Illinois	177,495	9,127	0	0.025
Indiana	91,308	21,725	0	0.120
Kansas	42,316	8,804	0	0.079
Kentucky	56,894	14,674	0	0.124
Louisiana	55,118	13,390	0	0.122
Maryland	77,017	8,981	0	0.048
Massachusetts	97,192	7,537	0	0.036
Michigan	132,947	22,151	0	0.097
Minnesota	81,009	11,523	0	0.096
Missouri	84,878	13,227	0	0.098
Nevada	36,118	10,501	0	0.073
New Jersey	115,019	1,740	0	0.006
North Carolina	264,206	18,766	0	0.085
Ohio	124,505	27,948	0	0.120
Oklahoma	165,376	11,893	0	0.112
Oregon	46,749	10,106	0	0.079
Pennsylvania	52,467	20,753	0	0.089
South Carolina	172,583	13,990	0	0.110
Tennessee	58,174	14,949	0	0.104
Texas	84,388	28,164	0	0.073
Virginia	303,928	11,155	0	0.055
Washington	110,981	11,506	0	0.064
Wisconsin	88,835	17,472	0	0.122

Tabla A.5 Logit Regressions for the probability of belonging to each state vs. New York (summary).

	Race/ethnicity					Years of residence				English				Education		
State	**Afr. Am.**	**Asians**	**Nat. Am.**	**Hisp.**	**Other**	**0-5 years**	**6-10 years**	**11-15 years**	**16+ years**	**Very well**	**well**	**Not well**	**Not At all**	**High School**	**Some College**	**Bachelor**
al	-0.73	0.44	-0.56	0.89	0.17	0.99	1.41	2.14	1.83	1.03	0.75	0.23	-0.23	0.40	0.40	0.78
	0.02	0.06	0.09	0.04	0.06	0.07	0.07	0.08	0.04	0.04	0.06	0.07	0.10	0.02	0.02	0.02
az	1.27	0.28	-2.59	-1.09	0.03	0.37	0.53	0.71	0.72	0.42	0.49	0.38	-0.18	0.28	-0.05	0.39
	0.03	0.03	0.06	0.02	0.05	0.04	0.04	0.04	0.02	0.02	0.03	0.04	0.05	0.02	0.02	0.02
ca	0.45	-1.20	-1.47	-1.30	-0.71	0.66	0.60	0.67	0.30	0.02	0.00	-0.05	-0.36	0.19	-0.21	0.00
	0.01	0.01	0.06	0.01	0.03	0.02	0.02	0.02	0.01	0.01	0.02	0.02	0.03	0.01	0.01	0.01
co	1.30	0.20	-0.92	-0.74	-0.08	0.56	0.49	0.77	0.85	0.80	0.66	0.59	0.46	0.22	-0.05	0.04
	0.03	0.04	0.09	0.02	0.05	0.05	0.04	0.05	0.03	0.03	0.04	0.04	0.07	0.02	0.02	0.02
ct	0.52	0.66	0.48	0.36	0.04	-0.05	0.04	0.29	0.35	0.09	0.05	0.26	0.45	-0.07	-0.01	0.03
	0.02	0.04	0.15	0.03	0.06	0.05	0.04	0.04	0.02	0.02	0.04	0.05	0.08	0.02	0.02	0.02
dc	-1.59	-0.49	-0.49	-0.21	-0.63	-0.04	0.27	0.62	0.56	-0.02	-0.08	-0.11	-0.09	0.16	0.01	-0.94
	0.02	0.04	0.20	0.04	0.08	0.06	0.06	0.06	0.03	0.03	0.06	0.08	0.14	0.05	0.05	0.04
fl	-0.16	0.83	-0.16	-0.53	0.08	-0.13	-0.05	0.18	0.17	0.15	0.24	0.34	0.04	-0.05	-0.11	0.23
	0.01	0.02	0.07	0.01	0.03	0.03	0.02	0.02	0.01	0.01	0.02	0.02	0.04	0.01	0.01	0.01
ga	-0.95	0.03	0.07	0.16	-0.01	0.25	0.43	0.88	1.14	0.54	0.30	0.13	-0.07	0.29	0.32	0.41
	0.01	0.03	0.10	0.02	0.05	0.04	0.03	0.04	0.02	0.02	0.03	0.04	0.05	0.02	0.02	0.02
hi	0.59	-4.18	-4.73	-1.25	-3.72	1.42	1.79	1.74	1.38	0.90	0.66	0.81	2.00	-0.41	-0.47	-0.04
	0.08	0.04	0.09	0.05	0.04	0.08	0.07	0.07	0.04	0.04	0.05	0.07	0.17	0.05	0.05	0.05
il	0.07	0.09	0.43	-0.09	0.35	0.69	0.60	0.76	0.80	0.01	-0.25	-0.26	-0.30	0.10	-0.09	0.10
	0.01	0.02	0.10	0.02	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.03	0.04	0.02	0.02	0.02
in	0.67	0.69	0.64	0.72	0.33	0.74	0.86	1.34	1.53	0.68	0.54	0.39	0.59	0.21	0.34	0.83
	0.02	0.04	0.11	0.03	0.05	0.06	0.05	0.06	0.03	0.03	0.04	0.05	0.09	0.02	0.02	0.02
ks	0.93	0.28	-0.79	0.09	-0.01	0.66	0.77	1.30	1.17	0.84	0.63	0.46	0.26	0.22	0.05	0.47
	0.04	0.06	0.09	0.04	0.06	0.09	0.07	0.08	0.04	0.04	0.06	0.07	0.10	0.03	0.03	0.03
ky	0.74	0.79	0.56	1.32	0.65	0.37	0.73	1.29	1.56	1.02	0.71	0.80	0.75	0.29	0.43	0.94
	0.03	0.06	0.14	0.04	0.07	0.07	0.06	0.08	0.04	0.03	0.06	0.08	0.12	0.02	0.02	0.02
la	-0.95	0.16	-0.45	0.75	0.05	1.42	1.99	2.45	1.99	0.20	-0.02	-0.02	-0.14	0.27	0.47	0.83
	0.02	0.05	0.10	0.04	0.06	0.08	0.07	0.09	0.04	0.03	0.06	0.07	0.12	0.02	0.02	0.02
md	-0.76	-0.30	-0.07	0.48	-0.11	0.11	0.43	0.82	0.82	0.28	0.19	0.27	0.31	0.15	0.19	0.16
	0.02	0.03	0.12	0.03	0.05	0.04	0.04	0.04	0.02	0.02	0.03	0.04	0.07	0.02	0.02	0.02
ma	1.16	0.66	0.53	1.01	0.01	-0.19	-0.08	0.25	0.21	0.00	-0.09	0.03	-0.04	0.03	0.07	0.00
	0.02	0.03	0.11	0.02	0.04	0.04	0.03	0.04	0.02	0.02	0.03	0.04	0.06	0.02	0.02	0.02
mi	0.24	0.19	-0.30	0.96	0.21	0.61	0.68	1.05	1.22	0.44	0.34	0.60	0.79	0.09	-0.09	0.37
	0.02	0.03	0.07	0.03	0.04	0.05	0.04	0.04	0.02	0.02	0.04	0.05	0.09	0.02	0.02	0.02
mn	1.47	0.24	-0.77	1.08	0.42	0.12	0.30	0.68	1.03	0.64	0.52	0.56	0.83	0.16	-0.12	0.28
	0.04	0.04	0.09	0.04	0.07	0.06	0.05	0.06	0.03	0.03	0.05	0.06	0.11	0.02	0.02	0.02
mo	0.40	0.36	-0.16	0.98	0.29	0.76	1.06	1.32	1.55	0.80	0.67	0.69	0.73	0.25	0.28	0.65
	0.02	0.04	0.09	0.03	0.05	0.06	0.05	0.06	0.03	0.03	0.05	0.06	0.12	0.02	0.02	0.02
nv	0.62	-0.45	-1.59	-0.72	-0.37	0.18	0.39	0.44	0.33	0.47	0.41	0.48	0.38	0.10	0.02	0.60
	0.03	0.03	0.08	0.02	0.05	0.05	0.04	0.04	0.03	0.03	0.04	0.04	0.06	0.02	0.02	0.03
nj	0.04	-0.01	0.54	-0.07	0.15	-0.02	0.07	0.21	0.23	-0.08	-0.06	-0.06	-0.02	-0.11	0.01	0.00
	0.02	0.02	0.11	0.02	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.03	0.05	0.02	0.02	0.02
nc	-0.45	0.40	-1.23	0.23	0.21	0.43	0.48	1.03	1.40	0.75	0.54	0.22	0.04	0.36	0.29	0.56
	0.01	0.03	0.07	0.02	0.05	0.04	0.03	0.04	0.02	0.02	0.03	0.04	0.06	0.02	0.02	0.02
oh	0.39	0.38	0.66	1.33	0.40	0.73	0.95	1.54	1.48	0.58	0.56	0.73	1.21	0.05	0.18	0.60
	0.02	0.03	0.10	0.03	0.04	0.05	0.04	0.05	0.02	0.02	0.04	0.05	0.10	0.02	0.02	0.02
ok	0.65	0.17	-2.92	0.07	-1.19	0.70	0.86	1.27	1.32	0.99	0.49	0.39	0.49	0.29	0.29	0.79
	0.03	0.06	0.06	0.03	0.04	0.08	0.06	0.08	0.04	0.04	0.06	0.06	0.11	0.03	0.03	0.03
or	2.30	0.13	-1.17	0.20	-0.25	0.44	0.40	0.55	0.59	0.66	0.55	0.41	0.10	0.28	-0.02	0.46
	0.05	0.04	0.08	0.03	0.05	0.06	0.05	0.05	0.03	0.03	0.04	0.05	0.07	0.02	0.02	0.02
pa	0.58	0.23	1.06	1.04	0.61	0.59	0.87	1.16	1.13	0.45	0.38	0.49	0.53	-0.05	0.26	0.47
	0.02	0.03	0.11	0.03	0.05	0.04	0.04	0.04	0.02	0.02	0.03	0.04	0.08	0.02	0.02	0.02
sc	-0.83	0.45	-0.11	0.57	0.29	0.61	1.05	1.66	1.66	0.87	0.51	0.14	-0.03	0.28	0.34	0.62
	0.02	0.05	0.11	0.03	0.07	0.06	0.05	0.07	0.04	0.03	0.06	0.06	0.08	0.02	0.02	0.02
tn	-0.12	0.52	0.16	0.87	0.47	0.51	0.93	1.32	1.56	0.90	0.62	0.42	0.43	0.27	0.42	0.77
	0.02	0.05	0.10	0.03	0.06	0.05	0.05	0.06	0.03	0.03	0.05	0.05	0.09	0.02	0.02	0.02
tx	-0.17	-0.42	-0.69	-1.46	-0.14	1.05	1.01	1.17	1.12	0.10	0.15	0.09	-0.36	0.32	0.13	0.38
	0.01	0.02	0.07	0.01	0.03	0.03	0.02	0.02	0.01	0.01	0.02	0.02	0.03	0.01	0.01	0.01
va	-0.48	-0.26	-0.13	0.40	0.01	0.32	0.40	0.86	0.80	0.43	0.31	0.33	0.41	0.23	0.23	0.30
	0.01	0.03	0.09	0.02	0.04	0.04	0.03	0.04	0.02	0.02	0.03	0.04	0.06	0.02	0.02	0.02
wa	1.53	-0.40	-1.51	0.35	-0.43	0.16	0.19	0.46	0.47	0.54	0.39	0.49	0.41	0.21	-0.14	0.27
	0.03	0.03	0.07	0.02	0.04	0.04	0.04	0.04	0.02	0.02	0.03	0.04	0.06	0.02	0.02	0.02
wi	1.25	0.44	-0.65	0.75	0.60	0.93	0.90	1.41	1.41	0.68	0.54	0.45	0.75	0.14	0.11	0.57
	0.03	0.05	0.08	0.03	0.07	0.07	0.06	0.07	0.04	0.03	0.05	0.06	0.12	0.02	0.02	0.02

Tabla A.6 Logit Regressions for the probability of belonging to each state vs. New York: estimated coefficients (standard errors below).

	Industry (code)													
State	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**Intercept**
al	-1.08	-0.52	-0.86	-0.43	-0.31	-0.18	0.41	0.11	0.00	-0.09	-0.38	-0.14	-1.00	1.03
	0.05	0.03	0.02	0.03	0.02	0.03	0.04	0.03	0.02	0.03	0.03	0.03	0.10	0.02
az	-0.76	-0.68	-0.25	-0.21	-0.32	-0.19	0.38	-0.21	-0.28	-0.32	-0.21	-0.29	-1.09	1.17
	0.05	0.02	0.02	0.03	0.02	0.03	0.04	0.02	0.02	0.02	0.03	0.03	0.09	0.02
ca	-1.16	-0.47	-0.50	-0.37	-0.29	-0.10	-0.19	-0.07	-0.36	-0.25	-0.27	-0.22	-1.57	-0.06
	0.03	0.01	0.01	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.07	0.01
co	-1.45	-0.82	-0.27	-0.35	-0.40	-0.30	-0.26	-0.19	-0.43	-0.54	-0.46	-0.18	-1.59	1.34
	0.04	0.02	0.02	0.03	0.02	0.03	0.03	0.02	0.02	0.02	0.03	0.03	0.08	0.02
ct	0.56	-0.16	-0.56	-0.05	-0.15	0.24	0.24	-0.11	-0.01	-0.06	-0.06	0.21	-0.64	1.54
	0.09	0.03	0.02	0.03	0.02	0.03	0.04	0.02	0.02	0.03	0.03	0.03	0.10	0.03
dc	-0.20	-0.62	0.71	0.75	0.48	-0.29	-0.83	-0.16	-1.22	-0.75	-1.25	-2.27	-3.04	3.99
	0.38	0.06	0.08	0.11	0.06	0.05	0.05	0.04	0.03	0.04	0.04	0.03	0.13	0.05
fl	-0.72	-0.72	-0.02	-0.43	-0.44	-0.17	0.06	-0.24	-0.36	-0.48	-0.32	-0.26	-1.19	0.37
	0.04	0.02	0.02	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.07	0.01
ga	-0.98	-0.79	-0.79	-0.58	-0.43	-0.41	-0.13	-0.15	-0.33	-0.29	-0.41	-0.30	-1.93	0.72
	0.05	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.07	0.02
hi	-1.32	-0.64	0.55	0.11	-0.15	-0.12	0.42	0.30	-0.10	-0.74	-0.25	-0.77	-4.09	3.97
	0.09	0.05	0.06	0.07	0.04	0.05	0.08	0.05	0.04	0.04	0.06	0.05	0.10	0.05
il	-0.63	-0.28	-0.79	-0.43	-0.20	-0.32	0.20	-0.06	-0.15	-0.20	-0.20	0.10	-0.25	0.43
	0.04	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.09	0.02
in	-0.49	-0.10	-1.03	-0.11	-0.06	-0.08	0.51	0.32	0.25	-0.09	-0.16	0.32	0.78	0.40
	0.05	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.03	0.03	0.13	0.02
ks	-1.61	-0.06	-0.63	-0.14	-0.07	-0.09	0.09	0.30	0.24	0.06	-0.05	0.06	-1.14	1.34
	0.05	0.03	0.02	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.03	0.10	0.03
ky	-1.29	-0.09	-0.62	-0.08	-0.05	-0.18	0.63	0.31	0.30	0.03	-0.10	0.03	-1.48	0.68
	0.05	0.03	0.02	0.04	0.02	0.03	0.04	0.03	0.02	0.03	0.03	0.03	0.09	0.02
la	-1.86	-0.60	-0.25	-0.28	-0.19	-0.05	0.54	0.20	0.02	-0.24	-0.33	-0.16	-1.40	1.10
	0.04	0.03	0.02	0.03	0.02	0.03	0.05	0.03	0.02	0.03	0.03	0.03	0.09	0.02
md	-0.18	-0.58	-0.01	-0.13	-0.23	0.03	0.20	0.04	-0.41	-0.15	-0.33	-0.69	-1.44	1.23
	0.06	0.02	0.02	0.03	0.02	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.08	0.02
ma	0.58	-0.04	-0.28	0.02	-0.01	0.26	0.26	0.17	-0.07	0.03	0.05	0.25	0.40	0.72
	0.07	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.12	0.02
mi	-0.53	-0.05	-1.03	-0.10	-0.13	0.14	0.53	0.31	0.08	-0.20	-0.17	0.33	1.20	0.35
	0.05	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.14	0.02
mn	-1.12	-0.15	-0.67	-0.16	-0.09	-0.01	0.42	0.10	0.09	-0.04	-0.13	0.42	0.95	0.77
	0.04	0.02	0.02	0.03	0.02	0.03	0.04	0.02	0.02	0.02	0.03	0.03	0.16	0.02
mo	-0.85	-0.23	-0.52	-0.15	-0.14	-0.15	0.28	0.05	0.08	-0.18	-0.24	0.05	-0.86	0.51
	0.04	0.02	0.02	0.03	0.02	0.03	0.03	0.02	0.02	0.02	0.03	0.03	0.10	0.02
nv	-1.28	-1.06	0.01	-0.41	-0.45	-0.44	0.12	-0.36	-0.54	-1.54	-0.32	-0.42	-1.29	2.27
	0.06	0.03	0.04	0.04	0.03	0.04	0.05	0.03	0.03	0.02	0.04	0.04	0.10	0.03
nj	0.30	-0.17	-0.46	-0.43	-0.21	-0.27	-0.01	-0.05	-0.17	-0.04	-0.06	-0.12	0.11	0.98
	0.07	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.10	0.02
nc	-0.92	-0.57	-0.77	-0.26	-0.20	0.05	0.37	0.07	0.00	-0.15	-0.14	0.08	-2.13	0.37
	0.05	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.07	0.02
oh	-0.32	0.02	-0.81	-0.15	-0.07	0.02	0.49	0.16	0.11	-0.11	-0.08	0.25	0.53	-0.08
	0.04	0.02	0.01	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.10	0.02
ok	-1.82	-0.17	-0.33	-0.23	-0.08	-0.10	0.25	0.20	0.16	-0.09	-0.24	-0.20	-1.48	1.13
	0.05	0.03	0.03	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.03	0.03	0.09	0.03
or	-1.78	-0.33	-0.61	-0.30	-0.28	-0.09	0.36	0.14	-0.10	-0.22	-0.18	-0.05	0.68	1.23
	0.05	0.03	0.02	0.03	0.02	0.03	0.04	0.03	0.02	0.03	0.03	0.03	0.17	0.03
pa	-0.49	0.05	-0.47	-0.01	-0.02	0.04	0.55	0.30	0.11	0.04	-0.05	0.27	0.90	-0.14
	0.04	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.12	0.02
sc	-0.51	-0.67	-0.90	-0.32	-0.32	-0.10	0.43	0.07	-0.07	-0.37	-0.35	-0.12	-2.25	1.24
	0.06	0.03	0.02	0.04	0.02	0.03	0.04	0.03	0.02	0.02	0.03	0.03	0.08	0.02
tn	-0.46	-0.39	-0.83	-0.37	-0.25	-0.40	0.33	0.11	0.01	-0.18	-0.32	0.11	-0.19	0.58
	0.05	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.03	0.03	0.11	0.02
tx	-1.76	-0.48	-0.49	-0.39	-0.25	-0.26	0.16	0.02	-0.18	-0.08	-0.28	-0.04	-1.34	-0.09
	0.03	0.02	0.01	0.02	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.02	0.07	0.01
va	-0.90	-0.67	-0.43	-0.10	-0.34	-0.06	-0.04	-0.03	-0.52	-0.25	-0.40	-0.68	-2.72	0.81
	0.05	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.07	0.02
wa	-1.71	-0.41	-0.51	-0.26	-0.21	-0.16	0.04	0.18	-0.19	-0.20	-0.20	-0.22	-1.92	0.91
	0.05	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.03	0.02	0.08	0.02
wi	-1.21	-0.06	-0.93	-0.12	-0.10	0.05	0.53	0.24	0.26	-0.08	-0.02	0.41	0.70	0.58
	0.04	0.03	0.02	0.03	0.02	0.03	0.04	0.02	0.02	0.02	0.03	0.03	0.14	0.02

Tabla A.6 (Cont.) Logit Regressions for the probability of belonging to each state vs. New York: estimated coefficients (standard errors below)

State	Segregation		Factor contributions*									
	Unconditional	Conditional	All	%	Race/ Ethnicity	%	Immigr./ English	%	Education	%	Industry	%
Alabama/al	0.042	0.117	0.075	179.1	0.076	181.5	0.001	3.2	0.000	-0.4	-0.002	-5.2
Arizona/az	0.082	0.083	0.001	0.7	-0.003	-3.2	0.006	6.8	0.000	-0.4	-0.002	-2.5
California/ca	0.099	0.071	-0.029	-29.0	-0.018	-18.2	-0.005	-5.3	-0.002	-2.1	-0.003	-3.4
Colorado/co	0.056	0.083	0.027	48.5	0.012	21.4	0.019	34.2	-0.001	-1.3	-0.003	-6.0
Connecticut/ct	0.053	0.081	0.028	53.0	0.022	42.4	0.006	10.8	0.000	0.1	0.000	-0.2
District of Columbia/dc	0.130	0.156	0.026	19.9	0.007	5.2	0.019	14.2	-0.007	-5.5	0.008	6.0
Florida/fl	0.042	0.049	0.007	16.5	0.003	7.3	0.003	7.5	0.000	-0.2	0.001	1.9
Georgia/ga	0.066	0.083	0.017	25.7	0.028	41.8	-0.007	-10.0	-0.001	-0.9	-0.003	-5.2
Hawaii/hi	0.069	0.114	0.045	65.2	0.032	46.6	0.026	38.2	-0.003	-4.0	-0.011	-15.6
Illinois/il	0.058	0.074	0.015	26.5	0.025	43.7	-0.007	-11.3	-0.003	-4.6	-0.001	-1.2
Indiana/in	0.025	0.123	0.098	384.6	0.078	308.9	0.017	68.8	0.001	2.9	0.001	3.9
Kansas/ks	0.036	0.110	0.073	202.2	0.055	150.6	0.018	50.5	-0.001	-1.4	0.001	2.4
Kentucky/ky	0.023	0.117	0.094	416.6	0.078	344.3	0.015	66.6	0.001	3.4	0.001	2.3
Louisiana/la	0.053	0.107	0.054	102.8	0.064	121.3	-0.008	-15.5	-0.002	-3.3	0.000	0.3
Maryland/md	0.057	0.079	0.021	36.7	0.022	38.0	0.002	3.8	-0.001	-2.2	-0.002	-2.9
Massachusetts/ma	0.042	0.070	0.028	66.9	0.031	74.3	-0.001	-3.4	0.000	0.0	-0.002	-4.0
Michigan/mi	0.028	0.087	0.059	209.5	0.042	149.5	0.014	50.7	0.002	7.4	0.001	1.8
Minnesota/mn	0.029	0.099	0.070	245.8	0.053	185.8	0.018	64.7	-0.001	-2.2	-0.001	-2.5
Missouri/mo	0.024	0.097	0.073	299.1	0.054	219.8	0.018	73.7	0.000	0.3	0.001	5.4
Nevada/nv	0.095	0.090	-0.005	-5.8	-0.012	-12.5	0.014	14.5	-0.002	-1.7	-0.006	-6.1
New Jersey/nj	0.073	0.075	0.001	1.9	0.003	3.9	-0.001	-0.9	0.000	-0.4	-0.001	-0.7
New York/ny	0.053	0.053	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0
North Carolina/nc	0.063	0.098	0.036	56.7	0.038	60.5	0.000	-0.4	-0.001	-1.7	-0.001	-1.6
Ohio/oh	0.020	0.084	0.064	315.3	0.044	218.6	0.017	85.4	0.002	10.8	0.000	0.4
Oklahoma/ok	0.039	0.109	0.070	178.7	0.054	139.9	0.015	39.1	-0.001	-3.8	0.001	3.5
Oregon/or	0.047	0.083	0.035	74.4	0.038	80.7	-0.002	-3.7	0.000	-0.9	-0.001	-1.8
Pennsylvania/pa	0.025	0.074	0.049	198.2	0.041	164.7	0.009	34.5	0.000	1.6	-0.001	-2.7
South carolina/sc	0.063	0.103	0.040	63.3	0.046	72.4	-0.002	-3.4	-0.002	-3.0	-0.002	-2.8
Tennessee/tn	0.034	0.110	0.076	222.6	0.066	192.1	0.008	24.2	0.003	9.9	-0.001	-3.6
Texas/tx	0.078	0.078	0.000	0.4	0.010	13.1	-0.007	-8.4	-0.001	-1.6	-0.002	-2.7
Virginia/va	0.046	0.068	0.021	45.7	0.019	41.6	0.004	8.7	-0.002	-3.3	-0.001	-1.4
Washington/wa	0.049	0.068	0.019	38.5	0.023	46.4	0.000	0.5	-0.001	-2.3	-0.003	-6.2
Wisconsin/wi	0.023	0.100	0.077	335.5	0.061	267.2	0.014	59.8	0.001	4.8	0.001	3.7

Table A7. Conditional segregation: total change and factor contributions (*M* index).

* Values report the contribution of each factor using the Shapley decomposition as the difference between conditional and unconditional segregation induced by each set of characteristics. Percentages indicate factor contributions as proportions of unconditional segregation.